# An introduction to the random matrix approach in number theory

Ashkan Nikeghbali
University of Zürich

Journées Louis Antoine, Rennes, Octobre 2017

In these lectures we aim at giving the reader a first glimpse at the so called random matrix approach in number theory. We also wish to illustrate the use of probabilistic ideas and techniques in analytic number theory, more specifically in understanding the distribution of values of some arithmetic functions which have to do with prime numbers, whose distribution is not random.

In probability theory, the Gaussian distribution plays an ubiquitous role: it is a universal distribution which emerges through the central limi theorem. Hence it is not surprising to see it appear in some of the celebrated limit theorems of analytic number theory, such as the Erdős-Kac central limit theorem for the number of distinct prime divisors of an integer and Selberg's central limit theorem for the logarithm of the Riemann zeta function on the critical line. The first thing we shall do is to put these statements under a more probabilistic looking form by associating with them, in a natural way, a sequence of arithmetically defined random variables or probability measures. But from a probabilistic point of view, these limit theorems do not look so natural, in the sense that they do not fit in a general probabilistic framework from which we could deduce these limit theorems (e.g. the central limit theorem for sums of independent random variables). So one tries to build a simple probabilistic model, in the sense that it is based on some independence assumption concerning the behaviour of prime numbers which we know does not exactly hold, but which we hope is close enough to the reality (let us say asymptotically true), so that it predicts the correct fluctuations. We shall give a couple of standard results which justify the choice of these probabilistic models. But we shall also see how these models fail to capture the true nature of our arithmetic sequences. We believe that this phenomenon is at the heart of the moments conjecture of Keating and Snaith.

The conjecture of Keating and Snaith is central in this lecture: it is a conjecture about the moments of the Riemann zeta function on the critical line (or equivalently about the Fourier transform of the logarithm of zeta) using random matrix theory. The final form of this conjecture is a mixture of two components: one purely arithmetic and one purely group theoretic (random unitary matrices). Why this conjecture, which has been extensively checked numerically and proved in the function field case, is true remains a mystery. We will show that this phenomenon is not an isolated one and that it is shared with other arithmetic objects.

Why this random matrix connection? This brings us to another universal distribution in probability, but a distribution which mathematicians have only recently been able to prove to be a universal law of nature: the sine kernel point process. This remarkable point process emerges asymptotically in the local (or microscopic) distribution of eigenvalues of the Gaussian Unitary Ensemble (GUE) or the Circular Unitary Ensemble (CUE). And many other ensemble of random matrices (this is the remarkable works initiated by Tao and Vu on the one hand, and by Erdős, Schlein, Yau and Yin on the other hand). In 1972, Montgomery conjectured that the zeros of the Riemann zeta function on the critical line also follow this distribution. This seems consistent with a Polya-Hilbert spectral interpretation of the zeros of the Riemann zeta function. Because of the importance of the sine kernel point process emerging from random matrix theory, we shall devote some part of the lecture to prove how it is obtained in random matrix theory. And then we give a totally probabilistic proof of the results by Keating and Sanith, which in fact will give many other new results.

We shall conclude this lecture with a probabilistic attempt to prove Ramachandra's conjecture, which is a statement about the values taken by the Riemann zeta function on the critical line. We shall note that this can be viewed as a local limit type theorem statement. This forces us to develop a new framework which does not rely on sums of (independent) random variable. We are able to prove within this framework an analogue of Ramachandra's conjecture for the corresponding random matrix statistics, the stochastic zeta function. As of today, we are only able to give a conditional proof of Ramachandra's conjecture (but a quantitative version of it).

These notes contain references which are by no means complete. The first Chapter is mostly from works I have done in collaboration with Emmanuel Kowalski and I have also used there some material from Emmanuel Kowalski's lecture notes[1] on probabilistic number theory. In fact, the reader can look at these notes for a classical proof of the Erdős-Kac theorem and for a proof of Selberg's central limit theorem. The random matrix part presentation is part of an ongoing' project with Joseph Najnudel.

[1] https://people.math.ethz.ch/~kowalski/probabilistic-number-theory.pdf

# Contents

CHAPTER 1

# Independence vs dependence

## 1. Two fundamental conjectures

We start by presenting the conjectured relations between random matrices and the Riemann zeta function through two of the most striking conjectures: Montgomery's pair correlation conjecture and Krating-Snaith's moments conjecture. In these notes, we shall give detailed proofs for the random matrix or probabilistic results and refer the reader to the volume [**21**] for more details and references. It was in 1972 that Montgomery conjectured that the distribution of the zeros of the Riemann zeta function on the critical line are similar to those of large random matrices. This then seemed to support the Polya-Hilbert philosophy for a spectral interpretation for the zeros of the Riemann zeta function (although today this has to be more nuanced given all recent results on universality in random matrix theory). In 1980's, A. Odlyzko made extensive numerical computations which seem to support Montgomery's conjecture. Then in the late 1990's, Katz and Sarnak proved in a gigantic work the conjectures in the function field case (i.e. zeta functions over finite fields). And then in 2000, Keating and Snaith used the very brave heuristic that the distribution of values of the Riemann zeta function on the critical line can be modeled by the value distribution of the characteristic polynomial of random unitary matrices on the unit circle in the $n$ limit. They used this "philosophy" to make a guess or conjecture on the moments of the Riemann zeta function on the critical line, a problem on which analytic number theorists have been stuck for about a century. Since this seminal work, many research papers have studied the characteristic polynomial of random unitary matrices and it is now an object of interest in its own, with connections to mathematical physics, Gaussian multiplicative chaos, combinatorics, branching processes, etc. For some of the fundamental results, one can find four or five different proofs using totally different approaches: the most striking one being the case of ratios of factors of characteristic polynomials for which there exist results using representation theory, supersymmetry, classical analysis, probability theory, etc. It should be noted that as one can expect, the problems have also become more complex over time.

The goal of this lecture is to try understand some of these conjectures and connections using probabilistic models and ideas. Hence the focus will be essentially on probabilistic techniques.

**1.1. The Montogomery conjecture.** It has been well known since Euclide that there are infinitely many prime numbers. But how many are there up to a value $x$, or in other words is there an asymptotic for

$$\pi(x) := \#\{p \leqslant x,\ p \text{ prime}\}?$$

THEOREM 1.1 (The prime number theorem (1896)). *We have*

$$\pi(x) \underset{x\to\infty}{\sim} \mathrm{Li}(x) = \int_2^x \frac{dt}{\log t} \sim \frac{x}{\log x}.$$

When $x$ is large, the probability that a number $n$ in the vicinity of $x$ is a prime is approximately $1/\log x$.

There are many open problems on prime numbers: e.g. the twin prime conjecture which states that

$$\#\{n \leqslant x :\ n \text{ and } n+2 \text{ are prime}\} \sim C\frac{x}{\log^2 x}$$

as $x \to \infty$, where

$$C = 2\prod_{p>2}\left(1 - \frac{1}{(p-1)^2}\right) \approx 1.3202\cdots$$

One important object related to the distribution of prime numbers is the Riemann zeta function (see

[28] for more details and proofs). In what follows $s = \sigma + it$ with $\sigma, t \in \mathbb{R}$. The Riemann zeta function is defined as

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \sigma = \mathrm{Re}(s) > 1.$$

The sum is absolutely convergent for $\sigma > 1$ and it is uniformly convergent for $\mathrm{Re}(s) \geq 1 + \delta$ for any $\delta > 0$. It then follows that $\zeta(s)$ is holomorphic in the domain $\sigma > 1$. The function is connected to prime numbers through the following Euler product representation.

THEOREM 1.2 (Euler product formula). *If $\sigma > 1$ we have*

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$$

*where the product runs over all primes p and the product is absolutely convergent.*

PROOF. This is an analytic reformulation of the fundamental theorem of arithmetic which states that every integer $n$ can be decomposed as product of primes in a unique way (up to the order of the factors). □

It is a very classical fact that the Riemann zeta function can be analytically continued to the whole complex plane except at the point $s = 1$ where it has a simple pole, and satisfies a functional equation:

THEOREM 1.3. *The $\zeta$ function has an analytic continuation to $\mathbb{C}$ except for a simple pole at $s = 1$ with residue* 1*. Moreover,*

$$\pi^{s/2}\zeta(s)\Gamma\left(\frac{s}{2}\right) = \pi^{(1-)s/2}\zeta(1-s)\Gamma\left(\frac{1-s}{2}\right)$$

*and $\zeta(-2k) = 0$ for $k \geq 1$ and $k \in \mathbb{N}$.*

It is sometimes more convenient to introduce the Riemann $\xi$ function which is an entire function:

$$\xi(s) = \frac{1}{2}s(s-1)\pi^{s/2}\zeta(s)\Gamma\left(\frac{s}{2}\right)$$

which is entire in $\mathbb{C}$. The functional equation gives

$$\xi(s) = \xi(1-s).$$

We already know that $\zeta(s) \neq 0$ for $\sigma > 1$ by the Euler product formula, and that $\Gamma(s) \neq 0$ for all $s$; thus $\xi(s)$ has no zeros when $\sigma > 1$. Moreover, the functional equation implies that $\xi(s)$ has no zeros for $\sigma < 0$. The zeros of $\xi$ are all in the strip $0 \leq \sigma \leq 1$ and they are the same as the zeros of $\zeta(s)$ in that strip.
We may also note that if $\rho$ is a zero of $\xi(s)$ then $1 - \rho$ is also a zero of $\xi(s)$. Moreover, since $\overline{\xi(s)} = \xi(\bar{s})$ then $\bar{\rho}$ and $1 - \bar{\rho}$ are also zeros of $\xi(s)$. The zeros are thus symmetrically arranged about the line $\sigma = 1/2$ and the real axis.

The Riemann hypothesis is the statement that $\sigma = \frac{1}{2}$ for all the non-trivial zeros of $\xi(s)$. This conjecture is of central importance in mathematics for understanding the distribution of the zeros of the Riemann zeta function amounts to understanding the distribution of prime numbers: there are many formulas relating the zeros to prime numbers (they are called explicit formulas in analytic number theory).

Using the theory of entire functions of finite order, one can prove the following Hadammard factorization for $\xi(s)$:

$$\xi(s) = e^{-Bs} \prod_\rho \left(1 - \frac{s}{\rho}\right) e^{s/\rho} \tag{1}$$

where $\rho = \beta + i\gamma$ runs over the zeros of $\zeta(s)$ in the strip $0 \leq \beta \leq 1$.

With the argument principle we can prove the following theorem. Let $N(T) = \#\{\rho = \beta + i\gamma,\ 0 \leqslant \beta \leqslant 1,\ 0 \leqslant \gamma \leqslant T\}$, i.e. $N(T)$ denotes the number of zeros $\rho$ in the rectangle whose base is the line $[0, 1]$ and whose height is $T$.

THEOREM 1.4 (Riemann-von Mangoldt estimate). *If $T$ is not an ordinate of a zero of $\zeta(s)$, then*

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T).$$

The following conjecture about the distribution of the zeros of the Riemann zeta function on the critical line is at the heart of the connection between number theory and random matrix theory.

THEOREM 1.5 (Montgomery, 1972). *We assume that the Riemann hypothesis is satisfied. We note $\frac{1}{2} + it_n$ then $n$-th zero of the zeta function with $t_n \geq 0$. We define $w_n$ as*

$$w_n = \frac{t_n}{2\pi}\log\frac{t_n}{2\pi}.$$

*Then for $f \in \mathscr{S}(\mathbb{R})$ such that $\operatorname{supp}\hat{f} \subset [-1,1]$ one has*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{1\leqslant m\neq n\leqslant N} f(\omega_m - \omega_n) = \int_{-\infty}^{\infty} f(x)R_2(x)dx,$$

*where*

$$R_2(x) = 1 - \left(\frac{\sin(\pi x)}{\pi x}\right)^2.$$

CONJECTURE 1.6 (Montgomery, 1972). *The above holds without any condition on the support of the Fourier transform of the test function $f$.*

The connection with random matrix theory occurred to Montgomery after discussing with Dyson. Indeed Dyson had considered similar computations in random matrix theory. We state quickly basic facts we need to illustrate the connection but we shall come back later with more rigour and details on these random matrix aspects.

The random matrix model Dyson considered is the unitary group $U(N)$ equipped with the Haar measure. The Haar measure on $U(N)$ induces a probability distribution for the angles of the eigenvalues, called eigenangles, as follows: $U \in U(N)$ chosen according to the Haar measure has $N$ eigenvalues $(e^{i\theta_1}, \cdots, e^{i\theta_N})$ and the joint probability density of $(\theta_1, \cdots, \theta_N)$ on $[0, 2\pi)^N$ (or alternatively on $(-\pi, \pi]$) is

$$p(\theta_1, \cdots, \theta_N) = \frac{1}{(2\pi)^N N!}\prod_{1\leqslant j<k\leqslant N}\left|e^{i\theta_j} - e^{i\theta_k}\right|^2.$$

Let $f$ be a function on $U(N)$ such that $\forall U \in U(N)$ and $A \in U(N)$,

$$f(U^*AU) = f(A);$$

we then have

$$\begin{aligned}\int_{U(N)} f(A)d\mu_{\text{Haar}} &\equiv \mathbb{E}_N[f(A)]\\ &= \int_{[0,2\pi]^N} f(e^{i\theta_1}, \cdots, ^{i\theta_N})p(\theta_1, \cdots, \theta_N)d\theta_1\cdots d\theta_N.\end{aligned}$$

For instance, for $U \in U(N)$ with eigenvalues denoted by $e^{i\theta_1}, \cdots, e^{i\theta_N}$ we could define

$$f(e^{i\theta_1}, \cdots, ^{i\theta_N}) = \prod_{k=1}^{N}(1 - e^{i\theta_k}).$$

For $f$ a suitable test function,

$$\lim_{N\to\infty}\mathbb{E}_N\left[\frac{1}{N}\sum_{1\leqslant m\neq n\leqslant N} f\left((\theta_m - \theta_n)\frac{N}{2\pi}\right)\right] = \int_{-\infty}^{\infty} f(x)R_2(x)dx,$$

where

$$R_2(x) = 1 - \left(\frac{\sin(\pi x)}{\pi x}\right)^2.$$

Hence it seems that the so called pair correlation statistics coincide for both the rescaled zeros of the Riemann zeta function and for the rescaled eigenvalues of random unitary matrices. The so called GUE conjecture for the Riemann zeta function states that in fact the point process consisting of the rescaled zeros of the Riemann zeta function and the rescaled eigenangles have the same correlation functions of all orders and hence have the same distribution. We shall give a precise meaning to the notion of correlation functions in the next chapter.

Is there a deep hidden picture which could explain this striking analogy? In light of recent results on universality, one might be less enthusiastic than a couple of decades ago. Indeed it has been observed that for many models of random matrices, the eigenvalues have a limiting short-scale behavior when the dimension goes to infinity which depends on the global symmetries of the model, but not on its detailed features. For example, the Gaussian Orthogonal Ensemble (GOE), for which the matrices are real symmetric with independent gaussian entries on and above the diagonal, corresponds to a limiting short-scale behavior for the eigenvalues that is also obtained for several other models of random real symmetric matrices. Similarly, the limiting spectral behavior of a large class of random hermitian and unitary ensembles, including the Gaussian Unitary Ensemble (GUE, with independent, complex gaussians above the diagonal), and the Circular Unitary Ensemble (CUE, corresponding to the Haar measure on the unitary group of a given dimension), involves a remarkable random point process, called the *determinantal sine-kernel process*. It is a point process for which the $k$-point correlation function is given by (see next chapter for more details)

$$\rho_k(x_1,\dots,x_k) = \det\left(\frac{\sin(\pi(x_p-x_q))}{\pi(x_p-x_q)}\right)_{1\leq p,q\leq k}.$$

Montgomery's conjecture states that the limiting short-scale behavior of the imaginary parts of the zeros of the Riemann zeta function is also described by a determinantal sine-kernel process. On the one hand this can be disappointing since many other systems have this feature but on the other hand this similar behavior supports the conjecture of Hilbert and Pólya, who suggested that the non-trivial zeros of the Riemann zeta functions should be interpreted as the spectrum of an operator $\frac{1}{2}+iH$ with $H$ an unbounded Hermitian operator.

In the following section we are going to see that the connection can still be deepened.

**1.2. The moments conjecture and the characteristic polynomial.** A major breakthrough in the so-called random matrix approach in number theory is the seminal paper of Keating and Snaith [**14**], where they conjecture that the characteristic polynomial of a random unitary matrix, restricted to the unit circle, is a good and accurate model to predict the value distribution of the Riemann zeta function on the critical line. In particular, using this philosophy, they were able to conjecture the exact asymptotics of the moments of the Riemann zeta function, a result which was considered to be out of reach with classical tools from analytic number theory. One simple and naive explanation for the success of the characteristic polynomial as a random model to the Riemann zeta function comes from Montgomery's conjecture that asserts that the zeros of the Riemann zeta function on the critical line (after rescaling) statistically behave like the eigenangles after rescaling (and hence the zeros of the characteristic polynomial) of large random unitary matrices.

More precisely the characteristic polynomial is usually defined in the following way: for $U\in U(n)$, define $Z_n$ as:

$$Z_n(X) = \det\left(\mathrm{Id}-U_n^{-1}X\right) = \det\left(\mathrm{Id}-U_n^{*}X\right). \tag{2}$$

It is not hard to see that

$$Z_n(X) = (-X)^n\det(U_n^*)\overline{Z_n(\frac{1}{\overline{X}})}.$$

Note that some authors take the convention

$$Z_n(X) = \det\left(\mathrm{Id}-U_nX\right),$$

in which case

$$Z_n(X) = (-X)^n\det(U_n)\overline{Z_n(\frac{1}{X})}.$$

No matter the convention, the important common feature is that the zeros of $Z_n$ are on the unit circle and that the unit circle plays the role of the critical line.

Now we want to illustrate the way the Keating-Snaith philosophy is used to make predictions for the distribution of the values of the Riemann zeta function.

So first what are typical value distribution problems for the Riemann zeta function? The first probabilistic result is due to Selberg: if $U$ is a uniform random variable in $[0,1]$, then[1]

$$\frac{\log\zeta(\frac{1}{2}+iTU)}{\sqrt{\frac{1}{2}\log\log T}}\overset{\text{law}}{\to}\mathcal{N}_{\mathbb{C}},$$

as $n\to\infty$, and with $\mathcal{N}_{\mathbb{C}}=\mathcal{N}_1+i\mathcal{N}_2$ where $\mathcal{N}_1,\mathcal{N}_2$ denote two standard Gaussian distributions. In other words the log of the Riemann zeta function behaves like a complex Gaussian distribution, with very slowly growing variance. We do not provide a proof of this limi theorem but the way it is proven is through the method of moments and more precisely through the following estimate:

THEOREM 1.7. *If $n$ is a positive integer, $0<a<1$ and $T^{a/n}\leq x<T^{1/n}$, then there exists a constant $C=C_{n,a}$ such that for all sufficiently large $T$*

$$\frac{1}{T}\int_T^{2T}\left|\log\zeta(\tfrac{1}{2}+it)-\sum_{p\leqslant x}\frac{p^{-it}}{\sqrt{p}}\right|^{2n}dt\leqslant C.$$

Here, following a classical heuristic in analytic number theory (see next section), $\frac{p^{-it}}{\sqrt{p}}$ can be thought of as $\frac{X_p}{\sqrt{p}}$ where $X_p$ is uniformly distributed on the unit circle and the $(X_p)_p$ are independent and it follows from a very classical number theory estimate that

$$\operatorname{var}\left(\sum_{p\leqslant x}\frac{X_p}{\sqrt{p}}\right)=\sum_{p\leqslant x}\frac{1}{p}\sim\log\log x.$$

A seemingly related problem but which in fact is much harder and still open is the Ramachandra conjecture:

CONJECTURE 1.8 (Ramachandra). *We have:*

$$\overline{\{\zeta(1/2+it),\quad t\in\mathbb{R}\}}=\mathbb{C}.$$

We shall give a few results in this direction in relation with random matrix theory and the moments conjecture that we now discuss.

Another open problem concerns the size or growth of $\zeta(\frac{1}{2}+it)$.

CONJECTURE 1.9 (Lindelöf).

$$\zeta\left(\frac{1}{2}+it\right)=O(t^{\varepsilon}),\quad\forall\varepsilon>0.$$

It is known that the Riemann hypothesis implies the Lindelöf hypothesis. One way to attack this conjecture is through the moments.

THEOREM 1.10. *The Lindelöf hypothesis is equivalent to*

$$\frac{1}{T}\int_0^T\left|\zeta\left(\frac{1}{2}+it\right)\right|^{2k}dt=O(T^{\varepsilon}),\quad\forall\varepsilon>0,\ k=1,2,3,\cdots.$$

PROOF. A proof can be found in Titchmarsh' book **[28]**, chapter 13. □

The above explains why it is important to understand the moments and since Hardy and Littlewood in 1918, number theorists have tried to estimate these moments. But as of today, only the cases $k=1$ (1918) and $k=2$ (1923) are known. In the 1990's, Brian Conrey and his collaborators came with the following conjecture:

CONJECTURE 1.11. *For any positive integer $k$*

$$\int_0^T\left|\zeta\left(\frac{1}{2}+it\right)\right|^{2k}dt\sim a_kg_kT(\log T)^{k^2},\tag{3}$$

[1] $\log\zeta(s)$ is be defined as the unique version of the logarithm of zeta which is real on $(1,\infty)$, well-defined and continuous everywhere, except on the closed half-lines at the left of the zeros and the pole at 1 of zeta.

*where $a_k$ is the arithmetic factor*

$$a_k := \prod_p \left(1-\frac{1}{p}\right)^{k^2} {}_2F_1\left(k,k,1;\frac{1}{p}\right), \tag{4}$$

*and $g_k$ is an unknown factor for which very little is known:*

(1) $g(1) = 1/1!$ *by Hardy and Littlewood.*
(2) $g(2) = 2/(2!)^2$ *by Ingham.*

*It was also conjectured by Conrey and collaborators that $g(3) = 42/(3!)^2$, and $g(4) = 24024/(4!)^2$.*

Keating and Snaith made the guess that the unknown factor $g_k/k^2!$ should be predicted by random matrix theory, and more precisely by the moments of the characteristic polynomial. They were able to compute

$$M_N(\lambda) = \int_{U(N)} \left|Z_N(U)(e^{i\theta})\right|^{2\lambda} d\mu_{\text{Haar}}, \quad \text{Re}(\lambda) > -\frac{1}{2}$$

and managed to prove that

$$M_N(\lambda) = \prod_{j=1}^N \frac{\Gamma(j)\Gamma(2\lambda+j)}{\Gamma^2(j+\lambda)} \sim \frac{G^2(1+\lambda)}{G(1+2\lambda)} N^{\lambda^2}$$

as $N \to \infty$. Then they made the analogy $N \leftrightarrow \log T$ to make the following full conjecture:

CONJECTURE 1.12 (Moment conjecture, Keating-Snaith). *It is conjectured that:*

$$\frac{1}{T}\int_0^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^{2\lambda} dt \sim a(\lambda)g(\lambda)(\log T)^{\lambda^2}$$

*where*

$$g(\lambda) = \frac{G^2(1+\lambda)}{G(1+2\lambda)}, \quad \text{Re}(\lambda) > -\frac{1}{2}$$

*with $G(1) = 1$ and $G(z+1) = \Gamma(z)G(z)$. The function $G$ is known as the Barnes function.*

Moreover from their moment computation Keating and Snaith could deduce the following analogue of Selberg's central limit theorem for $\theta \in [0, 2\pi]$:

$$\frac{\log Z_N(e^{i\theta})}{\sqrt{\frac{1}{2}\log N}} \overset{\text{law}}{\to} \mathcal{N}_{\mathbb{C}}.$$

The methods they used to compute the moments, which is the main result on the random matrix side, is the Selberg integrals. Indeed they write

$$Z_N := \prod_{k=1}^N (1-e^{-i\theta_k}),$$

where $\theta_1, \cdots, \theta_2$ are the eigenvalues of $U$. For $s \in \mathbb{C}$ with $\text{Re}(s) > 0$, (in fact $\text{Re}(s) > -1$), Weyls' integration formula gives

$$E[|Z_N|^s] = \frac{1}{(2\pi)^n n!}\int_{[0,2\pi]^n} \left|\prod_{k=1}^n (1-e^{-i\theta_k})\right|^s \prod_{1 \leqslant j < h \leqslant n} \left|e^{i\theta_j} - e^{i\theta_h}\right|^2 d\theta_1 \cdots d\theta_n.$$

Then they use the Selberg integrals

$$\begin{aligned} J(a,b,\alpha,\beta,\gamma,n) &= \int_{\mathbb{R}^n} \prod_{1 \leqslant j < l \leqslant n} \left|x_j - x_l\right|^{2\gamma} \prod_{j=1}^n (a+ix_j)^{-\alpha}(b-ix_j)^{-\beta} dx_1 \cdots dx_n \\ &= \frac{(2\pi)^n}{(a+b)^{(\alpha+\beta)n-\gamma_n(n-1)-n}} \\ &\times \prod_{j=0}^n \frac{\Gamma(1+\gamma+j\gamma)\Gamma(\alpha+\beta-(n+j-\gamma)\gamma-1)}{\Gamma(1+j)\Gamma(\alpha+j\gamma)\Gamma(\beta-j\gamma)} \end{aligned}$$

in the special case $a = b = \gamma = 1$.

We shall propose in this lecture a totally different alternative approach to this result which does not use Weyl's integration formula. But for now let us explore a little more the computations by Keating and Snaith:

$$\mathbb{E}[|Z_N|^{2\lambda}] = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+2\lambda)}{\Gamma^2(j+\lambda)}, \quad \mathrm{Re}(\lambda) > -\frac{1}{2}.$$

Recall that the Barnes $G$-function is an entire function satisfying $G(1) = 1$ and

$$G(z+1) = \Gamma(z)G(z) \quad \forall z \in \mathbb{C},$$

and the zeros and located at the negative integers. We also have

$$G(1+z) = (2\pi)^{z/2} e^{-z(z+1)/2} \prod_{k=1}^{\infty} \left\{ \left(1+\frac{z}{k}\right)^{k} \left(1+\frac{z}{k}\right)^{z^2/2} e^{-z} \right\}.$$

In particular for $z \in \mathbb{C}\backslash\mathbb{Z}-$ we have

$$\prod_{j=1}^{N} \Gamma(j+\theta) = \frac{G(1+N+\theta)}{G(1+\theta)}, \quad \forall \theta \in \mathbb{C}\backslash\mathbb{Z}- .$$

This implies the following for the expectation of $|Z_N|^{2\lambda}$

$$\mathbb{E}[|Z_N|^{2\lambda}] = \frac{G(1+N)G(1+N+2\lambda)G^2(1+\lambda)}{G(1+2\lambda)G^2(1+N+\lambda)}.$$

It is known (for instance **[11]**) that uniformly on compact sets of $\{(\gamma,\delta) \in \mathbb{C}\times\mathbb{C},\ \gamma+\delta \notin \mathbb{Z}_-\}$

$$\frac{G(1+N+\gamma+\delta)G(1+N)}{G(1+N+\delta)G(1+N+\gamma)} = (1+N)^{\gamma\delta}\left(1+O\left(\frac{1}{N}\right)\right).$$

In our case $\gamma = \delta = \lambda$ we have

$$\frac{G(1+N)G(1+2\lambda+N)}{G^2(1+N+\lambda)} = (1+N)^{\lambda^2}\left(1+O\left(\frac{1}{N}\right)\right).$$

This implies that

$$\lim_{N\to\infty} \frac{1}{N^{\lambda^2}} \mathbb{E}[|Z_N|^{2\lambda}] = \frac{G^2(1+\lambda)}{G(1+2\lambda)}.$$

With the notation

$$g(k) = \frac{G^2(1+\lambda)}{G(1+2\lambda)},$$

we have $\forall k \in \mathbb{N}$

$$\begin{aligned} g(\lambda) &= \frac{(\Gamma(k)\Gamma(k-1)\cdots\Gamma(1))^2}{\Gamma(2k)\Gamma(2k-1)\cdots\Gamma(1)}, \\ &= \prod_{j=0}^{k-1} \frac{j!}{(j+k)!}. \end{aligned} \tag{5}$$

Now let us try to explain at least heuristically the appearance of the arithmetic factor. If in the moments conjecture we take $\lambda = iu$ where $u \in \mathbb{R}$, we can rephrase (3) in terms of Fourier transforms

$$\underbrace{(\log T)^{u^2}}_{=\exp[u^2 \log\log T]} \underbrace{\frac{1}{T}\int_0^T \exp[2iu \log\left|\zeta(\tfrac{1}{2}+it)\right|]dt}_{=\mathbb{E}[\exp(2iu\log\left|\zeta(\frac{1}{2}+iTU)\right|)]} \sim a(iu)g(iu)$$

as $T \to \infty$ and where $U$ is a uniform random variable on $[0,1]$. We note

$$a(iu) = \lim_{N\to\infty} a_1(u,N)a_2(u,N)$$

where

$$a_1(u,N) = \prod_{p\leqslant N} (1-p^{-1})^{-u^2}, \quad a_2(u,N) = \prod_{p\leqslant N} {}_2F_1(iu,iu,1;p^{-1}).$$

We then use the following formula of Mertens:

$$\prod_{p\leqslant N}(1-p^{-1}) \underset{N\to\infty}{\sim} e^{-\gamma}(\log N)^{-1}$$

where $\gamma$ is Euler's constant. Now, if we note $\gamma_N = 2(\gamma+\log\log N)$, $N \geq 2$,

$$\lim_{N\to\infty} e^{u^2\gamma_N/2}a_2(u,N) = a(iu).$$

Now let $X$ be a random variable which is uniformly distributed on the unit circle, and let $x \in \mathbb{R}$ with $x > 1$. We then have

$$\mathbb{E}\left[\exp\left(-2iu\log\left|1-\frac{X}{\sqrt{x}}\right|\right)\right] = \mathbb{E}\left[\left|1-\frac{X}{\sqrt{x}}\right|^{-2iu}\right] = {}_2F_1(iu,iu,1;x^{-1}).$$

Indeed,

$$\left|1-\frac{X}{\sqrt{x}}\right|^2 = 1+\frac{1}{x}-\frac{2\operatorname{Re}(X)}{\sqrt{x}} \geqslant (1-x^{-1/2})^2 > 0.$$

Next, $\operatorname{Re}(X) = \cos(\theta)$, where $\theta$ is uniform on $[0,2\pi]$

$$\begin{aligned}\mathbb{E}[e^{iu\log|1-X/\sqrt{x}|^{-2}}] &= \frac{1}{2\pi}\int_0^{2\pi}(1+x^{-1}-2x^{-1/2}\cos\theta)^{-iu}d\theta \\ &= {}_2F_1(iu,iu,1;x^{-1}),\end{aligned}$$

this comes from the integral representation of ${}_2F_1$. Now, write

$$Y_N = \sum_{p\leqslant N}\log\left(\left|1-\frac{X_p}{\sqrt{p}}\right|^{-2}\right)$$

then it follows from the previous calculations that

$$e^{u^2\gamma_N/2}\mathbb{E}[e^{iuY_N}] = a(iu).$$

Here, $(p^{it})$ were assumed to behave like independent random variables which are uniformly distributed on the unit circle, but we know this is not true. The random matrix factor seems to be a correction factor accounting for the dependency that this naive model ignores. And this phenomenon is not specific to the Riemann zeta function.

## 2. The splitting phenomenon

In some sense, probabilistic number theory can be viewed as the study of fluctuations (e.g. central limit theorem, local limit theorem, large deviations, etc.) of arithmetically defined sequences of random variables or probability measures. One of the most famous results in this direction is Erdős-Kac theorem:

THEOREM 2.1 (Erdős-Kac). *For any positive integer $n \geq 1$, let $\omega(n)$ denote the number of distinct prime divisors of $n$. Then for any real numbers $a < b$, we have*

$$\lim_{N\to\infty}\frac{1}{N}|\{1 \leq n \leq N : a \leq \frac{\omega(n)-\log\log N}{\sqrt{\log\log N}} \leq b\}| = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}dx.$$

One can easily interpret this as a convergence in law for a sequence of random variables which are arithmetically defined. Indeed, consider $\Omega_N = \{1,\cdots,N\}$ with the uniform measure $\mathbb{P}_N$ on it, and define the random variable $X_N$ from $\Omega_N$ to $\mathbb{R}$ by

$$X_N = \frac{\omega(n)-\log\log N}{\sqrt{\log\log N}}.$$

Then the Erdős-Kac theorem is the statement that $X_N$ converges in distribution to a standard Gaussian random variable $\mathcal{N}$.

It is very often possible to predict the behaviour of many such arithmetic variables or measures with the help of simple probabilistic models. These probabilistic models are based on two relatively simple results that we present now.

THEOREM 2.2. *Let $N \geq 1$ and $\Omega_N = \{1, \cdots, N\}$ endowed with the uniform probability measure $\mathbb{P}_N$. Fix an an integer $q \geq 1$ and denote by $\pi_q : \mathbb{Z} \longrightarrow \mathbb{Z}/q\mathbb{Z}$ the canonical surjection which to $n$ associates its class modulo $q$. Let $X_N$ be the random variable which is defined on $\Omega_N$ by $X_N(n) = \pi_q(n)$ and with values in $\mathbb{Z}/q\mathbb{Z}$. Then the random variables $(X_N)$ converge in law to the uniform probability measure $\mu_q$ on $\mathbb{Z}/q\mathbb{Z}$. More precisely, for any function $f : \mathbb{Z}/q\mathbb{Z} \longrightarrow \mathbb{C}$, we have*

$$|\mathbb{E}[f(X_N)] - \mathbb{E}[f]| \leq \frac{2}{N}\|f\|_1 \tag{6}$$

*where*

$$\|f\|_1 = \sum_{a \in \mathbb{Z}/q\mathbb{Z}} |f(a)|.$$

PROOF. We have by definition

$$\mathbb{E}[f(X_N)] = \frac{1}{N} \sum_{1 \leq n \leq N} f(\pi_q(n))$$

and

$$\mathbb{E}[f] = \frac{1}{q} \sum_{a \in \mathbb{Z}/q\mathbb{Z}} f(a).$$

Next we note that

$$\frac{1}{N} \sum_{1 \leq n \leq N} f(\pi_q(n)) = \sum_{a \in \mathbb{Z}/q\mathbb{Z}} f(a) \times \frac{1}{N} \sum_{\substack{1 \leq n \leq N \\ n \equiv a (modq)}} 1.$$

Hence the last sum counts the number of integers $m$ for which we have $1 \leq mq + a \leq N$, which is $[\frac{N-a}{q}]$. Now the trivial estimate $x - 1 \leq [x] \leq x$ for $x \geq 0$ yields the desired bound

$$|\frac{1}{N} \sum_{1 \leq n \leq N} f(\pi_q(n)) - \frac{1}{q} \sum_{a \in \mathbb{Z}/q\mathbb{Z}} f(a)| \leq \frac{2}{N}\|f\|_1.$$

□

In the special case where $f(n) = \mathbf{1}_{\{a\}}(\pi_q(n))$, we obtain

$$\left| \mathbb{P}_N(\pi_q(n) = a) - \frac{1}{q} \right| \leq \frac{2}{N}.$$

This is a non-trivial inequality as long as $q$ is small enough compared with $N$. In particular, this shows that for $q$ fixed, the probability that an integer $n \leq N$ is congruent to $a$ modulo $q$ is approximately $1/N$. This asymptotic estimate justifies the commonly used heuristic that the probability that an integer is divisible by a 2 or 3 is approximately $1/2$ or $1/3$. Hence one could start to think of a model of the form $\sum_{p \leq N} B_p$, where the sum runs over prime numbers and where $B_p$ is a Bernoulli random variable defined as $\mathbb{P}[B_p = 1] = \frac{1}{p} = 1 - \mathbb{P}[B_p = 0]$. Next we would like to know whether it is reasonable to assume that these Bernoulli random variables could be taken to be independent. This is indeed asymptotically true as a consequence of the Chinese remainder theorem. Indeed it is a well known fact that if $q_1$ and $q_2$ are two positive coprime integers, then the map $\mathbb{Z}/q_1q_2\mathbb{Z} \longrightarrow \mathbb{Z}/q_1\mathbb{Z} \times \mathbb{Z}/q_2\mathbb{Z}$ $x \longmapsto (x(\text{mod } q_1), x(\text{mod } q_2))$ is a ring isomorphism. This ring isomorphism can be interpreted as the fact that the random variables $\pi_{q_1}$ and $\pi_{q_2}$ on $\Omega_N$ are asymptotically independent, that is

$$\lim_{N \to \infty} \mathbb{P}_N(\pi_{q_1}(n = a), \pi_{q_2}(n) = b) = \frac{1}{q_1q_2} = (\lim_{N \to \infty} \mathbb{P}_N(\pi_{q_1}(n) = a))(\lim_{N \to \infty} \mathbb{P}_N(\pi_{q_2}(n) = b)).$$

So it is reasonable to build a probabilistic model by assuming that divisibility by distinct primes are independent events. We can summarize this in the following proposition:

PROPOSITION 2.3. *Let $N \geq 1$ and $\Omega_N = \{1, \cdots, N\}$ endowed with the uniform probability measure $\mathbb{P}_N$. Let $k \geq 1$ be an integer and fix $q_1 \geq 1, \cdots, q_k \geq 1$ a family of coprime integers. As $N \to \infty$, the random vector*

$$(\pi_{q_1}, \cdots, \pi_{q_k}) : \; n \longmapsto (\pi_{q_1}(n), \cdots, \pi_{q_k}(n))$$

*from $\Omega_N$ with values in $\mathbb{Z}/q_1\mathbb{Z}\cdots\times\mathbb{Z}/q_k\mathbb{Z}$ converges in law to the product of the uniform probability measures $\mu_{q_i}$, or in other words converges to a random vector whose components are independent and uniformly distributed on $\mathbb{Z}/q_i\mathbb{Z}$, for $1\le i\le k$. We also have a quantitative statement as is Theorem 2.2: for any function*

$$f:\ \mathbb{Z}/q_1\mathbb{Z}\cdots\times\mathbb{Z}/q_k\mathbb{Z}\longrightarrow\mathbb{C},$$

*we have*

$$\left|\mathbb{E}[f(\pi_{q_1}(n),\cdots,\pi_{q_1}(n))]-\mathbb{E}[f]\right|\le\frac{2}{N}\|f\|_1. \tag{7}$$

PROOF. The proof follows the same lines as the proof of Theorem 2.2 after noting that the Chinese remainder theorem gives a ring isomorphism $\mathbb{Z}/q_1\mathbb{Z}\cdots\times\mathbb{Z}/q_k\mathbb{Z}\longrightarrow\mathbb{Z}/q\mathbb{Z}$ where $q=q_1\cdots q$ under which the images of the product of the uniform measures $\mu_{q_i}$ is $\mu_q$. □

REMARK 2.4. *It is important to note that the random variables $\pi_{q_1}$ and $\pi_{q_2}$ are not independent: it is not true that*

$$\mathbb{P}_N(\pi_{q_1}(n=a),\pi_{q_2}(n)=b)=\mathbb{P}_N(\pi_{q_1}(n)=a)\mathbb{P}_N(\pi_{q_2}(n)=b).$$

*Independence only holds asymptotically. This is the source of many subtle behaviours in analytic number theory.*

Now going back to our probabilistic model, we may assume that the Bernoulli variables $B_p$ are independent. Consequently $Z_N=\sum_{p\le N}B_p$ can be thought of as a probabilistic model for $\omega(n)$ (more precisely for the arithmetically defined random variable $X_N(n)=\omega(n)$ defined on $\Omega_N$). Using Mertens' estimates that $\sum_{p\le N}\frac{1}{p}\sim\log\log N$, we have $\mathbb{E}[Z_N]=\sum_{p\le N}\frac{1}{p}\sim\log\log N$ and $var(Z_N)=\sum_{p\le N}(p(1-p))^{-1}\sim\log\log N$. An application of the central limit theorem for independent random variables yields that $\frac{Z_N-\log\log N}{\sqrt{\log\log N}}$ converges in distribution to a standard Gaussian random variable. In fact one can use this probabilistic model together with quantitative estimates as above to show that the moments of are close to those of $\frac{X_N-\log\log N}{\sqrt{\log\log N}}$ and $\frac{Z_N-\log\log N}{\sqrt{\log\log N}}$ are asymptotically the same. But our point here is different and we would rather illustrate the limitations of the probabilistic model and see where the independence assumption in the probabilistic model can go wrong when compared to the true arithmetic sequence of random variables. We shall use for this a formula of Rényi and Turán:

PROPOSITION 2.5 (Rényi and Turán). *We have the following asymptotic formula:*

$$\mathbb{E}_N[e^{itX_N}]=(\log N)^{e^{it}-1}\left(\Phi(t)+o(1)\right), \tag{8}$$

*uniformly for $t\in\mathbb{R}$ and as $N\to\infty$, with the factor $\Phi(t)$ given by*

$$\Phi(t)=\frac{1}{\Gamma(e^{it})}\prod_p\left(1-\frac{1}{p}\right)^{e^{it}}\left(1+\frac{e^{it}}{p-1}\right),$$

*where the Euler product is absolutely convergent.*

The proof of the Rényi-Turán formula is much more difficult than the moments computation performed to prove the Erdős-Kac central limit theorem and the content is actually much deeper (the value $t=\pi$ implies the prime number theorem). Formula (8) follows from the following deeper theorem whose proof is based on the Selberg-Delange method applied to the Dirichlet series of $y^{\omega(n)}$

$$\sum_{n\ge1}\frac{y^{\omega(n)}}{n^s}=\prod_p\left(1+\frac{y}{p^s-1}\right).$$

PROPOSITION 2.6 ([27], Section II.6, Theorem 1). *For any $A>0$, we have for any $y\in\mathbb{C}$, with $|y|\le A$*

$$\sum_{k\le n}y^{\omega(k)}=n(\log n)^{y-1}(\lambda_0(y)+O(1/\log n)),$$

*where*

$$\lambda_0(y)=\frac{1}{\Gamma(y)}\prod_p\left(1-\frac{1}{p}\right)^y\left(1+\frac{y}{p-1}\right)$$

*and the constant in the $O$ symbol only depends on $A$.*

The following discussion is based on [19]. In formula (8) one recognizes $(\log N)^{e^{it}-1}$ as the characteristic function of a Poisson random variable with parameter $\log\log N$. Since $\Phi(t/\sqrt{\log\log N}) \to \Phi(0) = 1$, the Erdős-Kac central limit theorem immediately follows from an application of Lévy's criterion for convergence in law. In fact this type of convergence is called mod-Poisson convergence and there is now a well built general theory (of mod-phi convergence) from which one can deduce precise large deviations estimates, speed of convergence, approximation in the total variation distance (see e.g. [12], [2] and [6]) but we shall not touch upon this here. We would rather try to compare this with what the probabilistic model $Z_N$ predicts. Indeed, a straightforward computation using the independence of the Bernoulli variables shows that

$$\mathbb{E}[e^{itZ_N}] = (\log N)^{e^{it}-1} \prod_p \left(1 - \frac{1}{p}\right)^{e^{it}} \left(1 + \frac{e^{it}}{p-1}\right).$$

Consequently the probabilistic model fails to capture the mod-Poisson phenomenon. This is coming from the fact that the primes do not actually behave independently of each other and that the independence assumption, while enough to predict the central limit theorem, fails to account for the real nature of $\omega(n)$. At the level of non nomalized characteristic functions, extreme values play a more crucial role.

Now one may ask about an interpretation for the factor $\frac{1}{\Gamma(e^{it})}$. A natural explanation is that this is a correction factor to the independence assumption made in the probabilistic model. But does this factor correspond to some natural "model" as well? It is surprising and mysterious that the answer is positive. Indeed, consider the symmetric group of order $N$, $\mathcal{S}_N$, endowed with the uniform measure, and define on it the random variable $\ell_N$ that maps a permutation $\sigma$ to the number of cycles in its cyclic decompositions. Then it is well known in probability theory (see [1]) that $\ell_N$ has the same distribution as $\sum_{k \leq N} B_k$ where the $B_k$'s are independent Bernoulli random variables with parameters $1/k$. Using the product representation of $1/\Gamma(z)$, one obtains that

$$\mathbb{E}[e^{it\ell_N}] = N^{e^{it}-1} \left(\frac{1}{\Gamma(e^{it})} + o(1)\right).$$

This can be added to the list of analogy between multiplicative properties of permutations and random permutations. It remains an open problem to understand how permutations and primes can be mixed to obtain the Rényi-Turán formula (see [19] for more examples and explicit proofs where the role of random permutations appear in finite fields analogue).

As far as we are concerned, we would like to see whether the moments conjecture of Keating and Snaith is another manifestation of the above splitting phenomenon. In fact the discussion we already had above when explaining heuristically the moments conjecture shows that this is indeed the case. The heuristic we used to define

$$Y_N = -\sum_{p \leqslant N} \log\left(\left|1 - \frac{X_p}{\sqrt{p}}\right|\right)$$

as a random model for the modulus of the zeta function on the critical line can be justified by the following result:

THEOREM 2.7. *For $T \geq 0$, let $\Omega_T = [-T, T]$ be endowed with the uniform probability measure, that is the Lebesgue measure divided by $2T$ (note that we could as well take $\Omega_T = [0, T]$ equipped with $dt/T$). Let $X_T = (X_{p,T})_p$ be the $\hat{\mathbf{S}}^1 \equiv \prod_p \mathbf{S}^1$ valued random variable on $\Omega_T$, given by*

$$X_T(t) = (p^{-it})_p.$$

*Then as $T \to \infty$, the random variable $X_T$ converges in law to a random variable $X = (X_p)_p$ where the $X_p$ are independent and uniformly distributed on the unit circle $\mathbf{S}^1$.*

PROOF. Since $\hat{\mathbf{S}}^1$ is a compact Abelian group, Weyl's criterion shows that the statement is equivalent to the property that for any non trivial character $\chi : \hat{\mathbf{S}}^1 \longrightarrow \mathbf{S}^1$, we have

$$\lim_{T\to\infty} \mathbb{E}[\chi(X_{p,T})] = 0.$$

But an elementary property of the product of compact groups and the general form of the characters of $\mathbf{S}^1$ show that it is enough to prove that for characters $\chi$ of the form

$$\chi(z) = \prod_p z_p^{m_p},$$

for every finite non empty subset $S$ of prime numbers, $m_p \in \mathbb{Z}$, $m_p \neq 0$ and $p \in S$, and $z = (z_p)_p \in \hat{\mathbf{S}}^1$. This is now easily checked thanks to the elementary estimate that for $r > 0$, we have

$$|\mathbb{E}[r^{-it}]| \leq \min(1, \frac{1}{T|\log r|}).$$

□

So we can see that there is an analogy between the case of $\omega(n)$ and the case of the Riemann zeta function on the critical line. In the case of the Riemann zeta function, we have the random Euler product which is sometimes called the stochastic zeta function. The heuristic on which this model is based also comes from an equidistribution result for a sequence of arithmetically defined random variables (Theorem 2.7). This model is enough to predict Selberg's central limit theorem but is not enough for the Fourier transform of $\log(\zeta(1/2 + it))$. As for $\omega(n)$, a correction factor is needed which seems to account for the fact that primes are not behaving independently. In the case of $\omega(n)$, the correction factor seems to find its source in a compact group: it is related to some statistic coming from random permutations. In the case of the Riemann zeta function, it is conjectured that there is also a compact group with uniform measure: the unitary group. It is still a very open question on which there has been no significant progress to understand how the unitary group and primes should combine to obtain a proof of the conjecture.

CHAPTER 2

# The limiting behavior of the eigenvalues of random unitary matrices

The results we shall present here are very classical, but in some places (most notably when it comes to weak convergence of point processes) we have chosen our own proofs. For sake of completeness we have also put in Appendix 1 the classical methods the way they are usually exposed (e.g. like in [**21**]).

## 1. Preliminary facts about the distribution of eigenvalues

In this section we consider the most natural model of random unitary matrices. We call $U(n)$ the unitary group of size $n$, that is the set of matrices $U$ of size $n$ such that $UU^* = U^*U = Id$. Such matrices also represent isometries of $\mathbb{C}^n$, that is $\langle u(x), u(y)\rangle = \langle x, y\rangle$ for all $x, y \in \mathbb{C}^n$. Since for every $x \in \mathbb{C}^n$ we have $\|u(x)\| = \|x\|$, it is easily seen that $U(n)$ is a compact Lie group. There exists on $U(n)$ a natural probability measure which is given by the following fundamental result:

PROPOSITION 1.1. *Let $G$ be a compact Lie group. There exists a unique probability measure $\mathbb{P}$ (also noted $\mu_{Haar}$ or $\mu_H$ according to the context) on $G$ (endowed with its Borel sigma algebra) such that for every Borel measurable set $A$, and every $g \in G$, we have $\mathbb{P}(gA) = \mathbb{P}(Ag) = \mathbb{P}(A)$. The measure $\mathbb{P}$ is called the (probability) Haar measure, or uniform measure, on $G$.*

DEFINITION 1.2. *Let $n$ be a strictly positive integer, and endow $U(n)$ with the Haar measure $\mathbb{P}$. Then $(U(n), \mathbb{P})$ is called the Circular Unitary Ensemble of size n, and is noted CUE(n).*

One can easily check that in the context of the above Proposition, $\forall g \in G$, $\forall f \in L^1(G)$

$$\int_G f(gh) d\mathbb{P}(h) = \int_G f(hg) d\mathbb{P}(h) = \int_G f(h) d\mathbb{P}(h).$$

EXAMPLE 1.3. *Let us consider the following examples.*

(1) *Let $G$ be a finite group, $G = \{g_1, \cdots, g_n\}$ where the $g_i$'s are distinct. Then*

$$\mathbb{P} = \frac{1}{n}\sum_{j=1}^{n} \delta_{g_j}.$$

*is the Haar measure. Indeed, $\forall g$, $\forall f \in L^1(G)$, one has*

$$\frac{1}{n}\sum_{j=1}^{n} f(g_j) = \frac{1}{n}\sum_{j=1}^{n} f(gg_i) = \frac{1}{n}\sum_{j=1}^{n} f(g_i g),$$

*which comes from the fact that the map $h \mapsto hg$ from $G \to G$ is a bijection.*

(2) *$\mathbb{U} = \{e^{i\theta}, \ -\pi < \theta \leq \pi\}$ is an Abelian compact metric group. The normalized arc-length measure $d\theta/2\pi$ is a Haar measure.*

The following result is an immediate consequence of the definition of the Haar measure:

PROPOSITION 1.4. *Let $U$ be a deterministic matrix in $U(n)$ and let $M$ be a random matrix drawn from $CUE(n)$. Then $MU$, $UM$ and $UMU^{-1}$ are also in $CUE(n)$.*

Next we are interested in a very natural question in random matrix theory: what is the distribution of the eigenvalues of matrices in $CUE(n)$? It is well known that if $M \in U(n)$, then all its eigenvalues are on the unit circle $\mathbb{U}$. We can for instance order the eigenvalues in increasing order of the arguments in $(-\pi, \pi]$, for $j \in \{1, \ldots, n\}$,

$$\lambda_j(M) = e^{i\theta_j(M)},$$

where

$$-\pi < \theta_1(M) \leq \theta_2(M) \leq \cdots \leq \theta_n(M) \leq \pi.$$

If all eigenvalues are distinct, we can define for each $j \in \{1, \dots, n\}$, the eigenvector $v_j(M)$ associated with the eigenvalue $\lambda_j(M)$. Of course such an eigenvector is not unique and is defined up to a multiplicative factor. We can define it in a unique manner if we take it of norm one and such that (say) its first coordinate of largest modulus is real and strictly positive. Now let $V(M)$ be a matrix whose columns are the vectors $v_1(M), v_2(M), \cdots, v_n(M)$. Since $\|v_j(M)\| = 1$ for $j \in \{1, \dots, n\}$ and since the eigenspaces are orthogonal, we deduce that $V(M) \in U(n)$. Moreover for $j \in \{1, \dots, n\}$, $Mv_j(M) = \lambda_j(M)v_j(M)$, which implies that

$$MV(M) = V(M)\Lambda(M),$$

where $\Lambda(M)$ is the diagonal matrix whose $(j, j)$-th coefficient is $\lambda_j(M)$. In other words,

$$M = V(M)\Lambda(M)(V(M))^{-1},$$

that is $M$ is the conjugate of a diagonal matrix.

Next let us note $U^*(n)$ the set of unitary matrices of order $n$ having $n$ distinct eigenvalues, and let us note $\mathcal{D}_n$ the set of diagonal matrices of order $n$ whose coefficients (on the diagonal) are in $\mathbb{U}$, with strictly increasing arguments in $(-\pi, \pi]$. Last define $U_+(n)$ to be the set of matrices in $U(n)$ such that for each column, the first coefficient with highest modulus is a strictly positive real number. The above discussion shows that

$$\Pi : M \mapsto (\Lambda(M), V(M))$$

is a bijection from $U_n^*$ to $\mathcal{D}_n \times U_+(n)$, and the inverse bijection is given by

$$\Pi^{-1} : (\Lambda, V) \mapsto V\Lambda V^{-1}.$$

If we identify $\mathcal{D}_n$ with

$$\Delta_n := \{(e^{i\theta_1}, \dots, e^{i\theta_n}), -\pi < \theta_1 < \theta_2 < \cdots < \theta_n \leq \pi\},$$

we have the following formula, called Weyl's integration formula:

PROPOSITION 1.5. *If $M$ is in CUE(n), then a.s. $M$ has $n$ distinct eigenvalues. Moreover, the following hold:*

(1) *The probability distribution of $\Lambda(M)$ has density:*

$$D(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_n} \prod_{1 \leq j < k \leq n} |\lambda_k - \lambda_j|^2$$

*w.r.t. the uniform measure on $\Delta_n$, $Z_n > 0$ being a normalization constant.*

(2) *The distribution of $V(M)$ is the push forward of the Haar measure on $U(n)$ by the application which multiplies each column by a complex number of modulus 1 in such a way that we obtain a matrix in $U_+(n)$.*

(3) *The random variables $\Lambda(M)$ and $V(M)$ are independent.*

*Moreover if $\Lambda \in \mathcal{D}_n$ is distributed as in (1), and if $U \in U(n)$ is Haar distributed and if $\Lambda$ and $U$ are independent, then $U\Lambda U^{-1} \in CUE(n)$.*

The above proposition shows that the eigenvalues of CUE(n) tend to repulse each other: if $\lambda_1, \cdots, \lambda_{n-1}$ are fixed, then the density $D(\lambda_1, \dots, \lambda_n)$ tends to zero when $\lambda_n$ tends to one of the eigenvalues $\lambda_1, \cdots, \lambda_{n-1}$.

### 1.1. Correlation functions.

PROPOSITION 1.6. *Let $M \in CUE(n)$. The probability density function of $\Lambda(M)$ w.r.t the uniform measure on $\Delta_n$ is equal to*

$$D(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_n} \det((A_{j,k})_{1 \leq j,k \leq n}),$$

*where*

$$A_{j,k} = \sum_{\ell=0}^{n-1} (\overline{\lambda_j}\lambda_k)^{\ell}.$$

PROOF. We use the classical formula for Vandermonde's determinant:

$$\prod_{1 \leq j < k \leq n} (\lambda_k - \lambda_j) = \det(C),$$

where

$$C_{j,k} = \lambda_k^{j-1}.$$

Now we apply this result to the conjugate of $\lambda_1, \dots, \lambda_n$ and use the fact that the determinant of a matrix is equal to the determinant of its transpose to deduce that

$$\prod_{1 \le j < k \le n} \overline{(\lambda_k - \lambda_j)} = \det(B),$$

where

$$B_{j,k} = \overline{\lambda_j}^{k-1}.$$

We now multiply both equalities to obtain

$$\prod_{1 \le j < k \le n} |\lambda_k - \lambda_j|^2 = \det(BC),$$

where

$$(BC)_{j,k} = \sum_{\ell=1}^{n} B_{j,\ell} C_{\ell,k} = \sum_{\ell=1}^{n} \overline{\lambda_j}^{\ell-1} \lambda_k^{\ell-1},$$

i.e.

$$(BC)_{j,k} = A_{j,k},$$

which is the desired result.

□

One can state this result under a slightly different form in a way that the ordering of the angles does not matter.

THEOREM 1.7. *Let $M \in CUE(n)$ and let $E$ be the set of its eigenvalues. For every bounded measurable function $F : \mathbb{U}^n \to \mathbb{R}$, we have:*

$$\mathbb{E}\left( \sum_{\mu_1 \neq \mu_2 \neq \cdots \neq \mu_n \in E} F(\mu_1, \dots, \mu_n) \right) = \frac{1}{Z_n'} \int_{\mathbb{U}^n} \rho_n(\lambda_1, \dots, \lambda_n) F(\lambda_1, \dots, \lambda_n) d\nu(\lambda_1) \dots d\nu(\lambda_n),$$

*where $Z_n' > 0$ is a normalization constant, $\nu$ is the uniform measure on $\mathbb{U}$ and*

$$\rho_n(\lambda_1, \dots, \lambda_n) = \det\left( (K_0(\lambda_j, \lambda_k))_{1 \le j,k \le n} \right),$$

*with*

$$K_0(\lambda, \lambda') = \sum_{\ell=0}^{n-1} (\overline{\lambda}\lambda')^{\ell}.$$

PROOF. We first need to check that

$$\mathbb{E}\left( \sum_{\mu_1 \neq \mu_2 \neq \cdots \neq \mu_n \in E} F(\mu_1, \dots, \mu_n) \right) = \mathbb{E}\left( G(\lambda_1(M), \dots, \lambda_n(M)) \right),$$

where

$$G(\mu_1, \dots, \mu_n) = \sum_{\sigma \in \mathfrak{S}_n} F(\mu_{\sigma(1)}, \dots \mu_{\sigma(n)}),$$

which follows from the fact that the $n$-tuples $(\mu_1, \mu_2, \dots, \mu_n)$ of distinct elements of $E$ are exactly the $n!$ permutations of $(\lambda_1(M), \dots, \lambda_n(M))$. Now an application of the previous proposition with Weyl's integration formula yields

$$\mathbb{E}\left( \sum_{\mu_1 \neq \mu_2 \neq \cdots \neq \mu_n \in E} F(\mu_1, \dots, \mu_n) \right) = \frac{1}{Z_n} \int_{\Delta_n} G(\lambda_1, \dots, \lambda_n) \det((A_{j,k})_{1 \le j,k \le n}) d\alpha(\lambda_1, \dots, \lambda_n),$$

where $\alpha$ is the uniform measure on $\Delta_n$. Hence

$$\mathbb{E}\left( \sum_{\mu_1 \neq \mu_2 \neq \cdots \neq \mu_n \in E} F(\mu_1, \dots, \mu_n) \right) = \frac{1}{Z_n'} \int_{\Delta_n} G(\lambda_1, \dots, \lambda_n) \rho_n(\lambda_1, \dots, \lambda_n) d\nu(\lambda_1) \dots d\nu(\lambda_n),$$

where $Z_n' > 0$ is another normalization constant. Now we deduce that the previous expectation is equal to

$$\frac{1}{Z_n'} \sum_{\sigma \in \mathfrak{S}_n} \int_{\Delta_n} F(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(n)}) \rho_n(\lambda_1, \dots, \lambda_n) d\nu(\lambda_1) \dots d\nu(\lambda_n)$$

$$= \frac{1}{Z'_n} \sum_{\sigma \in \mathfrak{S}_n} \int_{\sigma(\Delta_n)} F(\lambda_1, \dots, \lambda_n) \rho_n(\lambda_{\sigma^{-1}(1)}, \dots, \lambda_{\sigma^{-1}(n)}) d\nu(\lambda_1) \dots d\nu(\lambda_n),$$

where $\sigma(\Delta_n)$ is the image of $\Delta_n$ by the map

$$(\lambda_1, \dots, \lambda_n) \mapsto (\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(n)}).$$

Next we note that $\rho_n$ is invariant under a permutation of the coordinates. Indeed when we permute the elements of $(\lambda_1, \dots, \lambda_n)$, the rows and the columns of the matrix $(K_0(\lambda_j, \lambda_k))_{1 \le j,k \le n}$ are permuted accordingly and hence the determinant does not change. The last expression is hence equal to

$$\frac{1}{Z'_n} \sum_{\sigma \in \mathfrak{S}_n} \int_{\sigma(\Delta_n)} F(\lambda_1, \dots, \lambda_n) \rho_n(\lambda_1, \dots, \lambda_n) d\nu(\lambda_1) \dots d\nu(\lambda_n).$$

Now we also note that the sets $\sigma(\Delta_n)$ are pairwise disjoint and that their union is equal to

$$\{(\lambda_1, \dots, \lambda_n) \in \mathbb{U}^n, \lambda_1 \neq \dots \neq \lambda_n\},$$

which is of full measure in $\mathbb{U}^n$ w.r.t. the uniform measure. Consequently we have:

$$\mathbb{E}\left( \sum_{\mu_1 \neq \mu_2 \neq \dots \neq \mu_n \in E} F(\mu_1, \dots, \mu_n) \right) = \frac{1}{Z'_n} \int_{\mathbb{U}^n} \rho_n(\lambda_1, \dots, \lambda_n) F(\lambda_1, \dots, \lambda_n) d\nu(\lambda_1) \dots d\nu(\lambda_n).$$

□

We shall see later that in fact the normalization constant $Z'_n$ is in fact equal to 1. The function $\rho_n(\lambda_1, \dots, \lambda_n)$ gives information about the probability of finding an eigenvalue in the neighborhood of the points $\lambda_1, \dots, \lambda_n$. The function $\rho_n : \mathbb{U}^n \to \mathbb{R}_+$ is called the $n$-point correlation function. One can go backwards from there and compute $r$-point correlation functions for $1 \le r \le n$.

PROPOSITION 1.8. *With the notation above, for all $r \in \{1, \dots, n\}$, and for any bounded measurable function $F$ from $\mathbb{U}^r$ into $\mathbb{R}$, we have:*

$$\mathbb{E}\left( \sum_{\mu_1 \neq \mu_2 \neq \dots \neq \mu_r \in E} F(\mu_1, \dots, \mu_r) \right) = \frac{1}{Z'_n} \int_{\mathbb{U}^r} \rho_r(\lambda_1, \dots, \lambda_r) F(\lambda_1, \dots, \lambda_r) d\nu(\lambda_1) \dots d\nu(\lambda_r),$$

*where the function $\rho_r$ from $\mathbb{U}^r$ to $\mathbb{R}_+$ satisfies the backwards induction relation:*

$$\rho_r(\lambda_1, \dots, \lambda_r) = \frac{1}{n-r} \int_{\mathbb{U}} \rho_{r+1}(\lambda_1, \dots, \lambda_r, \lambda) d\nu(\lambda),$$

*for all $r \in \{1, \dots, n-1\}$.*

PROOF. For $1 \le r \le n-1$, let us assume the result holds for $r+1$. For all bounded and measurable function $G$ from $\mathbb{U}^{r+1}$ to $\mathbb{R}$, we have

$$\mathbb{E}\left( \sum_{\mu_1 \neq \mu_2 \neq \dots \neq \mu_{r+1} \in E} G(\mu_1, \dots, \mu_{r+1}) \right)$$

$$= \frac{1}{Z'_n} \int_{\mathbb{U}^{r+1}} \rho_{r+1}(\lambda_1, \dots, \lambda_{r+1}) G(\lambda_1, \dots, \lambda_{r+1}) d\nu(\lambda_1) \dots d\nu(\lambda_{r+1}).$$

If $F$ is the map from $\mathbb{U}^r$ to $\mathbb{R}$ given by

$$F(\lambda_1, \dots, \lambda_r) = G(\lambda_1, \dots, \lambda_{r+1}),$$

we thus have

$$\mathbb{E}\left( \sum_{\mu_1 \neq \mu_2 \neq \dots \neq \mu_{r+1} \in E} F(\mu_1, \dots, \mu_r) \right)$$

$$= \frac{1}{Z'_n} \int_{\mathbb{U}^{r+1}} \rho_{r+1}(\lambda_1, \dots, \lambda_{r+1}) F(\lambda_1, \dots, \lambda_r) d\nu(\lambda_1) \dots d\nu(\lambda_{r+1})$$

$$= \frac{1}{Z'_n} \int_{\mathbb{U}^r} F(\lambda_1, \dots, \lambda_r) d\nu(\lambda_1) \dots d\nu(\lambda_r) \int_{\mathbb{U}} \rho_{r+1}(\lambda_1, \dots, \lambda_{r+1}) d\nu(\lambda_{r+1}).$$

The terms of the sum do not depend on $\mu_{r+1}$. But $\mu_{r+1}$ is an eigenvalue of $M$ which must be different from $\mu_1, \dots, \mu_r$, which leaves $n-r$ possible choices. Consequently

$$(n-r)\mathbb{E}\left(\sum_{\mu_1 \neq \mu_2 \neq \dots \neq \mu_r \in E} F(\mu_1, \dots, \mu_r)\right)$$
$$= \frac{1}{Z'_n} \int_{\mathbb{U}^r} F(\lambda_1, \dots, \lambda_r) d\nu(\lambda_1) \dots d\nu(\lambda_r) \int_{\mathbb{U}} \rho_{r+1}(\lambda_1, \dots, \lambda) d\nu(\lambda),$$

and we conclude with a backwards induction.

□

The above proposition applies to a large class of point processes, but we shall not go in this direction. In this lecture we shall rather explore the fact that the correlation function $\rho_n$ can be expressed as a determinant. The following lemma will reveal helpful:

LEMMA 1.9. *For all $\lambda, \lambda' \in \mathbb{U}$,*

$$K_0(\lambda, \lambda') = \int_{\mathbb{U}} K_0(\lambda, \lambda'') K_0(\lambda'', \lambda') d\nu(\lambda''),$$

*and*

$$\int_{\mathbb{U}} K_0(\lambda, \lambda) d\nu(\lambda) = n.$$

PROOF. The second identity is obvious for $K_0(\lambda, \lambda) = n$ for all $\lambda \in \mathbb{U}$. For the first identity we note that

$$\int_{\mathbb{U}} K_0(\lambda, \lambda'') K_0(\lambda'', \lambda') d\nu(\lambda'') = \int_{\mathbb{U}} \left(\sum_{\ell=0}^{n-1} (\overline{\lambda}\lambda'')^\ell\right) \left(\sum_{\ell=0}^{n-1} (\overline{\lambda''}\lambda')^\ell\right) d\nu(\lambda'')$$
$$= \sum_{0 \leq \ell, \ell' \leq n-1} \overline{\lambda}^\ell (\lambda')^{\ell'} \int_{\mathbb{U}} (\lambda'')^\ell (\overline{\lambda''})^{\ell'} d\nu(\lambda'').$$

But

$$\int_{\mathbb{U}} (\lambda'')^\ell (\overline{\lambda''})^{\ell'} d\nu(\lambda'') = \int_{\mathbb{U}} (\lambda'')^{\ell - \ell'} d\nu(\lambda'')$$

equals 1 if $\ell = \ell'$ and 0 otherwise. We can then conclude that

$$\int_{\mathbb{U}} K_0(\lambda, \lambda'') K_0(\lambda'', \lambda') d\nu(\lambda'') = \sum_{0 \leq \ell \leq n-1} (\overline{\lambda})^\ell (\lambda')^\ell = K_0(\lambda, \lambda').$$

□

We are now able to obtain a simple representation for the $r$-point correlation function:

PROPOSITION 1.10. *For every $r \in \{1, \dots, n\}$, and every bounded and measurable function $F$ from $\mathbb{U}^r$ to $\mathbb{R}$, we have:*

$$\mathbb{E}\left(\sum_{\mu_1 \neq \mu_2 \neq \dots \neq \mu_r \in E} F(\mu_1, \dots, \mu_r)\right) = \int_{\mathbb{U}^r} \rho_r(\lambda_1, \dots, \lambda_r) F(\lambda_1, \dots, \lambda_r) d\nu(\lambda_1) \dots d\nu(\lambda_r),$$

*where*

$$\rho_r(\lambda_1, \dots, \lambda_r) = \det\left((K_0(\lambda_j, \lambda_k))_{1 \leq j,k \leq r}\right).$$

PROOF. We prove the result with a backwards induction. We have already established the result for $r = n$. We assume that the expression given by the proposition holds for $r+1$, for $1 \leq r \leq n-1$. Then from previous results we have

$$\rho_r(\lambda_1, \dots, \lambda_r) = \frac{1}{n-r} \int_{\mathbb{U}} \det\left((K_0(\lambda_j, \lambda_k))_{1 \leq j,k \leq r+1}\right) d\nu(\lambda_{r+1}).$$

We expand the determinant and isolate the factors involving $\lambda_{r+1}$ to obtain:

$$\det\left((K_0(\lambda_j, \lambda_k))_{1 \leq j,k \leq r+1}\right) = \sum_{\sigma \in \mathfrak{S}_{r+1}, \sigma(r+1) = r+1} \epsilon(\sigma) K_0(\lambda_{r+1}, \lambda_{r+1}) \prod_{1 \leq j \leq r} K_0(\lambda_j, \lambda_{\sigma(j)})$$
$$+ \sum_{1 \leq a,b \leq r} \sum_{\sigma \in \mathfrak{S}_{r+1}, \sigma(a) = r+1, \sigma(r+1) = b} \epsilon(\sigma) K_0(\lambda_a, \lambda_{r+1}) K_0(\lambda_{r+1}, \lambda_b) \prod_{1 \leq j \leq r, j \neq a} K_0(\lambda_j, \lambda_{\sigma(j)}).$$

Next we integrate w.r.to $\lambda_{r+1}$ using the previous lemma:

$$(n-r)\rho_r(\lambda_1,\dots,\lambda_r) = n\sum_{\sigma\in\mathfrak{S}_r}\epsilon(\sigma)\prod_{1\le j\le r}K_0(\lambda_j,\lambda_{\sigma(j)}) + \sum_{1\le a,b\le r}\sum_{\sigma\in\mathfrak{S}_{r+1},\sigma(a)=r+1,\sigma(r+1)=b}\epsilon(\sigma)K_0(\lambda_a,\lambda_b)\prod_{1\le j\le r,j\neq a}K_0(\lambda_j,\lambda_{\sigma(j)}).$$

Now we note that the permutations $\sigma$ in $\mathfrak{S}_{r+1}$ such that $\sigma(a)=r+1$ and $\sigma(r+1)=b$ are in one to one correspondence with permutations $\sigma'$ in $\mathfrak{S}_r$ satisfying $\sigma'(a)=b$: to see it, define $\sigma'$ from $\sigma$ by setting $\sigma'(c)=\sigma(c)$ for all $c\in\{1,\dots,r\}$ such that $c\neq a$. With this bijection, we get:

$$(n-r)\rho_r(\lambda_1,\dots,\lambda_r) = n\sum_{\sigma\in\mathfrak{S}_r}\epsilon(\sigma)\prod_{1\le j\le r}K_0(\lambda_j,\lambda_{\sigma(j)}) + \sum_{1\le a,b\le r}\sum_{\sigma'\in\mathfrak{S}_r,\sigma'(a)=b}\epsilon(\sigma)\prod_{1\le j\le r,}K_0(\lambda_j,\lambda_{\sigma'(j)}).$$

Moreover if $\sigma''$ is the element of $\mathfrak{S}_{r+1}$ which satisfies $\sigma''(c)=\sigma'(c)$ for $c\le r$ and $\sigma''(r+1)=r+1$, one easily checks that $\sigma=\tau_{r+1,b}\circ\sigma''$, where $\tau_{r+1,b}$ is the transposition exchanging $r+1$ and $b$, and thus

$$\epsilon(\sigma)=-\epsilon(\sigma'')=-\epsilon(\sigma').$$

Consequently

$$(n-r)\rho_r(\lambda_1,\dots,\lambda_r) = n\sum_{\sigma\in\mathfrak{S}_r}\epsilon(\sigma)\prod_{1\le j\le r}K_0(\lambda_j,\lambda_{\sigma(j)}) - \sum_{1\le a,b\le r}\sum_{\sigma'\in\mathfrak{S}_r,\sigma'(a)=b}\epsilon(\sigma')\prod_{1\le j\le r,}K_0(\lambda_j,\lambda_{\sigma'(j)}),$$

that is

$$(n-r)\rho_r(\lambda_1,\dots,\lambda_r) = n\sum_{\sigma\in\mathfrak{S}_r}\epsilon(\sigma)\prod_{1\le j\le r}K_0(\lambda_j,\lambda_{\sigma(j)}) - \sum_{1\le a\le r}\sum_{\sigma'\in\mathfrak{S}_r}\epsilon(\sigma')\prod_{1\le j\le r,}K_0(\lambda_j,\lambda_{\sigma'(j)}),$$

The right hand side of the last equality is equal to

$$(n-r)\det\left((K_0(\lambda_j,\lambda_k))_{1\le j,k\le r}\right),$$

which proves the result for $r$ and hence by backwards induction the statement is always true. Now the previous proposition states that

$$\mathbb{E}\left(\sum_{\mu_1\neq\mu_2\neq\dots\neq\mu_r\in E}F(\mu_1,\dots,\mu_r)\right)=\frac{1}{Z_n'}\int_{\mathbb{U}^r}\rho_r(\lambda_1,\dots,\lambda_r)F(\lambda_1,\dots,\lambda_r)d\nu(\lambda_1)\dots d\nu(\lambda_r),$$

and so it is enough to show that $Z_n'=1$.

For this we take $r=1$ and $F\equiv 1$:

$$\sum_{\mu_1\in E}1=\frac{1}{Z_n'}\int_{\mathbb{U}}\rho_1(\lambda)d\nu(\lambda),$$

where

$$\sum_{\mu_1\in E}1=n$$

and

$$\int_{\mathbb{U}}\rho_1(\lambda)d\nu(\lambda)=\int_{\mathbb{U}}K_0(\lambda,\lambda)d\nu(\lambda)=n,$$

from which we immediately deduce the desired result.

□

A random point process whose correlation functions are of the form above is called a determinantal point process. Determinantal point processes play a very important role in random matrix theory and mathematical physics. The function $K_0$ is called the kernel of this determinantal point process. Note that $K_0$ is a complex valued function while the correlation functions $\rho_r$ are real valued. We thus give an alternative description:

THEOREM 1.11. *For all $r \in \{1, \dots, n\}$, and for $\theta_1, \dots, \theta_r \in \mathbb{R}$, the correlation function $\rho_r$ satisfies:*

$$\rho_r(e^{i\theta_1}, \dots, e^{i\theta_r}) = \det((K(\theta_j, \theta_k))_{1 \le j,k \le n}),$$

*where*

$$K(\theta, \theta') = \frac{\sin[n(\theta' - \theta)/2]}{\sin[(\theta' - \theta)/2]},$$

*for $\theta \neq \theta'$ modulo $2\pi$, and*

$$K(\theta, \theta') = n$$

*if $\theta = \theta'$ modulo $2\pi$.*

PROOF. The result is already proven if one replaces $K(\theta_j, \theta_k)$ with $K'(\theta_j, \theta_k)$, where

$$K'(\theta_j, \theta_k) = K_0(e^{i\theta_j}, e^{i\theta_k}) = \sum_{\ell=0}^{n-1} e^{i\ell(\theta_k - \theta_j)} = \frac{e^{in\theta} - 1}{e^{i\theta} - 1} = e^{i\theta(n-1)/2} \frac{e^{in\theta/2} - e^{-in\theta/2}}{e^{i\theta/2} - e^{-i\theta/2}},$$

if $\theta := \theta_k - \theta_j \neq 0$ modulo $2\pi$, and

$$K'(\theta_j, \theta_k) = n$$

if $\theta$ is a multiple of $2\pi$. Consequently we have

$$K'(\theta_j, \theta_k) = e^{-i\theta_j(n-1)/2} e^{i\theta_k(n-1)/2} K(\theta_j, \theta_k).$$

It follows that $(K'(\theta_j, \theta_k))_{1 \le j,k \le n}$ can be obtained from the matrix $(K(\theta_j, \theta_k))_{1 \le j,k \le n}$ by multiplying the $j$-th raw by $e^{-i\theta_j(n-1)/2}$ for all $j$, and the $k$-th column by $e^{i\theta_k(n-1)/2}$ for all $k$. One can then easily check that both matrices have the same determinant.

□

We note that in order to compute $\rho_r$ at $r$ points of the unit circle using the previous proposition, we need to consider the arguments of these points. These arguments are a priori only defined modulo $2\pi$. If $n$ is odd, then $K(\theta_j, \theta_k)$ does not depend on the choices made for $\theta_j$ and $\theta_k$, since $\sin(n\theta/2)/\sin(\theta/2)$ is $2\pi$-periodic in $\theta$. On the other hand if $n$ is even, this last function is $4\pi$-periodic, so that $K(\theta_j, \theta_k)$ can change sign, for instance if one adds $2\pi$ to $\theta_j$. However this change of sign does not alter the value of the determinant. For instance if $n = 2$, we have $K(\theta, \theta) = 2$,

$$K(\theta, \theta') = \frac{\sin(\theta' - \theta)}{\sin[(\theta' - \theta)/2]} = 2\cos[(\theta' - \theta)/2],$$

which changes sign if one adds $2\pi$ to $\theta$ or $\theta'$. However a simple determinant computation yields

$$\rho_2(e^{i\theta}, e^{i\theta'}) = 4(1 - \cos^2[(\theta' - \theta)/2]) = 4\sin^2[(\theta' - \theta)/2],$$

which only depends on $e^{i\theta}$ and $e^{i\theta'}$: we have

$$\rho_2(\lambda, \lambda') = |\lambda - \lambda'|^2,$$

in agreement with Proposition 1.5.

More generally we can give the following general formula when $r = 2$:

COROLLARY 1.12. *The 2-point correlation for CUE(n) is given by*

$$\rho_2(e^{i\theta}, e^{i\theta'}) = n^2 - \frac{\sin^2[n(\theta' - \theta)/2]}{\sin^2[(\theta' - \theta)/2]},$$

*whenever $e^{i\theta} \neq e^{i\theta'}$.*

A simple Taylor expansion yields, for $n \ge 2$,

$$\rho_2(e^{i\theta}, e^{i\theta'}) \sim n^2(n^2 - 1)(\theta' - \theta)^2/12$$

when $\theta' - \theta$ tends to zero. We recover the fact the for CUE(n), the eigenvalues exhibit a repulsion proportional to the square of their distance.

**1.2. The macroscopic behavior of the eigenvalues of CUE(n).** We introduce the empirical distribution of the eigenvalues; if $\lambda_1, \dots, \lambda_n$ are the eigenvalues, we define

$$\mu_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j},$$

where $\delta_{\lambda_j}$ is the Dirac mass at $\lambda_j$. We would now like to prove that this measure converges to the uniform measure on the unit circle. For this we consider its Fourier transform: for $k \in \mathbb{Z}$,

$$\hat{\mu}_n(k) = \int_{\mathbb{U}} z^k d\mu_n(z) = \frac{1}{n} \sum_{j=1}^{n} \lambda_j^k = \frac{1}{n} \operatorname{Tr}(M^k),$$

where $M \in CUE(n)$. This shows that it is important to understand the behavior of traces of powers. We state and prove a simple result which is enough for our purposes:

PROPOSITION 1.13. *Let $M \in CUE(n)$. For $k = 0$, $\operatorname{Tr}(M^k) = n$, and for $k \neq 0$,*

$$\mathbb{E}[\operatorname{Tr}(M^k)] = 0, \ \mathbb{E}[|\operatorname{Tr}(M^k)|^2] = |k| \wedge n.$$

It is interesting to compare this result to the case of $n$ points taken independently and uniformly on the unit circle. If $\lambda'_1, \dots \lambda'_n$ are chosen in this way, then

$$\sum_{j=1}^{n} (\lambda'_j)^k = n$$

if $k = 0$, and

$$\mathbb{E}\left[\sum_{j=1}^{n} (\lambda'_j)^k\right] = 0, \ \mathbb{E}\left[\left|\sum_{j=1}^{n} (\lambda'_j)^k\right|^2\right] = n$$

if $k \neq 0$.

PROOF. The case $k = 0$ is trivial so we shall focus on the case $k \neq 0$. Since the distribution of $M$ and $\lambda M$ is the same for every $\lambda \in \mathbb{U}$, we have:

$$\mathbb{E}[\operatorname{Tr}(M^k)] = \mathbb{E}[\operatorname{Tr}((\lambda M)^k)] = \lambda^k \mathbb{E}[\operatorname{Tr}(M^k)],$$

hence

$$\mathbb{E}[\operatorname{Tr}(M^k)] = 0.$$

To compute the second moment we express the traces as functions of the eigenvalues to obtain:

$$\mathbb{E}[|\operatorname{Tr}(M^k)|^2] = \mathbb{E}\left[n + \sum_{1 \le p \neq q \le n} \lambda_p^k \overline{\lambda_q^k}\right] = n + \int_{\mathbb{U}^2} (\lambda \overline{\lambda'})^k \rho_2(\lambda, \lambda') d\nu(\lambda) d\nu(\lambda'),$$

where $d\nu$ is the uniform measure on $\mathbb{U}$. An application of Proposition 1.10 yields:

$$\begin{aligned}
\mathbb{E}[|\operatorname{Tr}(M^k)|^2] &= n + \int_{\mathbb{U}^2} (\lambda \overline{\lambda'})^k \left(K_0(\lambda,\lambda) K_0(\lambda',\lambda') - K_0(\lambda,\lambda') K_0(\lambda',\lambda)\right) d\nu(\lambda) d\nu(\lambda') \\
&= n + \int_{\mathbb{U}^2} (\lambda \overline{\lambda'})^k \left(n^2 - \sum_{0 \le \ell, \ell' \le n-1} (\lambda \overline{\lambda'})^{\ell} (\lambda' \overline{\lambda})^{\ell'}\right) d\nu(\lambda) d\nu(\lambda') \\
&= n + n^2 \int_{\mathbb{U}^2} (\lambda \overline{\lambda'})^k d\nu(\lambda) d\nu(\lambda') - \sum_{0 \le \ell, \ell' \le n-1} \int_{\mathbb{U}^2} (\lambda \overline{\lambda'})^{k+\ell-\ell'} d\nu(\lambda) d\nu(\lambda').
\end{aligned}$$

In the last sum of the equality above, only indexes corresponding to $\ell$ and $\ell'$ such that $k + \ell - \ell' = 0$ yield a non zero term (in which case the term is equal to 1). For $|k| \ge n$ there are no $\ell$ and $\ell'$ such that $k + \ell - \ell' = 0$. For $0 < k < n$, the pairs of indexes satisfying this equality are $(0, k), (1, k+1), \dots, (n-1-k, n-1)$: there are $n - k$ of them. For $-n < k < 0$, the pairs of indexes satisfying this equality are $(-k, 0), (1-k, 1), \dots, (n-1, n-1+k)$: there are $n + k$ of them. In all cases there thus $(n - |k|) \vee 0$ such pairs, which yields

$$\mathbb{E}[|\operatorname{Tr}(M^k)|^2] = n - [(n - |k|) \vee 0] = |k| \wedge n.$$

□

We can now prove the following convergence result:

PROPOSITION 1.14. *Let $(M_n)_{n\geq 1}$ be a sequence of random matrices, with $M \in CUE(n)$. We note $\mu_n$ the empirical distribution of the eigenvalues. Then almost surely $\mu_n$ converges in law, when $n \to \infty$, to the uniform measure on the unit circle.*

PROOF. Let $q > 0$ be an integer. It follows from the previous proposition that for $k \neq 0$,

$$\mathbb{P}\left(|\hat{\mu}_n(k)| \geq 1/q\right) \leq q^2 \mathbb{E}\left[|\hat{\mu}_n(k)|^2\right] = \frac{q^2}{n^2}\mathbb{E}[|\operatorname{Tr}(M^k)|^2] = \frac{q^2(|k| \wedge n)}{n^2} \leq \frac{|k|q^2}{n^2}.$$

An application of the Borel Cantelli lemma yields that a.s. for $k \neq 0$ and for all $q \geq 1$, $|\hat{\mu}_n(k)| < 1/q$ if $n$ is sufficiently large. In other words, a.s. for all $k \neq 0$,

$$\hat{\mu}_n(k) \underset{n\to\infty}{\longrightarrow} 0 = \hat{\nu}(k),$$

where $\nu$ is the uniform measure on the unit circle. Moreover it is immediate that

$$\hat{\mu}_n(0) = \hat{\nu}(0) = 1,$$

consequently $\mu_n$ converges a.s. in law to $\nu$.

□

We observe from the above results that for $k \neq 0$, the $L^2$ norm of the trace of $M^k$ does not depend on $n$, if $n \geq |k|$. This means that this moment becomes constant and thus converges when $n$ increases. This is in fact a special case of a more general result due to Diaconis and Shahshahani ([**9**], [**10**]):

PROPOSITION 1.15 (Diaconis and Shahshahani). *Let $p \geq 1$, $a_1, \dots, a_p, b_1, \dots, b_p \geq 0$ be integers and let $M \in CUE(n)$. We have:*

$$\mathbb{E}\left[\prod_{j=1}^{p}\left(\operatorname{Tr}(M^j)\right)^{a_j}\left(\overline{\operatorname{Tr}(M^j)}\right)^{b_j}\right] = \mathbb{1}_{\forall j \in \{1,\dots,p\}, a_j = b_j}\prod_{j=1}^{p} j^{a_j}(a_j)!,$$

*whenever*

$$n \geq \sum_{j=1}^{p} j(a_j + b_j).$$

We do not prove this result but rather give as a consequence a result on the convergence of traces of powers.

COROLLARY 1.16. *Let $p > 0$ be an integer and let $M \in CUE(n)$. Then the following convergence in law holds:*

$$\left(\operatorname{Tr}(M_n^k)\right)_{1\leq k\leq p} \underset{n\to\infty}{\longrightarrow} (Z_k)_{1\leq k\leq p},$$

*where $(Z_k)_{k\geq 1}$ are complex independent Gaussian random variables satisfying*

$$\mathbb{E}[Z_k] = \mathbb{E}[Z_k^2] = 0, \ \mathbb{E}[|Z_k|^2] = k.$$

PROOF. It is equivalent to show the convergence of joint moments of $\left(\operatorname{Tr}(M_n^k)\right)_{1\leq k\leq p}$. According to the previous proposition each joint moment is constant for $n$ sufficiently large. We thus need to show that for integers $a_1, \dots, a_p, b_1, \dots, b_p \geq 0$,

$$\mathbb{E}\left[\prod_{j=1}^{p} Z_j^{a_j}\overline{Z_j}^{b_j}\right] = \mathbb{1}_{\forall j \in \{1,\dots,p\}, a_j = b_j}\prod_{j=1}^{p} j^{a_j}(a_j)!.$$

Since the variables $(Z_j)_{j\geq 1}$ are independent, it is enough to show that

$$\mathbb{E}[Z_j^{a_j}\overline{Z_j}^{b_j}] = \mathbb{1}_{a_j = b_j} j^{a_j}(a_j)!.$$

The case $a_j \neq b_j$ follows from the rotation invariance property of the distributions of $Z_j$. Moreover we note that $Z_j$ is distributed like $\sqrt{j}Z_1$, and thus for $a_j = b_j$, we have

$$\mathbb{E}[|Z_j|^{2a_j}] = j^{a_j}\mathbb{E}[|Z_1|^{2a_j}].$$

Next we note that $Z_1 = \mathcal{N}_1 + i\mathcal{N}_2$, where $\mathcal{N}_1$ and $\mathcal{N}_2$ are two independent centred Gaussian random variables with variance 1/2. Hence we have $|Z_1|^2 = \mathcal{N}_1^2 + \mathcal{N}_2^2$ and

$$\mathbb{E}[|Z_1|^{2a_j}] = \int_0^\infty e^{-t}t^{a_j}dt = (a_j)!$$

which completes the proof. □

We thus see that the traces of powers of $CUE(n)$ random matrices converge in law without any normalization. This is very different from the situation of $n$ i.i.d. points uniformly chosen on the unit circle where according to the central limit theorem one needs a $\sqrt{n}$ normalization.

The Diaconis-Shahshahani result deals with the situation where the powers are fixed and the size of the matrices goes to infinity. One may ask what happens when the size is fixed and the powers become large. We already know from Proposition 1.13 that $\mathbb{E}[|\operatorname{Tr}(M_n^k)|^2] = n$ if $|k| > n$: this is the same second moment obtained when one replaces the eigenvalues with $n$ i.i.d. points uniformly distributed on the unit circle. Remarkably this generalizes to:

PROPOSITION 1.17 (Rains). *Let $M \in CUE(n)$ and let $k$ be such that $|k| \geq n$. Then the eigenvalues of $M_n^k$ are distributed like $n$ independent random variables uniformly distributed on the unit circle.*

PROOF. Since the set of trigonometric polynomials is dense in the space of continuous functions on $\mathbb{U}$, it is enough to prove that for a symmetric function $F$ from $\mathbb{U}^n$ to $\mathbb{C}$, which is polynomial in its variables and its inverses, we have

$$\int_{\mathbb{U}^n} F(\lambda_1^k, \dots, \lambda_n^k)\rho_n(\lambda_1, \dots, \lambda_n) d\nu(\lambda_1) \dots d\nu(\lambda_n) = K_n \int_{\mathbb{U}^n} F(\lambda_1^k, \dots, \lambda_n^k) d\nu(\lambda_1) \dots d\nu(\lambda_n),$$

$K_n$ depending only on $n$, and $\rho_n$ being the $n$-point correlation function for $CUE(n)$. Now we use the fact that there exists $C_n > 0$ such that

$$\rho_n(\lambda_1, \dots, \lambda_n) = C_n \prod_{1 \leq j < k \leq n} |\lambda_j - \lambda_k|^2 = C_n \prod_{1 \leq j < k \leq n} (\lambda_j - \lambda_k)(\lambda_j^{-1} - \lambda_k^{-1}) = R(\lambda_1, \dots, \lambda_n),$$

where $R$ is a polynomial function in $\lambda_1, \dots, \lambda_n, \lambda_1^{-1}, \dots, \lambda_n^{-1}$, symmetric in $\lambda_1, \dots, \lambda_n$, and such that the exponent of each of the variables $\lambda_1, \dots, \lambda_n$ is between $-n+1$ and $n-1$. Now if $\alpha\lambda_1^{m_1} \dots \lambda_n^{m_n}$, $-n+1 \leq m_1, \dots, m_n \leq n-1$ is one of the terms of $R$, the integral

$$\int_{\mathbb{U}^n} \alpha\lambda_1^{m_1} \dots \lambda_n^{m_n} F(\lambda_1^k, \dots, \lambda_n^k) d\nu(\lambda_1) \dots d\nu(\lambda_n)$$

decomposes into a linear combination of integrals

$$\int_{\mathbb{U}^n} \lambda_1^{m_1'} \dots \lambda_n^{m_n'} d\nu(\lambda_1) \dots d\nu(\lambda_n)$$

where the exponents $m_1', \dots, m_n'$ are respectively congruent to $m_1, \dots, m_n$ modulo $k$. This last integral is non zero only in the case where $m_1', \dots, m_n'$ are zero, which means that $m_1, \dots, m_n$ are multiple of $k$. But since their absolute value is strictly smaller than $n$, this implies that $m_1, \dots, m_n$ are all zero. Hence we have:

$$\begin{aligned} &\int_{\mathbb{U}^n} F(\lambda_1^k, \dots, \lambda_n^k)\rho_n(\lambda_1, \dots, \lambda_n) d\nu(\lambda_1) \dots d\nu(\lambda_n) \\ &= \int_{\mathbb{U}^n} F(\lambda_1^k, \dots, \lambda_n^k) R(\lambda_1, \dots, \lambda_n) d\nu(\lambda_1) \dots d\nu(\lambda_n) \\ &= \int_{\mathbb{U}^n} K_n F(\lambda_1^k, \dots, \lambda_n^k) d\nu(\lambda_1) \dots d\nu(\lambda_n), \end{aligned}$$

where $K_n$ is the constant term of the polynomial $R$. This completes the proof.

□

**1.3. The behavior of eigenvalues at the microscopic scale.** We first make the observation that a unitary matrix with simple eigenvalues has $n$ eigenangles in any interval of length $2\pi$, such as $(-\pi, \pi]$. The average spacing of these eigenangles is thus $2\pi/n$. So if we wish to observe a non trivial behavior of the eigenangles we need to properly normalize the spacing between eigenvalues by multiplying the angles by a factor $n/2\pi$. This motivates our next proposition.

PROPOSITION 1.18. *Let $E_n$ be the set of random points, obtained by multiplying by $n/2\pi$ all eigenangles in $(-\pi, \pi]$ of $CUE(n)$ random matrices. Then for all $r \in \{1, \dots, n\}$, and for every measurable and bounded function $F$ from $\mathbb{R}^r$ to $\mathbb{R}$, we have:*

$$\mathbb{E}\left(\sum_{x_1 \neq \dots \neq x_r \in E_n} F(x_1, \dots, x_r)\right) = \int_{(-n/2, n/2]^r} F(y_1, \dots, y_r)\rho_r^{(n)}(y_1, \dots, y_r) dy_1 \dots dy_r,$$

*where*

$$\rho_r^{(n)}(y_1,\dots,y_r)=\det((K^{(n)}(y_j,y_k))_{1\le j,k\le r}),$$

*with*

$$K^{(n)}(y,y')=\frac{\sin[\pi(y'-y)]}{n\sin[\pi(y'-y)/n]}$$

*for $y'\neq y$ and*

$$K^{(n)}(y,y)=1.$$

PROOF. Define $G$ from $\mathbb{U}^r$ to $\mathbb{R}$ as

$$G(\lambda_1,\dots,\lambda_r)=F(x_1,\dots,x_r),$$

where for all $j\in\{1,\dots,r\}$,

$$x_j=n\eta_j/2\pi,$$

$\eta_j\in(-\pi,\pi]$ being the argument of $\lambda_j$. We have

$$\mathbb{E}\left(\sum_{x_1\neq\dots\neq x_r\in E_n}F(x_1,\dots,x_r)\right)=\mathbb{E}\left(\sum_{\lambda_1\neq\dots\neq\lambda_r\in E}G(\lambda_1,\dots,\lambda_r)\right),$$

where $E$ is the set of the eigenvalues of the considered $CUE(n)$ matrix. Using previous notation we can write:

$$\mathbb{E}\left(\sum_{x_1\neq\dots\neq x_r\in E_n}F(x_1,\dots,x_r)\right)=\int_{\mathbb{U}^r}G(\lambda_1,\dots,\lambda_r)\ \det((K(\theta_j,\theta_k))_{1\le j,k\le r})d\nu(\lambda_1)\dots d\nu(\lambda_n),$$

where $\theta_j\in(-\pi,\pi]$ is the argument of $\lambda_j$. Now we write the integral w.r. to $\theta_1,\dots\theta_n$ to obtain

$$\mathbb{E}\left(\sum_{x_1\neq\dots\neq x_r\in E_n}F(x_1,\dots,x_r)\right)=\int_{(-\pi,\pi]^r}G(e^{i\theta_1},\dots,e^{i\theta_r})\det((K(\theta_j,\theta_k))_{1\le j,k\le r})\frac{d\theta_1}{2\pi}\dots\frac{d\theta_r}{2\pi}.$$

Now we make the change of variables $y_j=n\theta_j/2\pi$ and get

$$\begin{aligned}&\mathbb{E}\left(\sum_{x_1\neq\dots\neq x_r\in E_n}F(x_1,\dots,x_r)\right)\\&=\int_{(-n/2,n/2]^r}G(e^{2i\pi y_1/n},\dots,e^{2i\pi y_r/n})\det((K(2\pi y_j/n,2\pi y_k/n))_{1\le j,k\le r})\frac{dy_1}{n}\dots\frac{dy_r}{n},\end{aligned}$$

that is

$$\begin{aligned}&\mathbb{E}\left(\sum_{x_1\neq\dots\neq x_r\in E_n}F(x_1,\dots,x_r)\right)\\&=\int_{(-n/2,n/2]^r}F(y_1,\dots,y_r)\det\left[\left(\frac{1}{n}K(2\pi y_j/n,2\pi y_k/n)\right)_{1\le j,k\le r}\right]dy_1\dots dy_r.\end{aligned}$$

Next we note that for $y\neq y'$ in $(-n/2,n/2]$, we have

$$K(2\pi y/n,2\pi y'/n)=\frac{\sin[\pi(y'-y)]}{\sin[\pi(y'-y)/n]},$$

and

$$K(2\pi y/n,2\pi y/n)=n,$$

which implies

$$\frac{1}{n}K(2\pi y/n,2\pi y'/n)=K^{(n)}(y,y').$$

□

The set of random points $E_n$ is also a determinantal point process with kernel $K^{(n)}$. The following celebrated convergence result is due to Dyson:

THEOREM 1.19 (Dyson). *For $y, y' \in \mathbb{R}$, we have:*

$$K^{(n)}(y,y') \underset{n\to\infty}{\longrightarrow} K^{(\infty)}(y,y'),$$

*where*

$$K^{(\infty)}(y,y') = \frac{\sin[\pi(y'-y)]}{\pi(y'-y)}$$

*for $y \neq y'$ and*

$$K^{(\infty)}(y,y) = 1.$$

The kernel $K^{(\infty)}$ which involves the function $x \mapsto (\sin x)/x$: is called the *sine kernel*. The convergence of $K^{(n)}$ towards $K^{(\infty)}$ hints at some convergence of the point process $E_n$ to a determinantal point process with kernel $K^{(\infty)}$. We now give a rigorous framework for this result. To this end we shall need the following lemma:

LEMMA 1.20. *For $1 \leq r \leq n$ and for all $y_1, \ldots, y_r \in (-n/2, n/2]$, we have*

$$0 \leq \rho_r^{(n)}(y_1, \ldots, y_r) \leq 1.$$

PROOF. A straightforward computation for geometric sums shows that for $y, y' \in (-n/2, n/2]$,

$$K^{(n)}(y,y') = \frac{1}{n} \sum_{j \in A_n} e^{ij\pi(y'-y)/n},$$

where $A_n$ is the set of integers between $-n$ et $n$, with a different parity than $n$. Hence $K^{(n)}(y,y')$ is equal to the usual Hermitian scalar product between the vectors $v_y$ et $v_{y'}$ of $\mathbb{C}^{A_n}$, where

$$v_y := \left(\frac{1}{\sqrt{n}} e^{ijy/n}\right)_{j \in A_n}.$$

Hence $\rho_r^{(n)}(y_1, \ldots, y_r)$ is the Gram determinant of the vectors $v_{y_1}, \ldots, v_{y_r}$, which implies

$$0 \leq \rho_r^{(n)}(y_1, \ldots, y_r) \leq ||v_{y_1}||^2 \ldots ||v_{y_r}||^2.$$

Since $A_n$ has $n$ elements and that all the coordinates of $v_{y_j}$ have modulus $1/\sqrt{n}$, these vectors have norm 1, and the result of the lemma follows.

□

We now state and prove the main convergence result.

THEOREM 1.21. *Let $E_n$ denote the set of eigenvalues taken in $(-\pi, \pi]$ and multiplied by $n/2\pi$ of a random unitary matrix of size $n$ following the Haar measure. Let us also define for $y \neq y^{'}$*

$$K^{(\infty)}(y,y^{'}) = \frac{\sin[\pi(y^{'}-y)]}{\pi(y^{'}-y)}$$

*and*

$$K^{(\infty)}(y,y) = 1.$$

*Then there exists a point process $E_\infty$ such that for all $r \in \{1, \ldots, n\}$, and for all Borel measurable and bounded functions $F$ with compact support from $\mathbb{R}^r$ to $\mathbb{R}$, we have, as $n \to \infty$,*

$$\mathbb{E}\left(\sum_{x_1 \neq \cdots \neq x_r \in E_n} F(x_1, \ldots, x_r)\right) \to \int_{\mathbb{R}^r} F(y_1, \ldots, y_r) \rho_r^{(\infty)}(y_1, \ldots, y_r) dy_1 \ldots dy_r,$$

*where*

$$\rho_r^{(\infty)}(y_1, \ldots, y_r) = \det((K^{(\infty)}(y_j, y_k))_{1 \leq j,k \leq r}).$$

*Moreover the point process $E_n$ converges to $E_\infty$ in the following sense: for all Borel measurable bounded functions $f$ with compact support from $\mathbb{R}$ to $\mathbb{R}$,*

$$\sum_{x \in E_n} f(x) \underset{n\to\infty}{\longrightarrow} \sum_{x \in E_\infty} f(x),$$

*where the convergence above holds in law.*

PROOF. In the course of the proof we shall use the above lemma which states that for $1 \leq r \leq n$, and for all $y_1, \dots, y_r \in (-n/2, n/2]$, we have

$$0 \leq \rho_r^{(n)}(y_1, \dots, y_r) \leq 1,$$

and that the correlation functions of a Poisson point process with intensity 1 are all equal to 1 (and hence are the correlation functions $\rho_r^{(n)}$ are dominated by those of a Poisson point process with intensity 1).

We first note the following identity: for any integer $p \geq 0$,

$$\left( \sum_{x \in E_n} f(x) \right)^p = \sum_{m=1}^{u_p} \sum_{x_1 \neq x_2 \neq \dots \neq x_{r_{p,m}} \in E_n} G_{f,p,m}(x_1, \dots, x_{r_{p,m}}),$$

where $u_p$ depends only on $p$, $r_{p,m}$ on $p$ and $m \leq u_p$, and $G_{f,p,m}$ being a measurable, bounded function with compact support from $\mathbb{R}^{r_{p,m}}$ to $\mathbb{R}$, and depending only on $f$, $p$ and $m$. For instance

$$\left( \sum_{x \in E_n} f(x) \right)^3 = \sum_{x_1 \in E_n} (f(x_1))^3 + 3 \sum_{x_1 \neq x_2 \in E_n} (f(x_1))^2 f(x_2) + \sum_{x_1 \neq x_2 \neq x_3 \in E_n} f(x_1) f(x_2) f(x_3),$$

with

$$u_3 = 3, r_{3,1} = 1, r_{3,2} = 2, r_{3,3} = 3,$$
$$G_{f,3,1}(x_1) = (f(x_1))^3,$$
$$G_{f,3,2}(x_1, x_2) = 3(f(x_1))^2 f(x_2),$$
$$G_{f,3,3}(x_1, x_2, x_3) = f(x_1) f(x_2) f(x_3).$$

We can hence write

$$\mathbb{E}\left[ \left( \sum_{x \in E_n} f(x) \right)^p \right] = \sum_{m=1}^{u_p} \int_{(-n/2, n/2]^{r_{p,m}}} G_{f,p,m}(y_1, \dots, y_{r_{p,m}}) \rho_{r_{p,m}}^{(n)}(y_1, \dots, y_{r_{p,m}}) \, dy_1 \dots dy_{r_{p,m}},$$

provided the above expression converges absolutely, which we now check. Since $G_{f,p,m}$ is measurable, bounded with compact support, we can find $A_{f,p,m} > 0$ such that

$$|G_{f,p,m}(y_1, \dots, y_{r_{p,m}})| \leq A_{f,p,m} \mathbb{1}_{|y_1|, \dots, |y_{r_{p,m}}| \leq A_{f,p,m}}$$

for $y_1, \dots, y_{r_{p,m}} \in \mathbb{R}$. Moreover from the remark above on the correlation functions, we have

$$|\rho_{r_{p,m}}^{(n)}(y_1, \dots, y_{r_{p,m}})| \mathbb{1}_{y_1, \dots, y_{r_{p,m}} \in (-n/2, n/2]} \leq 1.$$

Consequently the expression we are dealing with can be bounded from above by

$$\sum_{m=1}^{u_p} \int_{[-A_{f,p,m}, A_{f,p,m}]^{r_{p,m}}} A_{f,p,m} \leq \sum_{m=1}^{u_p} (2A_{f,p,m})^{r_{p,m}+1},$$

which is finite. Moreover our upper bound is independent of $n$. Now since the kernel $K^{(n)}$ converges pointwise to $K^{(\infty)}$, we also have

$$\rho_{r_{p,m}}^{(n)}(y_1, \dots, y_{r_{p,m}}) \mathbb{1}_{y_1, \dots, y_{r_{p,m}} \in (-n/2, n/2]} \underset{n \to \infty}{\longrightarrow} \rho_{r_{p,m}}^{(\infty)}(y_1, \dots, y_{r_{p,m}}),$$

and we can apply the dominated convergence theorem to obtain

$$\mathbb{E}[(X_f^{(n)})^p] \underset{n \to \infty}{\longrightarrow} M_{f,p}^{(\infty)}$$

where

$$X_f^{(n)} = \sum_{x \in E_n} f(x)$$

and

$$M_{f,p}^{(\infty)} = \sum_{m=1}^{u_p} \int_{\mathbb{R}^{r_{p,m}}} G_{f,p,m}(y_1, \dots, y_{r_{p,m}}) \rho_{r_{p,m}}^{(\infty)}(y_1, \dots, y_{r_{p,m}}) \, dy_1 \dots dy_{r_{p,m}}.$$

We also note that

$$\mathbb{E}[|X_f^{(n)}|^p] \leq \mathbb{E}[(X_{|f|}^{(n)})^p] \leq \mathbb{E}\left[ \left( \sum_{x \in N} |f(x)| \right)^p \right],$$

where $N$ is a Poisson point process defined on $\mathbb{R}$ with intensity 1 (the last inequality follows from the fact that the correlation functions of $E_n$ are smaller or equal than 1, and hence smaller or equal than the correlation functions of $N$).

Now for every $\lambda \in \mathbb{R}$, each term of the series

$$\sum_{p\geq 0} \frac{(i\lambda)^p}{p!} \mathbb{E}[(X_f^{(n)})^p]$$

is uniformly dominated in absolute value and independently of $n$, by the corresponding term in the series

$$\sum_{p\geq 0} \frac{|\lambda|^p}{p!} \mathbb{E}\left[\left(\sum_{x\in N} |f(x)|\right)^p\right] = \mathbb{E}\left[\exp\left(|\lambda| \sum_{x\in N} |f(x)|\right)\right].$$

If we choose $A_f > 0$ in such a way that $|f| \leq A_f$ and such that the support of $f$ is contained in $[-A_f, A_f]$, we have

$$\mathbb{E}\left[\exp\left(|\lambda| \sum_{x\in N} |f(x)|\right)\right] \leq \mathbb{E}\left[\exp\left(|\lambda| A_f \operatorname{Card}(N \cap [-A_f, A_f])\right)\right] = \mathbb{E}[e^{|\lambda| A_f Y_{2A_f}}],$$

$Y_{2A_f}$ standing for a Poisson random variable with parameter $2A_f$. The latter is finite and we can thus apply the dominated convergence theorem to obtain

$$\mathbb{E}\left[e^{i\lambda X_f^{(n)}}\right] \underset{n\to\infty}{\longrightarrow} \sum_{p\geq 0} \frac{(i\lambda)^p}{p!} M_{f,p}^{(\infty)},$$

the last series in display being absolutely convergent and bounded from above by

$$\begin{aligned}\left|1 - \sum_{p\geq 0} \frac{(i\lambda)^p}{p!} M_{f,p}^{(\infty)}\right| &\leq \sum_{p\geq 1} \frac{|\lambda|^p}{p!} M_{|f|,p}^{(\infty)} \leq \sum_{p\geq 1} \frac{|\lambda|^p}{p!} \sup_{n\geq 1} \mathbb{E}[|X_f^{(n)}|^p] \\ &\leq \sum_{p\geq 1} \frac{|\lambda|^p}{p!} \mathbb{E}\left[\left(\sum_{x\in N} |f(x)|\right)^p\right] = \mathbb{E}\left[\exp\left(|\lambda| \sum_{x\in N} |f(x)|\right)\right] - 1 \\ &\leq \mathbb{E}[e^{|\lambda| A_f Y_{2A_f}}] - 1 = e^{2A_f(e^{|\lambda| A_f} - 1)} - 1.\end{aligned}$$

Consider now a finite number $f_1, f_2, \dots, f_q$ of measurable and bounded functions with compact support, and let $A > 0$ be such that $|f_j| \leq A\mathbb{1}_{[-A,A]}$ for $j \in \{1, \dots, q\}$, and take $\lambda, \lambda_1, \dots, \lambda_q \in \mathbb{R}$. It follows from the definition of $X_f^{(n)}$ that

$$\sum_{j=1}^{q} \lambda_j X_{f_j}^{(n)} = X_g^{(n)}$$

where

$$g := \sum_{j=1}^{q} \lambda_j f_j,$$

which implies that

$$\mathbb{E}\left[e^{i\lambda \sum_{j=1}^q \lambda_j X_{f_j}^{(n)}}\right] \underset{n\to\infty}{\longrightarrow} \sum_{p\geq 0} \frac{(i\lambda)^p}{p!} M_{g,p}^{(\infty)}.$$

Now since $g$ is bounded by $A \sum_{j=1}^q |\lambda_j|$ and since the support of $g$ is included in $[-A, A]$, we have

$$\left|1 - \sum_{p\geq 0} \frac{(i\lambda)^p}{p!} M_{g,p}^{(\infty)}\right| \leq e^{2A(1+\sum_{j=1}^q |\lambda_j|)\left(e^{|\lambda| A(1+\sum_{j=1}^q |\lambda_j|)} - 1\right)} - 1.$$

If $\nu_1, \dots, \nu_q$ are real numbers not all equal to zero, we set

$$\lambda = \sum_{j=1}^{q} |\nu_j|, \lambda_j = \nu_j / \lambda,$$

which implies $\sum_{j=1}^{q} |\lambda_j| = 1$, and

$$\mathbb{E}\left[e^{i\sum_{j=1}^{q} \nu_j X_{f_j}^{(n)}}\right] \underset{n\to\infty}{\longrightarrow} Q(f_1,\dots,f_q,\nu_1,\dots,\nu_q)$$

where

$$|Q(f_1,\dots,f_q,\nu_1,\dots,\nu_q) - 1| \leq e^{4A(e^{2|\lambda|A}-1)} - 1.$$

For fixed $f_1,\dots,f_q$, the quantity $Q(f_1,\dots,f_q,\nu_1,\dots,\nu_q)$ hence tends to 1 when $(\nu_1,\dots,\nu_q)$ tends to zero. It follows from Lévy's convergence theorem that the vector $(X_{f_1}^{(n)},\dots,X_{f_q}^{(n)})$ converges in law, when $n$ goes to infinity, to a random variable with values in $\mathbb{R}^q$ with characteristic function given by

$$(\nu_1,\dots,\nu_q) \mapsto Q(f_1,\dots,f_q,\nu_1,\dots,\nu_q).$$

Consequently we have shown that there exists a family of random variables $(X_f^{(\infty)})_{f\in\mathcal{A}}$ indexed by the set $\mathcal{A}$ of bounded and measurable functions with compact support from $\mathbb{R}$ to $\mathbb{R}$, satisfying

$$(X_f^{(n)})_{f\in\mathcal{A}} \underset{n\to\infty}{\longrightarrow} (X_f^{(\infty)})_{f\in\mathcal{A}},$$

in the sense of finite dimensional distributions. Now for $x \geq 0$, define

$$X^{(n)}(x) = X_{\mathbb{1}_{[0,x]}}^{(n)},\ X^{(\infty)}(x) = X_{\mathbb{1}_{[0,x]}}^{(\infty)},$$

and for $x < 0$,

$$X^{(n)}(x) = -X_{\mathbb{1}_{(x,0)}}^{(n)},\ X^{(\infty)}(x) = -X_{\mathbb{1}_{(x,0)}}^{(\infty)}.$$

Note that for $y \geq x$, $X^{(n)}(y) - X^{(n)}(x)$ represents the number of points of $E_n$ in the interval $(x,y]$. Moreover we saw that $(X^{(n)}(x))_{x\in\mathbb{Q}}$ converges in law (in the sense of finite dimensional distributions) to $(X^{(\infty)}(x))_{x\in\mathbb{Q}}$. It follows from Skorokhod's representation theorem that there exist random variables $(Y^{(n)}(x))_{x\in\mathbb{Q}}$ and $(Y^{(\infty)}(x))_{x\in\mathbb{Q}}$, with respectively the same distributions as $(X^{(n)}(x))_{x\in\mathbb{Q}}$ and $(X^{(\infty)}(x))_{x\in\mathbb{Q}}$, such that almost surely, $Y^{(n)}(x)$ converges to $Y^{(\infty)}(x)$ for all $x \in \mathbb{Q}$. By construction $(Y^{(n)}(x))_{x\in\mathbb{Q}}$ is almost surely integer valued and increasing as a function of $x$: the same thing holds for $(Y^{(\infty)}(x))_{x\in\mathbb{Q}}$. Moreover by taking the limits from the right, we can extend $(Y^{(n)}(x))_{x\in\mathbb{Q}}$ et $(Y^{(\infty)}(x))_{x\in\mathbb{Q}}$ to càdlàg functions defined on $\mathbb{R}$. It is clear that $(Y^{(n)}(x))_{x\in\mathbb{R}}$ has then the same law as $(X^{(n)}(x))_{x\in\mathbb{R}}$, because $(X^{(n)}(x))_{x\in\mathbb{R}}$ is also càdlàg, with the same law when restricted to $\mathbb{Q}$. We can thus conclude that like for $(X^{(n)}(x))_{x\in\mathbb{R}}$, $(Y^{(n)}(x))_{x\in\mathbb{R}}$ is also the distribution function of some $\sigma$-finite measure $\mathcal{M}_n$, with the same law as the sum of the Dirac measures taken at the points of $E_n$. Almost surely, for $x \in \mathbb{Q}$, $(Y^{(n)}(x))$ converges to $(Y^{(\infty)}(x))$: hence this convergences also holds at all continuity points of $(Y^{(\infty)}(x))$. Consequently $\mathcal{M}_n$ converges weakly, in the sense of convergence in law on compact subsets, to a limiting random measure $\mathcal{M}_\infty$, with distribution function $Y^{(\infty)}$. On can thus write

$$\mathcal{M}_n = \sum_{k\in\mathbb{Z}} \delta_{t_k^{(n)}},\ \mathcal{M}_\infty = \sum_{k\in\mathbb{Z}} \delta_{t_k^{(\infty)}},$$

where $\{t_k^{(n)}, k \in \mathbb{Z}\}$ is a set of points with the same distribution as $E_n$. The weak convergence of $\mathcal{M}_n$ to $\mathcal{M}_\infty$ implies that for $r \geq 0$, $F$ continuous with compact support from $\mathbb{R}^r$ to $\mathbb{R}$,

$$\sum_{k_1\neq k_2\neq\dots\neq k_r} F(t_{k_1}^{(n)},\dots,t_{k_r}^{(n)}) \underset{n\to\infty}{\longrightarrow} \sum_{k_1\neq k_2\neq\dots\neq k_r} F(t_{k_1}^{(\infty)},\dots t_{k_r}^{(\infty)}).$$

Indeed the left hand side can be written as:

$$\sum_{m=1}^{u_r} \int_{\mathbb{R}^{s_{r,m}}} H_{F,r,m}(y_1,\dots,y_{s_{r,m}}) d\mathcal{M}_n(y_1)\dots d\mathcal{M}_n(y_{s_{r,m}}),$$

where $u_r$ depends only on $r$, $s_{r,m}$ on $r$ and on $m$ and $H_{F,r,m}$ depends on $F, r, m$, and the right hand side can be written in similar way with $\mathcal{M}_n$ replaced with $\mathcal{M}_\infty$. For instance

$$\begin{aligned}\sum_{k_1 \neq k_2 \neq k_3} F(t_{k_1}^{(n)}, \dots, t_{k_3}^{(n)}) &= \int_{\mathbb{R}^3} F(y_1, y_2, y_3) d\mathcal{M}_n(y_1) d\mathcal{M}_n(y_2) d\mathcal{M}_n(y_3) \\ &\quad - \int_{\mathbb{R}^2} [F(y_1, y_2, y_2) + F(y_2, y_1, y_2) + F(y_1, y_2, y_2)] d\mathcal{M}_n(y_1) d\mathcal{M}_n(y_2) \\ &\quad + 2 \int_{\mathbb{R}} F(y_1, y_1, y_1) d\mathcal{M}_n(y_1).\end{aligned}$$

If we assume that $F$ is positive, it follows from Fatou's lemma that

$$\begin{aligned}\mathbb{E}\left[\sum_{k_1 \neq k_2 \neq \dots \neq k_r} F(t_{k_1}^{(\infty)}, \dots t_{k_r}^{(\infty)})\right] &\leq \liminf_{n \to \infty} \mathbb{E}\left[\sum_{k_1 \neq k_2 \neq \dots \neq k_r} F(t_{k_1}^{(n)}, \dots t_{k_r}^{(n)})\right] \\ &= \liminf_{n \to \infty} \int_{\mathbb{R}^r} F(y_1, \dots, y_r) \rho_r^{(n)}(y_1, \dots, y_r) dy_1 \dots dy_r \\ &= \int_{\mathbb{R}^r} F(y_1, \dots, y_r) \rho_r^{(\infty)}(y_1, \dots, y_r) dy_1 \dots dy_r.\end{aligned}$$

Reproducing the same computations as in the beginning of our proof yields, for $f$ continuous and positive with compact support, and $p$ a positive integer:

$$N_{f,p}^{(\infty)} := \mathbb{E}\left[\left(\sum_{k \in \mathbb{Z}} f(t_k^{(\infty)})\right)^p\right] \leq M_{f,p}^{(\infty)}.$$

The bounds that we previously obtained for $M_{f,p}^{(\infty)}$, and which obviously apply to $N_{f,p}^{(\infty)}$ as well, allow us to deduce that for all $\lambda \in \mathbb{R}$,

$$\mathbb{E}\left[\exp\left(i\lambda \sum_{k \in \mathbb{Z}} f(t_k^{(\infty)})\right)\right] = \sum_{p \geq 0} \frac{(i\lambda)^p}{p!} N_{f,p}^{(\infty)}.$$

Moreover an application of the dominated convergence theorem yields

$$\begin{aligned}\mathbb{E}\left[\exp\left(i\lambda \sum_{k \in \mathbb{Z}} f(t_k^{(\infty)})\right)\right] &= \lim_{n \to \infty} \mathbb{E}\left[\exp\left(i\lambda \sum_{k \in \mathbb{Z}} f(t_k^{(n)})\right)\right] \\ &= \lim_{n \to \infty} \mathbb{E}\left[e^{i\lambda X_f^{(n)}}\right] = \sum_{p \geq 0} \frac{(i\lambda)^p}{p!} M_{f,p}^{(\infty)}.\end{aligned}$$

We can hence conclude that the coefficients of both series in $\lambda$ are equal, i.e. $M_{f,p}^{(\infty)} = N_{f,p}^{(\infty)}$. Going back to the expression of the expansion of the moment of order $p$ that we gave earlier in the proof, we see that the equality can hold only if

$$\mathbb{E}\left[\sum_{k_1 \neq k_2 \neq \dots \neq k_r} F(t_{k_1}^{(\infty)}, \dots t_{k_r}^{(\infty)})\right] = \int_{\mathbb{R}^r} F(y_1, \dots, y_r) \rho_r^{(\infty)}(y_1, \dots, y_r) dy_1 \dots dy_r,$$

for all $F, r$ such that $r = r_{p,m}, F = G_{f,p,m}$, with $1 \leq m \leq u_p$. Indeed the left hand side is always smaller or equal than the right hand side, and if one of the inequalities were a strict inequality, we would obtain by summing up all terms that $N_{f,p}^{(\infty)} < M_{f,p}^{(\infty)}$. The only term for which $r_{p,m} = p$ gives

$$\mathbb{E}\left[\sum_{k_1 \neq k_2 \neq \dots \neq k_p} F(t_{k_1}^{(\infty)}, \dots t_{k_p}^{(\infty)})\right] = \int_{\mathbb{R}^p} F(y_1, \dots, y_p) \rho_p^{(\infty)}(y_1, \dots, y_p) dy_1 \dots dy_p,$$

where

$$F(y_1, \dots, y_p) = f(y_1) \dots f(y_p).$$

This then extends to all functions $F$ which are measurable, positive and continuous with compact support: indeed there always exists an $f$ which is continuous with compact support on $\mathbb{R}$ such that $F \leq G$, with

$$G(y_1, \dots, y_p) = f(y_1) \dots f(y_p).$$

Since we have inequalities for both functions $F$ et $G - F$ and an equality for their sum $G$, we in fact have an equality everywhere.

We see that with a monotone class argument the previous equality then extends to functions $F$ which are measurable, bounded and with compact support. This shows the existence of a point process $E_\infty$ with the same correlation functions as those given in the statement of the Proposition, provided we do not exclude a priori point processes with multiple points. More precisely we take for $E_\infty$ the set of points $t_k^{(\infty)}$ of the support of the measure $\mathcal{M}_\infty$, taken with their multiplicities. Going back to our earlier computations, we see that for functions $f$ which are measurable, bounded with compact support from $\mathbb{R}$ to $\mathbb{R}$, and taking into account multiplicities, we have:

$$\mathbb{E}\left[\left(\sum_{x \in E_\infty} f(x)\right)^p\right] = M_{f,p}^{(\infty)}$$

and

$$\mathbb{E}\left[\exp\left(i\lambda \sum_{x \in E_\infty} f(x)\right)\right] = \sum_{p \geq 0} \frac{(i\lambda)^p}{p!} M_{f,p}^{(\infty)}.$$

Consequently

$$\mathbb{E}\left[e^{i\lambda X_f^{(n)}}\right] \underset{n\to\infty}{\longrightarrow} \mathbb{E}\left[\exp\left(i\lambda \sum_{x \in E_\infty} f(x)\right)\right],$$

which corresponds to the convergence in law stated in the Proposition.

It only remains to show that $E_\infty$ does not have multiple points. Indeed, if $E_\infty$ is the set of points $(t_k^{(\infty)})_{k\in\mathbb{Z}}$, taken with multiplicities, then for any mesurable bounded function $F$ with compact support from $\mathbb{R}^2$ in $\mathbb{R}$,

$$\mathbb{E}\left(\sum_{k_1 \neq k_2} F(t_{k_1}^{(\infty)}, t_{k_2}^{(\infty)})\right) = \int_{\mathbb{R}^2} F(y_1, y_2)\rho_2^{(\infty)}(y_1, y_2)dy_1dy_2.$$

Taking $F(y_1, y_2) = \mathbb{1}_{y_1 = y_2}$ above yields

$$\mathbb{E}\left[\mathrm{Card}\left\{(k_1, k_2) \in \mathbb{Z}^2, k_1 \neq k_2, t_{k_1}^{(\infty)} = t_{k_2}^{(\infty)}\right\}\right] = 0,$$

which shows that $E_\infty$ does almost surely not have multiple points.

□

The point process $E_\infty$ is called the sine kernel determinantal point process. It appears in many contexts in random matrix theory (e.g. universality results).

PROPOSITION 1.22. *The formula given in the previous proposition uniquely characterizes the distribution of $E_\infty$. This distribution is invariant under translation: for all $y \in \mathbb{R}$, the image of $E_\infty$ under a translation by $y$ has the same distribution as $E_\infty$. The expected number of points of $E_\infty$ in an interval $[a, b]$ is $b - a$, and its variance can be bounded by $\log(b - a)$ if $b - a \geq 2$.*

In comparison, a Poisson point process with intensity 1 has variance $b - a$. The 2-point correlation function $\rho_2(x, y)$ of $E_\infty$ is equivalent to $\pi^2(y - x)^2/3$ when $y - x$ goes to zero, which shows that the points of $E_\infty$ tend to repel each other on small scales.

PROOF. Going through the beginning of the proof of the previous proposition, one checks that the formula giving the correlation functions actually characterizes the finite dimensional distributions of

$$\sum_{x \in E_\infty} f(x),$$

for $f$ measurable, bounded and with compact support. In particular this gives the finite dimensional distributions of the càdlàg process

$$y \mapsto \sum_{x \in E_\infty} \mathbb{1}_{y \geq 0, x \in [0,y]} - \sum_{x \in E_\infty} \mathbb{1}_{y < 0, x \in (y,0)},$$

and hence the distribution of the stochastic process itself in the space of càdlàg functions from $\mathbb{R}$ to $\mathbb{R}$. Consequently the distribution of $E_\infty$ is characterized for the points in $E_\infty$ are the time at which the process jumps.

Now the translation invariance of the law of $E_\infty$ is an immediate consequence of the fact that the correlation functions characterize the law of $E_\infty$ and that for all $r \geq 1, y, y_1, \dots, y_r \in \mathbb{R}$,

$$\rho_r^{(\infty)}(y_1 + y, \dots, y_r + y) = \rho_r^{(\infty)}(y_1, \dots, y_r),$$

which itself follows from:

$$K^{(\infty)}(y_1 + y, y_2 + y) = K^{(\infty)}(y_1, y_2).$$

The computation for the expected number of points $N_{a,b}$ of $E_\infty \cap [a, b]$ is a straightforward consequence of $\rho_1 \equiv 1$. To estimate the variance, we write

$$\mathbb{E}[(N_{a,b})^2] = \mathbb{E}\left[N_{a,b} + \sum_{x_1 \neq x_2 \in E_\infty} \mathbb{1}_{x_1, x_2 \in [a,b]}\right] = (b - a) + \int_{[a,b]^2} \rho_2(y_1, y_2) dy_1 dy_2,$$

and hence

$$\begin{aligned}\mathrm{Var}[N_{a,b}] &= (b - a) - (b - a)^2 + \int_{[a,b]^2} \rho_2(y_1, y_2) dy_1 dy_2 \\ &= (b - a) - \int_{[a,b]^2} [K^{(\infty)}(y_1, y_2)]^2 dy_1 dy_2.\end{aligned}$$

But for all $y \in \mathbb{R}$, we have

$$K^{(\infty)}(0, y) = \frac{\sin(\pi y)}{\pi y} = \int_{-1/2}^{1/2} e^{2i\pi y t} dt,$$

that is $y \mapsto K^{(\infty)}(0, y)$ is the Fourier transform of the indicator function of the interval $[-1/2, 1/2]$. An application of Plancherel's formula and the translation invariance of $K^{(0)}$ implies that for all $y_1 \in \mathbb{R}$,

$$\int_{\mathbb{R}} [K^{(\infty)}(y_1, y_2)] dy_2 = 1,$$

and hence

$$\int_{[a,b] \times \mathbb{R}} [K^{(\infty)}(y_1, y_2)]^2 dy_1 dy_2 = (b - a).$$

Consequently we can write

$$\mathrm{Var}[N_{a,b}] = \int_{[a,b] \times (\mathbb{R} \setminus [a,b])} [K^{(\infty)}(y_1, y_2)]^2 dy_1 dy_2,$$

and by symmetry,

$$\mathrm{Var}[N_{a,b}] = 2 \int_a^b \int_b^\infty [K^{(\infty)}(y_1, y_2)]^2 dy_1 dy_2.$$

Now we know

$$[K^{(\infty)}(y_1, y_2)]^2 \leq \left(1 \wedge [1/(y_2 - y_1)^2]\right),$$

hence

$$\mathrm{Var}[N_{a,b}] \leq 2 \int_a^b \int_b^\infty \frac{dy_1 dy_2}{1 \vee (y_2 - y_1)^2}.$$

For $y_1 \leq b - 1$, we have

$$\int_b^\infty \frac{dy_2}{1 \vee (y_2 - y_1)^2} = \int_b^\infty \frac{dy_2}{(y_2 - y_1)^2} = \frac{1}{b - y_1}$$

and for $b - 1 \leq y_1 \leq b$,

$$\int_b^\infty \frac{dy_2}{1 \vee (y_2 - y_1)^2} = 2 + y_1 - b \leq 2.$$

Finally

$$\mathrm{Var}[N_{a,b}] \leq 4 \int_a^b \frac{dy_1}{1 \vee (b - y_1)} = O(\log(b - a)),$$

if $b - a \geq 2$.

□

EXAMPLE 1.23 (Exercise). *Let $M \in CUE(n)$ and note $\tilde{\theta}_j = \frac{n}{2\pi}\theta_j$. For suitable test functions, prove that*

$$\frac{1}{n}\mathbb{E}\left[\sum_{j \neq k} f(\tilde{\theta}_j - \tilde{\theta}_k)\right] \underset{N \to \infty}{\to} \int_{-\infty}^{\infty} f(v) \left[1 - \left(\frac{\sin(\pi v)}{\pi v}\right)^2\right] dv.$$

## 2. Virtual Isometries

In this section we shall freely use well known results from uniformly distributed vectors on spheres and distributional properties of elements of random unitary matrices (all the facts we shall need are explained in detail in Appendix 2 and Appendix 3).

**2.1. How to generate the Haar measure.** Now we wish to address the question of generating Haar distributed unitary matrices recursively. We build our intuition on the special case of uniformly distributed random permutations. A permutation is a bijection $\sigma \in \mathfrak{S}_n$ of $\{1, \cdots, n\}$

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

with $\#\mathfrak{S} = n!$. The Haar (or uniform) measure is defined as

$$\mu(\sigma) = \frac{1}{n!}.$$

A transposition $\tau_{ij}$ is an element of $\mathfrak{S}_n$ such that

$$\tau(j) = i, \quad \tau(k) = k, \quad \text{for } k \notin \{i, j\}.$$

LEMMA 2.1. *Let $\sigma \in \mathfrak{S}_\mathfrak{n}$. There exists a unique decomposition of $\sigma$ of the following form*

$$\sigma = \tau_{n,m_n} \circ \tau_{n-1,m_{n-1}} \circ \cdots \circ \tau_{1,m_1}$$

*where $\tau_{h,m_h}$ is the transposition which commutes $h$ and $m_h$ with $m_h \in \{1, \cdots, h\}$.*

PROOF. A simple induction. □

LEMMA 2.2. *For $k \in \{1, \cdots, n\}$, let $\tau_{k,m_k}$ be randomly and independently chosen according to $\mathbb{P}[m_k = j] = 1/k$, for $j \in \{1, \cdots, k\}$. Then $\sigma = \tau_{n,m_n} \circ \tau_{n-1,m_{n-1}} \circ \cdots \circ \tau_{1,m_1}$ is uniformly distributed.*

It follows from a simple application of the definition of the Haar measure that if a matrix $M \in CUE(n)$, then its first column is uniformly distributed on the unit sphere $\mathscr{S}_\mathbb{C}^n$ of $\mathbb{C}^n$ (see Appendix 2 and Appendix 3 for background on uniformly distributed vectors on spheres and distribution of elements of a random unitary matrix).

PROPOSITION 2.3. *Let $M \in U(N+1)$ be such that its first column $M_1$ is uniformly distributed on $\mathscr{S}_\mathbb{C}^{N+1}$. If $V_N \in U(N)$ is chosen independently of $M$ according to the Haar measure $\mu_{U(N)}$, then the matrix*

$$V_{N+1} := M \begin{pmatrix} 1 & 0 \\ 0 & V_N \end{pmatrix}$$

*is distributed with the Haar measure $\mu_{U(N+1)}$.*

PROOF. Due to the uniqueness property of the Haar measure, we only need to show that for a fixed $U \in U(N+1)$

$$UM \begin{pmatrix} 1 & 0 \\ 0 & V_N \end{pmatrix} \overset{\text{law}}{=} M \begin{pmatrix} 1 & 0 \\ 0 & V_N \end{pmatrix}.$$

In the following, a matrix $A$ will often be written $(A_1 \| \tilde{A})$, where $A_1$ is its first column. As $U \in U(N+1)$, $(UM)_1 = UM_1$ is distributed uniformly on the complex unit sphere $\mathscr{S}_\mathbb{C}^{N+1}$, so we can write $UM = (P_1 \| \tilde{P})$, with $P_1$ uniformly distributed on $\mathscr{S}_\mathbb{C}^{N+1}$ and $\tilde{P}$ having a distribution on the orthogonal hyperplane of $P_1$. We then need to show that

$$(P_1 \| \tilde{P}) \begin{pmatrix} 1 & 0 \\ 0 & V_N \end{pmatrix} \overset{\text{law}}{=} (M_1 \| \tilde{M}) \begin{pmatrix} 1 & 0 \\ 0 & V_N \end{pmatrix},$$

where all matrices are still independent. As $M_1$ and $P_1$ are identically distributed, by conditioning on $M_1 = P_1 = v$ (here $v$ is any fixed element of $\mathscr{S}_\mathbb{C}^{N+1}$) it is sufficient to show that

$$(v \| P') \begin{pmatrix} 1 & 0 \\ 0 & V_N \end{pmatrix} \overset{\text{law}}{=} (v \| M') \begin{pmatrix} 1 & 0 \\ 0 & V_N \end{pmatrix},$$

where $M'$ (resp $P'$) is distributed like $\tilde{M}$ (resp $\tilde{P}$) conditionally to $M_1 = v$ (resp $P_1 = v$). Let $A$ be any element of $U(N+1)$ such that $A(v) = (1, 0, \ldots, 0)$. Since $A$ is invertible, we just need to show that

$$A(v \| P') \begin{pmatrix} 1 & 0 \\ 0 & V_N \end{pmatrix} \overset{\text{law}}{=} A(v \| M') \begin{pmatrix} 1 & 0 \\ 0 & V_N \end{pmatrix},$$

that is to say

$$P''V_N \overset{\text{law}}{=} M''V_N,$$

where $P''$ and $M''$ are distributed on $U(N)$ independently of $V_N$. By independence and conditioning on $P''$ (resp $M''$), we get $P''V_N \overset{\text{law}}{=} V_N$ (resp $M''V_N \overset{\text{law}}{=} V_N$) by definition of the Haar measure $\mu_{U(N)}$. This gives the desired result.

□

Now we need to make a natural choice for the transformation $M$: we need a simple unitary transformation which maps the first vector of the basis to a uniformly distributed vector on the complex unit sphere. This is the topic of next section.

**2.2. Reflections over $\mathbb{R}$.** We will begin by briefly recalling the definition of reflections over $\mathbb{R}$ since this is the one which is used if one wants to carry our construction to the orthogonal group. We also wish to state them here in order to understand why real reflections or Householder transformations would not be suitable if the ground field is the field of complex numbers.

DEFINITION 2.4. *Let $H$ denote a real vector space of dimension $n$. A* ***reflection*** *of $H$ (sometimes called a* ***Householder transformation*** *or an* ***elementary reflector****) is defined as an orthogonal transformation of $H$ which fixes each element of a hyperplane (i.e. linear subspace of codimension 1).*

Let $r$ be a reflection. Let us assume that $r \neq 1$ so that $\ker(1-r) = K$ is a hyperplane and $\dim \operatorname{im}(1-r) = 1$.

Since $r \in O(\mathbb{R})$ we must have $\det r \in \{\pm 1\}$. Clearly 1 is an eigenvalue with multiplicity $n-1$. Let $\lambda \neq 1$ denote the other eigenvalue. As $\det r$ is the product of the eigenvalues we conclude $\lambda = -1$ and $\det r = -1$. Consequently there exists a vector $a \in H$ such that $r(a) = -a$ and note that $r$ is a map of order 2 (namely $r^2 = 1$). We will write $r_a$ for the reflection which maps $a \mapsto -a$.

It is easy to find a formula for $r_a$: First note that $\mathbb{R}a \oplus (\mathbb{R}a)^\perp = H$. There exists $\phi \in H^*$ (where $H^*$ is the space of linear forms on $H$) such that $x - r_a(x) = \phi(x)a$ for every $x \in H$ and $\ker \phi = \ker(1 - r_a) = K$. Since $\ker(\langle \cdot, a\rangle) = K$, there exists a non-zero $\lambda \in \mathbb{C}^*$ such that $\phi = \lambda\langle \cdot, a\rangle$. Therefore $x - r_a(x) = \lambda\langle x, a\rangle a$. Now $r_a(a) = -a$ so we have $\lambda = \frac{2}{\langle a,a\rangle}$, hence

$$r_a(x) = x - 2\frac{\langle x, a\rangle}{\langle a, a\rangle}a. \tag{9}$$

This is the equation of a real orthogonal reflection that sends $a$ to $-a$ and hence a householder transformation is parametrized by a vector $a$.

Finally, we note that if $m, e \in H$ are two distinct vectors with $\|m\| = \|e\| = 1$, there exists a unique reflection which maps $m$ on $e$, namely $r_{m-e}$. But it is easy to see that given two vectors of norm 1 in a complex Hilbert space, then there does not necessarily exist a Householder transformation (9) which maps one on the other. Hence we need to introduce another type of reflection.

**2.3. Reflections over $\mathbb{C}$.** Over the complex numbers the theory is different since the reflections we consider are not necessarily of finite order and they are parametrized by one vector and a phase (that is, a complex number of modulus 1).

We now assume we are given a Hilbert space $H$ with $\dim H = n$. For $u, v \in H$, we use the standard inner product

$$\langle u, v\rangle := \sum_{k=1}^{n} u_k \overline{v}_k.$$

If $F \subseteq H$, $F$ a subset of $H$, we write $F^\perp := \{x \in H \mid \langle u, x\rangle = 0 \text{ for all } u \in F\}$.

If $F$ and $G$ are subspaces of $H$, we write $H = F \oplus^\perp G$ to indicate $H = F + G = \{x + y \mid x \in F, y \in G\}$ and $\langle x, y\rangle = 0$ for all $x \in F$ and $y \in G$. It is then easy to check that $F = G^\perp$ and $(G^\perp)^\perp = G$ and $\dim G + \dim G^\perp = \dim H = n$.

Let $U(H)$ denote the set of unitary operators, i.e. linear bijections $u : H \to H$ which preserve the inner product: $\langle ux, uy\rangle = \langle x, y\rangle$ for every $x, y \in H$. Let 1 denote the identity map. We first recall the elementary and well known result:

LEMMA 2.5. *If $u \in U(H)$ then* $\operatorname{im}(1-u) = \ker(1-u)^\perp$.

DEFINITION 2.6. *Let $u$ denote a linear transform of a complex vector space $H$ of finite dimension. We call $u$ a **complex reflection** if it is the identity or if it is unitary and* $\operatorname{rank}(1-u) = 1$. *We will just write **reflection** if the base field is clear.*

REMARK 2.7. *Usually in the literature, a linear transformation $g \in GL(H)$ is a reflection if the order of $g$ is finite and* $\operatorname{rank}(1-u) = 1$. *For our applications we do not require that $g$ have finite order.*

As in the real case, we can compute a formula for a general complex reflection.

PROPOSITION 2.8. *Suppose that $r$ is a reflection of the space $H$ and that the vector $a \in H$ spans* $\operatorname{im}(1-r)$. *Then there exists $\alpha \in \mathbb{U}$ such that for every $x \in H$,*

$$r(x) = x - (1-\alpha)\frac{\langle x, a\rangle}{\langle a, a\rangle}a.$$

PROOF. We have by construction and Lemma 2.5 the equivalences $\operatorname{im}(1-u) = \mathbb{C}a$ and $\ker(1-u) = (\mathbb{C}a)^{\perp}$. Let $\phi \in H^*$ denote the linear form defined by $u(x) = x - \phi(x)a$. Clearly $\phi(x) = 0$ if and only if $(1-u)x = 0$, so $\ker\phi = \ker(1-u)$. But the linear form $\langle\cdot, a\rangle$ also vanishes on $\ker(1-u) = (\mathbb{C}a)^{\perp}$, so we must have some $\lambda \in \mathbb{C}^*$ such that $\phi = \lambda\langle\cdot, a\rangle$. Hence $u(x) = x - \lambda\langle x, a\rangle a$.

To determine $\lambda$, we note that $u(a) = \alpha a$ for some $\alpha \in \mathbb{C}$, $|\alpha| = 1$, so we must have $\lambda = \frac{1-\alpha}{\langle a,a\rangle}$ as required. □

DEFINITION 2.9. *For non-zero $a \in H$ and $\alpha \in \mathbb{C}$, $|\alpha| = 1$, we define the reflection $r_{a,\alpha}(x)$ by*

$$r_{a,\alpha}(x) = x - (1-\alpha)\frac{\langle x, a\rangle}{\langle a, a\rangle}a.$$

Note that $r_{a,\alpha}$ has eigenvalue 1 with multiplicity $n-1$ and eigenvalue $\alpha$ with multiplicity 1. The following facts are easy to check and we omit the proofs.

PROPOSITION 2.10. *For any non-zero $a \in H$ and $\alpha, \beta \in \mathbb{U}$ we have the following.*

(1) $r_{a,\alpha}r_{a,\beta} = r_{a,\alpha\beta}$.
(2) *For every $g \in U(H)$, $gr_{a,\alpha}g^* = r_{ga,\alpha}$.*
(3) *For every non-zero $\lambda \in \mathbb{C}$, $r_{\lambda a,\alpha} = r_{a,\alpha}$.*
(4) $r_{a,\alpha}^{-1} = r_{a,\bar{\alpha}}$.

PROPOSITION 2.11. *Let $a, b \in H$ be non-zero vectors and $\alpha, \beta \in \mathbb{U}$. Then the reflections $r_{a,\alpha}$ and $r_{b,\beta}$ commute if and only if $\mathbb{C}a = \mathbb{C}b$ or $\langle a, b\rangle = 0$.*

PROOF. This follows immediately by writing

$$r_{a,\alpha}r_{b,\beta}(x) = x - (1-\alpha)\frac{\langle x, a\rangle}{\langle a, a\rangle}a - (1-\beta)\frac{\langle x, b\rangle}{\langle b, b\rangle} + (1-\alpha)(1-\beta)\frac{\langle b, a\rangle\langle x, b\rangle}{\langle a, a\rangle\langle b, b\rangle}a$$

which shows that the reflections commute if and only if

$$\langle b, a\rangle\langle x, b\rangle a = \langle a, b\rangle\langle x, a\rangle b. \qquad \square$$

Now we note that given two distinct vectors $e, m \in H$ of unit length, there exists a unique complex reflection $r$ such that $r(e) = m$, which is $r_{m-e,\alpha}$ where $\alpha = -\frac{1-\langle m,e\rangle}{1-\overline{\langle m,e\rangle}}$. Such $r$ is given by the equation

$$r(x) = x - \frac{\langle x, m-e\rangle}{1-\overline{\langle m, e\rangle}}(m-e).$$

**2.4. Virtual isometries and projections of unitary matrices.** We now explain how to construct an infinite dimensional structure which contains in a natural way each finite dimensional unitary group. It is crucial to our construction that the unitary matrices between different dimensions be closely linked. This will allow us to give a meaningful definition to almost sure convergence of random matrices as the dimension grows to infinity.

PROPOSITION 2.12. *Let $H$ be a complex Hilbert space, $E$ a finite dimensional subspace of $H$ and $F$ a subspace of $E$. Then for any unitary operator $u$ acting on $H$ which fixes every vector in $E^{\perp}$, there exists a unique unitary operator $\pi_{E,F}(u)$ on $H$ which satisfies the following two conditions.*

(1) *$\pi_{E,F}(u)$ fixes every vector in $F^{\perp} \supseteq E^{\perp}$.*
(2) *The image of $H$ under $u - \pi_{E,F}(u)$ is contained in the image of $F^{\perp}$ under $u - 1$.*

*Moreover if $G$ is a subspace of $F$, then $\pi_{F,G} \circ \pi_{E,F}(u)$ is a well-defined unitary operator on $H$ and is equal to $\pi_{E,G}(u)$.*

PROOF. First we prove uniqueness. Let $x \in F \cap (u-1)(F^{\perp})$. There exists $y \in F^{\perp}$ such that $x = u(y) - y$. Since $x \in F$ and $y \in F^{\perp}$ we have

$$\|u(y)\|^2 = \|y\|^2 + \|x\|^2$$

by the Pythagorean theorem. Since $u$ is unitary $\|u(y)\| = \|y\|$ so we would require $\|x\| = 0$, hence $F \cap (u-1)(F^{\perp}) = \{0\}$.

Now if $v_1$ and $v_2$ are two unitary operators satisfying the properties of $\pi_{E,F}(u)$, then:

(1) $v_1$ and $v_2$ fix globally $F$ since they fix $F^{\perp}$.
(2) $v_1 - v_2$ vanishes on $F^{\perp}$ by construction.
(3) The range of $v_1 - v_2$ is included in $(u-1)(F^{\perp})$ since the range of each of $v_1 - u$ and $u - v_2$ are.

We conclude from (2) and (3) that the image of $v_1 - v_2$ is contained in $F \cap (u-1)(F^{\perp}) = \{0\}$ and so the operators must agree.

Next we prove the tower property of $\pi_{E,F}(u)$, assuming that the operator exists.

Let $G \subset F \subset E \subset H$. We write $v = \pi_{E,F}(u)$ and $w = \pi_{F,G}(v)$. These operators are well-defined and satisfy

(1) $v$ fixes every vector in $F^{\perp}$
(2) $(u-v)(H) \subseteq (u-1)(F^{\perp})$
(3) $w$ fixes every vector in $G^{\perp}$
(4) $(v-w)(H) \subseteq (v-1)(G^{\perp})$.

This yields the elementary calculation

$$\begin{aligned}(u-w)(H) &\subseteq \operatorname{span}\{(u-v)(H), (v-w)(H)\} \\ &\subseteq \operatorname{span}\{(u-1)(F^{\perp}), (v-1)(G^{\perp})\} \\ &\subseteq \operatorname{span}\{(u-1)(F^{\perp}), (u-1)(G^{\perp}), (u-v)(G^{\perp})\} \\ &\subseteq \operatorname{span}\{(u-1)(G^{\perp}), (u-v)(H)\} \\ &\subseteq (u-1)(G^{\perp}).\end{aligned}$$

Since $w$ fixes each vector of $G^{\perp}$, it is equal to $\pi_{E,G}(u)$.

Now we prove the existence of the operator by induction. It is sufficient to prove the existence of $\pi_{E,F}$ in the particular case where $E = \operatorname{span}\{F, e\}$, where $e$ is a unit vector orthogonal to $F$. In this case, if $u$ is a unitary operator fixing each vector of $E^{\perp}$, then the operator $v = \pi_{E,F}(u)$ can be constructed explicitly as follows.

If $u(e) = e$ then on taking $v = u$ which fixes $\operatorname{span}\{E^{\perp}, e\}$, hence it fixes $F^{\perp}$ and $(u-v)(H) = (u-1)(F^{\perp})$.

If $u(e) \neq e$ then for all $x \in H$ we define

$$v(x) := u(x) - \frac{\langle u(x), e - u(e)\rangle}{1 - \langle u(e), e\rangle}(e - u(e)) = \widetilde{r} \circ u$$

where $\widetilde{r}$, as indicated, is the unique reflection mapping $u(e) \mapsto e$. Hence $v$ is a unitary transformation.

Now let $x \in E^{\perp}$ so that $u(x) = x$ and $e - u(e) \in E$ since $E$ is globally fixed by $u$. Here $\langle u(x), e - u(e)\rangle = \langle x, e - u(e)\rangle = 0$ so $v(x) = x$. Moreover by construction $v(e) = e$, so consequently $v$ fixes each vector in $F^{\perp}$.

Finally, for all $x \in H$, we have $u(x) - v(x) = \gamma(e - u(e))$ for some $\gamma \in \mathbb{C}$, so $(u-v)(H) \subseteq (u-1)(F^{\perp})$, hence $v$ satisfies the conditions of $\pi_{E,F}(u)$. □

REMARK 2.13. *It follows from Proposition 2.10 that if $r := (\widetilde{r})^{-1}$ is the reflection mapping $e$ onto $u(e)$ then*

$$u = r\pi_{E,F}(u).$$

*Similarly one can easily prove that $u = \pi_{E,F} \circ r'$ where $r'$ is the unique reflection such that $r'(u^{-1}(e)) = e$.*

The existence of the projection map described above suggests how to define "virtual isometries", the infinite dimensional objects alluded to above. Indeed, let $H = \ell^2(\mathbb{C})$ and $(e_\ell)_{\ell \geq 1}$ be the canonical

Hilbert basis of $H$. Then for all $n \geq 1$ the space of unitary operators fixing each element of $V_n := \operatorname{span}(e_1, \dots, e_n)^\perp$ can be canonically identified with the unitary group $U(n)$.

By identification, for $n \geq m \geq 1$ the projection $\pi_{V_n,V_m}$ gives a map from $U(n)$ to $U(m)$, more simply noted $\pi_{n,m}$, so that, for $n \geq m \geq p \geq 1$,

$$\pi_{n,p} = \pi_{m,p} \circ \pi_{n,m}.$$

DEFINITION 2.14. *A **virtual isometry** is a sequence $(u_n)_{n\geq 1}$ of unitary matrices such that for all $n \geq 1$, $u_n \in U(n)$ and $\pi_{n+1,n}(u_{n+1}) = u_n$. The space of virtual isometries will be denoted $U^\infty$.*

REMARK 2.15. *$U^\infty$ does not appear to have a natural group structure.*

REMARK 2.16. *In this definition we made a particular choice of vector spaces lying in $\ell^2(\mathbb{C})$. Although it appears to be necessary to make a choice for the probabilistic arguments later, the ideas in this construction apply also to other Hilbert spaces with a sequence of basis vectors.*

PROPOSITION 2.17. *Let $(x_n)_{n\geq 1}$ be a sequence of vectors with $x_n \in \mathbb{C}^n$ lying on the unit sphere (i.e. $\|x_n\|_{\ell^2(\mathbb{C}^n)} = 1$). Then there exists a unique virtual isometry $(u_n)_{n\geq 1}$ such that $u_n(e_n) = x_n$ for every $n \geq 1$. In particular, $u_n$ is given by*

$$u_n = r_n r_{n-1} \cdots r_1$$

*where for each $1 \leq j \leq n$, $r_j = 1$ if $x_j = e_j$ and otherwise is the unique reflection mapping $e_j \mapsto x_j$.*

PROOF. This follows directly from the proof of Proposition 2.12 and the remarks which follow it. □

REMARK 2.18. *In the particular case where every $x_n = e_{t(n)}$ for some $t(n) \in \{1, \dots, n\}$, then $(u_n)_{n\geq 1}$ is the sequence of matrices associated to a virtual permutation* [**23**]. *The sequence of permutations $(\sigma_n)_{n\geq 1}$ is constructed by the so-called Chinese restaurant process* [**25**]*: for all $n \geq 1$,*

$$\sigma_n = \tau_{n,t(n)} \tau_{n-1,t(n-1)} \cdots \tau_{1,1}$$

*where $\tau_{k,j} = 1$ if $j = k$ and otherwise is the transposition $(j, k)$.*

REMARK 2.19. *The above proposition shows in particular the fact that any unitary matrix of $U(n)$ can be expanded into a product of reflections.*

Proposition 2.17 shows in particular that $U^\infty$ is non-empty. Moreover, our construction will allow us to construct measures on $U^\infty$. Indeed, it is natural to look for the analogue of Haar measure on $U^\infty$. First, we recall the following correspondence from [**4**] (this can also be obtained as a consequence of the results above and the subgroup algorithm of Diaconis and Shahshahani [**9**]).

PROPOSITION 2.20 ([**4**]). *Let $(x_n)_{n\geq 1}$ be a random sequence of vectors with $x_n \in \mathbb{C}^n$ and let $(u_n)_{n\geq 1}$ be the unique virtual isometry such that $u_n(e_n) = x_n$ for all $n \geq 1$. Then for each $n \geq 1$ the matrix $u_n$ is distributed according to Haar measure if and only if $x_1, \dots, x_n$ are independent and for each $1 \leq j \leq n$, $x_j$ is distributed according to the uniform measure on the unit sphere in $\mathbb{C}^j$.*

As a consequence, we deduce the compatibility between Haar measure on $U(n)$, $n \geq 1$, and the projections $\pi_{n,m}$ for $n \geq m \geq 1$.

PROPOSITION 2.21. *For all $n \geq m \geq 1$, the push-forward of the Haar measure on $U(n)$ under $\pi_{n,m}$ is the Haar measure on $U(m)$.*

REMARK 2.22. *One has to check that $\pi_{n,m}$ is measurable as is done in* [**4**]

The compatibility property of Proposition 2.21 allows us to define the Haar measure on $U^\infty$.

PROPOSITION 2.23. *Let $\mathcal{U}$ denote the $\sigma$-algebra on $U^\infty$ generated by the sets*

$$W(k, B) := \{(u_n)_{n\geq 1} \mid u_k \in B\}$$

*where $k \geq 1$ and $B \subseteq U(k)$ is a Borel set. Let $(\mu_n)_{n\geq 1}$ be a family of probability measures, $\mu_n$ defined on $(U(n), \mathcal{B}(U(n)))$, where $\mathcal{B}(U(n))$ is the Borel set on $U(n)$, and such that the push-forward of $\mu_{n+1}$ by $\pi_{n+1,n}$ is equal to $\mu_n$. Then there exists a unique probability measure $\mu$ on $(U^\infty, \mathcal{U})$ such that push-forward of $\mu$ by the nth coordinate map is equal to $\mu_n$ for all $n \geq 1$.*

PROOF. This is the Kolmogorov extension theorem applied to the family $(\mu_n)_{n\geq 1}$. See [**4**]. □

Now combining Proposition 2.20 and Proposition 2.23 we obtain the following.

PROPOSITION 2.24 ([4]). *There exists a unique probability measure $\mu^{(\infty)}$ on the space $(U^\infty, \mathcal{U})$ such that its push-forward by the coordinate maps is equal to the Haar measure on the corresponding unitary group. In particular, let $(x_n)_{n\geq 1}$ be a random sequence of vectors with each $x_n \in \mathbb{C}^n$ almost surely on the unit sphere, and let $(u_n)_{n\geq 1}$ be the unique virtual isometry such that $u_n(e_n) = x_n$. Then the distribution of $(u_n)_{n\geq 1}$ is equal to $\mu^{(\infty)}$ if and only if the $x_n$ are independent random variables and for all $n \geq 1$, $x_n$ is uniformly distributed on the unit sphere.*

### 3. Revisiting and refining the Ketaing-Snaith analysis of the characteristic polynomial

Now we give a simple recursive result for the characteristic polynomial from which many interesting results will follow.

THEOREM 3.1. *Let $(u_n)_{n\geq 1}$ be a virtual isometry and for $n \geq 1$ let $x_n = u_n(e_n)$. Note $v_n = x_n - e_n$. Let $(f_k^{(n)})_{1\leqslant k\leqslant n}$ be an orthonormal basis of $\mathbb{C}^n$ consisting of the eigenvectors of $u_n$, with $(\lambda_k^{(n)})_{1\leqslant k\leqslant n}$ the corresponding eigenvalues. Let $P_n(z)$ be the characteristic polynomail of $u_n$, given by*

$$Z_n(z) = \det(z1 - u_n),$$

*and let us decompose $x_{n+1} \in \mathbb{C}^{n+1}$ as*

$$x_{n+1} = \sum_{k=1}^{n} \mu_k^{(n)} f_k^{(n)} + \gamma_n e_{n+1}.$$

*Then for all $n \geq 1$ such that $x_{n+1} \neq e_{n+1}$, one has $\gamma_n \neq 1$, and we have the recursive relation:*

$$Z_{n+1}(z) = \frac{Z_n(z)}{1-\bar{\gamma}_n}\left[(z-\gamma_n)(1-\bar{\gamma}_n) + (z-1)\sum_{k=1}^{n}\left|\mu_k^{(n)}\right|^2 \frac{\lambda_k^{(n)}}{z-\lambda_k^n}\right]$$

*for all $z \notin \{\lambda_1^{(n)}, \cdots, \lambda_n^{(n)}\}$, $\gamma_n = \langle x_{n+1}, e_{n+1}\rangle$.*

PROOF. Since $(u_n)_{n\geq 1}$ is a virtual isometry and $x_{n+1} \neq e_{n+1}$ we have that

$$u_{n+1} = r_{n+1}(u_n \oplus 1),$$

where $r_{n+1}$ is the unique reflection such that $r_{n+1}(e_{n+1}) = x_{n+1}$ and where the notation $\oplus$ stands for diagonal blocks of matrices. The matrix on $r_{n+1}$ is given by (see appendix)

$$r_{n+1} = \mathbf{1}_{n+1} - \frac{1}{1-\bar{\gamma}_n} v_{n-1} v_{n+1}^T.$$

Hence for $z \notin \{\lambda_1^{(n)}, \cdots, \lambda_n^{(n)}\}$

$$\begin{aligned} Z_{n+1}(z) &= \det(z\mathbf{1}_{n+1} - u_n \oplus 1)\det\left(\mathbf{1}_{n+1} + \frac{1}{1-\bar{\gamma}_n}(z\mathbf{1}_{n+1} - u_n \oplus 1)^{-1} v_{n+1}\bar{v}_{n+1}^T (u_n \oplus 1)\right) \\ &= (z-1)Z_n(z) \end{aligned}$$

From a well known fact about rank one matrices we have (see Appendix 6 ) $\det(\mathbf{1} + A) = 1 + \mathrm{Tr}(A)$ and so

$$= (z-1)Z_n(z)\left(1 + \frac{1}{1-\bar{\gamma}_n}\,\mathrm{tr}\{(z\mathbf{1}_{n+1} - u_n \oplus 1)^{-1}\} v_{n+1}\bar{v}_{n+1}^T (u_n \oplus 1)\}\right).$$

Now we write all matrices in the basis $(e_{n+1}, f_1^{(n)}, \cdots, f_n^{(n)})$. The matrix of $(u_n \oplus 1)$ is

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \lambda_1^{(n)} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^{(n)} \end{pmatrix}.$$

The matrix of $(z\mathbf{1}_{n+1} - u_n \oplus 1)^{-1}$

$$\begin{pmatrix} 1/(z-1) & 0 & \cdots & 0 \\ 0 & 1/(z-\lambda_1^{(n)}) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1/(z-\lambda_n^{(n)}) \end{pmatrix}.$$

Since

$$v_{n+1} = x_{n+1} - e_{n+1} = \sum_{k=1}^{n} \mu_k^{(n)} f_k^{(n)} + (\gamma_n - 1)e_{n+1}$$

we have

$$v_{n+1}\bar{v}_{n+1}^T = \begin{pmatrix} |\gamma_n - 1|^2 & & \\ & \left|\mu_1^{(n)}\right|^2 & \\ & & \left|\mu_n^{(n)}\right|^2 \end{pmatrix}.$$

Note that if $A = (a_{ij})_{1\leqslant i,j \leqslant n}$ is a matrix, then $A\,\mathrm{diag}(\alpha_1, \cdots, \alpha_n) = (\alpha_1 c_1, \cdots, \alpha_n c_n)$ where $c_j$'s are the columns of $A$ and $\mathrm{diag}(\alpha_1, \cdots, \alpha_n)A = \begin{pmatrix} \alpha_1 l_1 \\ \vdots \\ \alpha_n l_n \end{pmatrix}$ where $l_i$'s are the rows of $A$. Consequently

$$Z_{n+1}(z) = \frac{Z_n(z)}{1-z}\left[1 + \frac{1}{1-\bar{\gamma}_n}\left(\frac{|1-\gamma_n|^2}{z-1} + \sum_{k=1}^{n}\left|\mu_k^{(n)}\right| \frac{\lambda_k^{(n)}}{z-\lambda_k^{(n)}}\right)\right]$$

the result follows by a simple expansion. □

COROLLARY 3.2. *The following hold:*

(1) *Let $(x_n)_{n\geq 1}$ be independent random variables such that $x_n$ is uniformly distributed on $\mathscr{S}_{\mathbb{C}}^n$. Let $(u_n)_{n\geq 1}$ be the unique virtual isometry such that $x_n = u_n(e_n)$. Let $Z_n = \det(\mathbf{1} - u_n)$. Then*

$$Z_n = (1-\gamma_{n-1})(1-\gamma_{n-2})\cdots(1-\gamma_0)$$

*where $\gamma_k = \langle x_{k+1}, e_{k+1}\rangle$.*

(2) *Let $U \in U(n)$ be Haar distributed, then*

$$\begin{aligned}\det(1_n - U) &\overset{\text{law}}{=} \prod_{k=1}^{n}(1 + e^{i\theta_k}\sqrt{\mathrm{B}_{1,k-1}}) \\ &\overset{\text{law}}{=} \prod_{k=1}^{n}(1 - e^{i\theta_k}\sqrt{\mathrm{B}_{1,k-1}})\end{aligned}$$

*where $\theta_1, \cdots, \theta_n$ are uniformly distributed on $[0, 2\pi]$, $\mathrm{B}_{1,k-1}$ is Beta distributed with parameters $1$, $k-1$ with the convention that $\mathrm{B}_{1,0} = 1$ a.s., and all variables in sight are independent.*

PROOF. In the previous proposition if we take $z = 1$, we have

$$Z_{n+1}(1) = \frac{Z_n(1)}{1-\bar{\gamma}_n}(1-\gamma_n)(1-\bar{\gamma}_n) = (1-\gamma_n)Z_n(1)$$

which implies that

$$Z_{n+1} = Z_n(1-\gamma_n).$$

The the first result follows with a simple induction.

The distribution decomposition follows from the well known fact (see Appendix 2 for background on uniform distributions) that if $x_k$ is uniformly distributed on $\mathbb{S}_{\mathbb{C}}^k$, then for all $j$, $\langle x_k, e_j\rangle$ is distributed like

$$e^{i\theta_k}\sqrt{\mathrm{B}_{1,k-1}}.$$

Then

$$Z_n \overset{\text{law}}{=} \prod_{k=1}^{n}(1 - e^{i\theta_k}\sqrt{\mathrm{B}_{1,k-1}}) \overset{\text{law}}{=} \prod_{k=1}^{n}(1 + e^{i\theta_k}\sqrt{\mathrm{B}_{1,k-1}})$$

which follows from the fact that $-e^{i\theta_k} \overset{\text{law}}{=} e^{i\theta_k}$.

□

Now we see that the joint moment generating function for the characteristic polynomial follows easily from our decomposition into a product of independent random variables.

LEMMA 3.3. *Let $X = 1 + e^{i\theta}\sqrt{\beta}$ where $\theta$ is uniformly distributed on $[0, 2\pi]$, independent of $\beta$ which follows a Beta distribution with parameters 1 and $k-1$ for $k \geq 1$. Then for all $t$ and $s$ such that $\mathrm{Re}(t \pm s) > -2$ and $\mathrm{Re}(t) > -1$ we have*

$$\mathbb{E}[|X|^t e^{is\arg(X)}] = \frac{\Gamma(k)\Gamma(k+t)}{\Gamma(k+\frac{t+s}{2})\Gamma(k+\frac{t-s}{2})}.$$

PROOF. We use notation from Appendix 4 on hypergeometric functions. We begin by noting that

$$X^{(t+s)/2}X^{-(t-s)/2} = |X|^t e^{is\arg(X)}$$

and thus

$$\begin{aligned}\mathbb{E}[|X|^t e^{is\arg(X)}] &= \mathbb{E}[X^{(t+s)/2}X^{-(t-s)/2}] \\ &= \mathbb{E}[(1+e^{i\theta}\sqrt{\beta})^a(1+e^{-i\theta}\sqrt{\beta})^b], \quad a = \tfrac{t+s}{2}, \quad b = \tfrac{t-s}{2}.\end{aligned}$$

Note that a.s. $\sqrt{\beta} < 1$

$$(1+x)^u = \sum_{k=0}^{\infty} \frac{(-1)^k(-u)_k x^k}{k!} \qquad |x| < 1,$$

with

$$(-u)_k = (-u)(-u+1)\cdots(-u+k-1).$$

As $|e^{i\theta}\sqrt{\beta}| < 1$ a.s., we have

$$\mathbb{E}[|X|^t e^{is\arg(X)}] = \mathbb{E}\left[\sum_{k=0}^{\infty}\frac{(-1)^k(-a)_k}{k!}\beta^{k/2}e^{ik\theta}\sum_{l=0}^{\infty}\frac{(-1)^l(-b)_l}{l!}\beta^{l/2}e^{-il\theta}\right].$$

Since $\mathbb{E}[e^{im\theta}] = 0$ if $m \neq 0$

$$= \sum_{l=0}^{\infty}\frac{(-1)^{2l}(-a)_l(-b)_l}{l!l!}E[\beta^l].$$

But

$$\mathbb{E}[\beta^l] = \frac{\Gamma(1+l)\Gamma(k)}{\Gamma(1)\Gamma(l+k)} = \frac{l!}{(k)_l}$$

so that in fact we have

$$\begin{aligned}&= \sum_{l=0}^{\infty}\frac{(-a)_l(-b)_l}{l!}\frac{1}{(k)_l} \\ &= {}_2F_1(-a,-b,k;1) \\ &= \frac{\Gamma(k)\Gamma(k+t)}{\Gamma(k+\frac{t+s}{2})\Gamma(k+\frac{t-s}{2})}\end{aligned}$$

by Gauss' formula. □

THEOREM 3.4. *We have for $\mathrm{Re}(t) > -1$:*

$$E[|Z_n|^t e^{is\arg(X)}] = \prod_{k=1}^{n}\frac{\Gamma(k)\Gamma(k+t)}{\Gamma(k+\frac{t+s}{2})\Gamma(k+\frac{t-s}{2})}.$$

PROOF. Follows from independence and the previous lemma. □

We now want to show that

$$\frac{\log Z_n}{\sqrt{\frac{1}{2}\log n}} \overset{\text{law}}{\to} \mathcal{N}_1 + i\mathcal{N}_2$$

as $n \to \infty$. One way of doing this, is to consider

$$\log\left(\prod_{k=1}^{n}(1+e^{i\theta_k}\sqrt{\mathrm{B}_{1,k-1}})\right) = \sum_{k=1}^{n}\log(1+e^{i\theta_k}\sqrt{\mathrm{B}_{1,k-1}})$$

and we have a sum of independent random variables. We would rather use another decomposition result.

THEOREM 3.5. *Let $(\beta_{j,j-1})$ be independent Beta variables with parameters $j$ and $j-1$ (with the construction that $\beta_{1,0} = 1$ a.s.). Define $W_1, \cdots, W_n$ as independent random variables that are independent of $(\beta_{j,j-1})_{1\le j\le n}$ with $W_j$ having density*

$$\sigma_{2(j-1)}(dv) = \frac{2^{2(j-1)}((j-1)!)^2}{\pi(2j-2)!}\cos^{2(j-1)}(v)\mathbf{1}_{(-\pi/2,\pi/2)}(v)dv.$$

*The joint distribution of $(\mathrm{Im}(\log(Z_n)), |Z_n|)$ is*

$$(\mathrm{Im}(\log Z_n), |Z_n|) \overset{\text{law}}{=} \left(\sum_{j=1}^n W_j, \prod_{j=1}^n \beta_{j,j-1} 2\cos(W_j)\right).$$

LEMMA 3.6. *Let $W_j$ have density*

$$K_j \cos^{2(j-1)}(v)1_{(-\pi/2,\pi/2)}(v)dv$$

*where $J_j$ is a normalizing constant. Next, let $X_j := \beta_{j,j-1}2\cos(W_j)e^{iW_j}$. Then for $\mathrm{Re}(t) > -1$, one has*

$$E[|X_j|^t e^{is\arg(X_j)}] = \frac{\Gamma(j)\Gamma(j-t)}{\Gamma(j+\frac{t+s}{2})\Gamma(j+\frac{t-s}{2})}.$$

Recall the classical results on Wallis' integrals:

$$\int_{-\pi/2}^{\pi/2} \cos^{2(j-1)}(v)dv = 2\int_0^{\pi/2}\cos^{2(j-1)}(v)dv = 2\int_0^{\pi/2}\sin^{2(j-1)}(v)dv = 2I_{2(j-1)}$$

with

$$I_n = \int_0^{\pi/2}\sin^n(v)dv.$$

Integration by parts yields

$$I_{n+2} = \frac{n+1}{n+2}I_n.$$

Thus $\forall p \in \mathbb{N}^*$

$$I_{2p} = \frac{(2p-1)(2p-3)\cdots 1}{(2p)(2p-2)\cdots 2}\frac{\pi}{2} = \frac{(2p)!}{2^{2p}(p!)^2}\frac{\pi}{2}.$$

PROOF OF THE LEMMA. One has

$$\begin{aligned}\mathbb{E}[|X_j|^t e^{is\arg(X_j)}] &= \mathbb{E}[|\beta_{j,j-1}2\cos W_j|^t e^{isW_j}] \\ &= \mathbb{E}[\left|\beta_{j,j-1_j}\right|^t]E[|2\cos W_j|^t e^{isW_j}] \\ &= \mathbb{E}[\left|\beta_{j,j-1_j}\right|^t]\frac{\Gamma^2(j)}{\pi\Gamma(j-1)!}\int_{-\pi/2}^{\pi/2} e^{isx}(e^{ix}+e^{-ix})^t(e^{ix}+e^{-ix})^{2(j-1)}dx\end{aligned}$$

We already saw that

$$\mathbb{E}[\left|\beta_{j,j-1_j}\right|^t] = \frac{\Gamma(j+t)\Gamma(2j-1)}{\Gamma(j)\Gamma(2j-1+t)}.$$

Moreover, we note that

$$e^{isx}(e^{ix}+e^{-ix})^t(e^{ix}+e^{-ix})^{2(j-1)} = (1+e^{2ix})^{j-1+(t+s)/2}(1+e^{-2ix})^{j-1+(t-s)/2}.$$

We now expand the right hand side as a hypergeometric function for $x \neq 0$

$$\begin{aligned}(1+e^{2ix})^{j-1+(t+s)/2} &= {}_1F_0(-(j-1+\tfrac{t+s}{2}); -e^{2ix}), \\ (1+e^{-2ix})^{j-1+(t-s)/2} &= {}_1F_0(-(j-1+\tfrac{t-s}{2}); -e^{-2ix}).\end{aligned}$$

We now integrate between $(-\pi/2, \pi/2)$

$$\begin{aligned}&\int_{-\pi/2}^{\pi/2} e^{isx}(e^{ix}+e^{-ix})^{2(j-1)}(e^{ix}+e^{-ix})^t dx \\ &= \int_{-\pi/2}^{\pi/2} {}_1F_0(-(j-1+\tfrac{t+s}{2}); -e^{2ix})\,{}_1F_0(-(j-1+\tfrac{t-s}{2}); -e^{-2ix})dx\end{aligned}$$

Since

$$\int_{-\pi/2}^{\pi/2} e^{2ikx} dx = 0$$

if $k \neq 0$ and $\pi$ otherwise. Hence only diagonal terms in the double sum will survive to give

$$\pi {}_2F_1 \left( \begin{matrix} -(j-1+\frac{t+s}{2}), (j-1+\frac{t-s}{2}) \\ 1 \end{matrix} ; 1 \right),$$

which by Gauss's formula for $\mathrm{Re}(t) > -1$

$$\pi \frac{\Gamma(2j-1+t)}{\Gamma(j+\frac{t+s}{2})\Gamma(j+\frac{t-s}{2})}$$

this ends the proof. □

THEOREM 3.7. *We note again $Z_n = \det(\mathbf{1}_n - U)$, where $U \in U(n)$ and is Haar distributed. Then*

$$\frac{\log Z_n}{\sqrt{\frac{1}{2}\log n}} \underset{n\to\infty}{\overset{\text{law}}{\to}} \mathcal{N}_1 + i\mathcal{N}_2,$$

*where $\mathcal{N}_1$ and $\mathcal{N}_2$ arw two standard independent Gaussian random variables.*

PROOF. Let us first recall a few facts about cumulants. The equation

$$\mathbb{E}[e^{itX}] = \sum_{n=0}^{\infty} \mathbb{E}[X^n] i^n t^n$$

is not true in all generality.
Assume that $X$ is a random variable such that $\mathbb{E}[e^{tX}]$ exists for all $t < \varepsilon$. Then we define the cumulant of order $n$ and note it $\chi_N$ by

$$g(t) := \log \mathbb{E}[e^{tX}] = \sum_{n=0}^{\infty} \chi_n \frac{t^n}{n!}.$$

If instead we consider the characteristic function

$$h(t) := \log \mathbb{E}[e^{-itX}] = \sum_{n=0}^{\infty} \chi_n \frac{(it)^n}{n!}.$$

More generally, if $X$ has a moment or order $h$, then

$$\log(\mathbb{E}[e^{itx}]) = \sum_{n=0}^{h} \chi_n \frac{(it)^n}{n!} + o(t^h).$$

Note that $\chi_1 = \mathbb{E}[X]$. If $\mathbb{E}[X] = 0$, then $\chi_1 = 0$ and we have $\chi_2 = \mathrm{var}(X) = \mathbb{E}[X^2]$. Note that for the Gaussian distribution all cumulants $\chi_n$ for $n \geq 3$ are zero. Finally, note that, if $X$ anbd $Y$ are independent, then

$$\chi_n(X+Y) = \chi_n(X) + \chi_n(Y).$$

We also introduce the polygamma functions

$$\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}, \quad \psi^{(k)}(z) = \frac{d^{k+1}}{dz^{k+1}} \log \Gamma(z).$$

It is known that as $z \to \infty$, $|\arg z| < \pi$ we have

$$\psi(z) \sim \log z - \frac{1}{2z} - \sum_{n=1}^{\infty} \frac{B_{2n}}{z^{2n}}$$

$$\psi^{(k)}(z) \sim (-1)^{k-1} \left[ \frac{(k-1)!}{z^k} + \frac{k!}{2z^{k+1}} + \sum_{n=0}^{\infty} B_{2n} \frac{(2n+k-1)!}{(2n)! z^{2n+k}} \right]$$

where $B_{2n}$ are the Bernoulli numbers given by

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}.$$

We denote $T_j := \log(\beta_{j,j-1} 2\cos W_j)$

$$(\log|Z_n|, \arg Z_n) \overset{\text{law}}{=} \left(\sum_{j=1}^n T_j, \sum_{j=1}^n W_j\right).$$

Furthermore

$$\mathbb{E}[e^{isW_j}] = \frac{\Gamma^2(j)}{\Gamma(j+\frac{s}{2})\Gamma(j-\frac{s}{2})}.$$

For $\mathrm{Re}(t) > -1$

$$\mathbb{E}[e^{tT_j}] = \frac{\Gamma(j)\Gamma(j+t)}{\Gamma^2(j+\frac{t}{2})}.$$

Let $\psi_{j,k}$ denote the $k$th cumulant of $T_j$ and let $R_{j,k}$ denote the $k$th cumulant of $W_j$. Then

$$g(t) := \log(E[e^{tT_j}]) = \log(\Gamma(j)) + \log(\Gamma(j+t)) - 2\log\Gamma(j+\tfrac{t}{2})$$

which implies that

$$g^{(k)}(t) = \psi^{(k-1)}(j+t) - \frac{1}{2^{k-1}}\psi^{(k-1)}(j+\tfrac{t}{2})$$

and

$$\psi_{j,k} = g^{(k)}(0) = \frac{2^{k-1}-1}{2^{k-1}}\psi_j^{(k-1)}.$$

Similarly,

$$R_{j,k} = \frac{(-1)^{k/2+1}}{2^{k-1}}\psi_j^{(k-1)}$$

when $k$ is even and 0 when $k$ is odd. Since $T_j$ and $W_j$ are independent, the $k$th cumulant of $\sum_{j=1}^n T_j$ is

$$\frac{2^{k-1}-1}{2^{k-1}}\sum_{j=1}^n \psi_{(j)}^{(k-1)}.$$

The $k$th cumulant of $\sum_{j=1}^n W_j$ is

$$\frac{(-1)^{k/2+1}}{2^{k-1}}\sum_{j=1}^n \psi_{(j)}^{(k-1)}\mathbf{1}_{k \text{ even}}.$$

The first cumulant is zero so the sums are centered. The second cumulant gives the variance and this shows that

$$\operatorname{var}\left(\sum_{j=1}^n T_j\right) \sim \frac{1}{2}\sum_{j=1}^n \frac{1}{j} \sim \frac{1}{2}\log n,$$
$$\operatorname{var}\left(\sum_{j=1}^n W_j\right) \sim \frac{1}{2}\sum_{j=1}^n \frac{1}{j} \sim \frac{1}{2}\log n.$$

Let

$$L_N = \frac{1}{S_N^3}\sum_{n=1}^N \mathbb{E}[|T_n|^3], \quad S_N^2 = \sum_{j=1}^N \mathbb{E}[|T_j|^2] \sim \frac{1}{2}\log n$$
$$L_N{}' = \frac{1}{\sigma_N^3}\sum_{n=1}^N \mathbb{E}[|W_n|^3], \quad \sigma_N^3 = \sum_{j=1}^N \mathbb{E}[|W_j|^2] \sim \frac{1}{2}\log n.$$

Since

$$\sum_{n=1}^N \mathbb{E}[|T_n|^3] < \infty, \quad \sum_{n=1}^N \mathbb{E}[|W_n|^3] < \infty$$

we have that $L_N \to 0$ and $L_N' \to 0$ as well. Next, how to check that the following holds?

$$\sum_{n=1}^\infty \mathbb{E}[|T_n|^3] < \infty, \quad \sum_{n=1}^N \mathbb{E}[|W_n|^3] < \infty$$

For instance, we know that $\mathbb{E}[|W_j|^3] \leq \mathbb{E}[|W_j|^4]^{3/4}$. One checks that

$$\mathbb{E}[|W_j|^4]^{3/4} = (R_{j,4} + 3R_{2,j}^2)^{3/4} \sim c/j^{3/2}$$

which implies that

$$\sum \mathbb{E}[|W_j|^3] < \infty,$$

and the same thing for $T_j$. □

From now on we write

$$\Phi(x) = \int_{-\infty}^{x} e^{-t^2/2} \frac{dt}{\sqrt{2\pi}},$$

and use classical results around central limit theorems and the law of the iterated logarithm.

COROLLARY 3.8. *The following estimates on the rate of convergence of the CLT hold:*

(1) *for the real part,*

$$\left| \mathbb{P}\left( \frac{\mathrm{Re}(\log Z_n)}{\sqrt{\frac{1}{2}\log n}} \leqslant x \right) - \Phi(x) \right| \leqslant \frac{C}{(\log n)^{3/2}(1+|x|)^3}$$

(2) *for the imaginary part,*

$$\left| \mathbb{P}\left( \frac{\mathrm{Im}(\log Z_n)}{\sqrt{\frac{1}{2}\log n}} \leqslant x \right) - \Phi(x) \right| \leqslant \frac{C}{(\log n)^{3/2}(1+|x|)^3}$$

*where $C > 0$ is a constant.*

COROLLARY 3.9. *Let $(U_n)_{n\geq 1}$ be a virtual isometry which is Haar distributed (i.e. with distribution $\mu_\infty$ on $U^\infty$). Define $Z_n := Z_n(1)$. Then $\mu_\infty$-a.s.*

(1) *for the real part,*

$$\limsup_{n\to\infty} \frac{\mathrm{Re}(\log Z_n)}{\sqrt{\log n \log\log\log n}} = 1,$$

(2) *for the imaginary part,*

$$\liminf_{n\to\infty} \frac{\mathrm{Im}(\log Z_n)}{\sqrt{\log n \log\log\log n}} = -1.$$

## 4. Almost sure convergence of eigenvalues

We consider a sequence $(u_n)_{n\geq 1}$ of virtual isometries. We assume throughout that for each $n \geq 1$, the $n$ eigenvalues of $u_n$ are distinct; this holds almost surely for virtual isometries constructed according to the Haar measure.

We recall that the eigenvalues of $u_n$, $\lambda_1^{(n)}, \lambda_2^{(n)}, \dots, \lambda_n^{(n)}$, are ordered in such a way that $\lambda_k^{(n)} = e^{i\theta_k^{(n)}}$, and

$$0 < \theta_1^{(n)} < \cdots < \theta_n^{(n)} < 2\pi.$$

Moreover, the eigenangles enjoy a property of periodicity: for all $k \in \mathbb{Z}$, $\theta_{k+n}^{(n)} = \theta_k^{(n)} + 2\pi$.

As all the eigenvalues are distinct, each eigenvalue corresponds to a one-dimensional eigenspace. We can therefore write $f_1^{(n)}$, ..., $f_n^{(n)}$ for the family of unit length eigenvectors of $u_n$, which are well-defined up to a complex phase: the notation $f_k^{(n)}$ is then extended $n$-periodically to all $k \in \mathbb{Z}$.

Let $x_n = u_n(e_n)$ and let $r_n$ denote the unique reflection on $\mathbb{C}^n$ mapping $e_n$ to $x_n$. Therefore, we have $u_{n+1} = r_{n+1} \circ (u_n \oplus 1)$. It is natural to decompose $x_{n+1}$ into the basis given by $\iota(f_1^{(n)}), ..., \iota(f_n^{(n)}), e_{n+1}$, where $\iota : \mathbb{C}^n \to \mathbb{C}^{n+1}$ is the inclusion which maps $(x_1, \dots, x_n)$ to $(x_1, \dots, x_n, 0)$. Identifying $f_k^{(n)}$ and $\iota(f_k^{(n)})$, we then have

$$x_{n+1} = \sum_{k=1}^{n} \mu_k^{(n)} f_k^{(n)} + \nu_n e_{n+1}$$

for some $\mu_k^{(n)}$ ($1 \leq k \leq n$) and $\nu_n$ such that $|\mu_1^{(n)}|^2 + \cdots + |\mu_n^{(n)}|^2 + |\nu_n|^2 = 1$. Again, is can be convenient to consider $\mu_k^{(n)}$ for all $k \in \mathbb{Z}$, by a $n$-periodic extension of the sequence. The following result gives the spectral decomposition of $u_{n+1}$ in function of the decomposition of $u_n$ and $x_{n+1}$:

THEOREM 4.1 (Spectral decomposition). *On the event that the coefficients $\mu_1^{(n)}, \ldots, \mu_n^{(n)}$ are all different from zero and that the $n$ eigenvalues of $u_n$ are all distinct (which holds almost surely under the uniform measure on $U^\infty$), the eigenvalues of $u_{n+1}$ are the unique roots of the rational equation*

$$\sum_{j=1}^n |\mu_j^{(n)}|^2 \frac{\lambda_j^{(n)}}{\lambda_j^{(n)} - z} + \frac{|1-\nu_n|^2}{1-z} = 1 - \overline{\nu}_n$$

*on the unit circle. Furthermore, they interlace between* 1 *and the eigenvalues of $u_n$*

$$0 < \theta_1^{(n+1)} < \theta_1^{(n)} < \theta_2^{(n+1)} < \cdots < \theta_n^{(n)} < \theta_{n+1}^{(n+1)} < 2\pi.$$

*and if the phases of the eigenvectors are suitably chosen, they satisfy the relation*

$$(h_k^{(n+1)})^{\frac{1}{2}} f_k^{(n+1)} = \sum_{j=1}^n \frac{\mu_j^{(n)}}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} f_j^{(n)} + \frac{\nu_n - 1}{1 - \lambda_k^{(n+1)}} e_{n+1}$$

*where*

$$h_k^{(n+1)} = \sum_{j=1}^n \frac{|\mu_j^{(n)}|^2}{|\lambda_j^{(n)} - \lambda_k^{(n+1)}|^2} + \frac{|\nu_n - 1|^2}{|1 - \lambda_k^{(n+1)}|^2}$$

*is the unique strictly positive real number which makes $f_k^{(n+1)}$ a unit vector.*

PROOF. Let $f$ be an eigenvector of $u_{n+1}$ with corresponding eigenvalue $z$. Then we have

$$f = \sum_{j=1}^n a_j f_j^{(n)} + b e_{n+1}$$

where $a_1, ..., a_n, b$ are (as yet unknown) complex numbers, not all zero. Our goal is to write these coefficients in terms of $x_{n+1}$ and the eigenvalues of $u_n$.

We have

$$\begin{aligned} zf &= u_{n+1} f \\ &= u_{n+1} \left( \sum_{j=1}^n a_j f_j^{(n)} + b e_{n+1} \right) \\ &= \sum_{j=1}^n a_j u_{n+1} f_j^{(n)} + b u_{n+1} e_{n+1} \\ &= \sum_{j=1}^n a_j \lambda_j^{(n)} r_{n+1} f_j^{(n)} + b x_{n+1}. \end{aligned}$$

We recall that for all $t \in \mathbb{C}^{n+1}$, $r_{n+1}(t)$ is given by

$$r_{n+1}(t) = t + \frac{\langle t, x_{n+1} - e_{n+1} \rangle}{\langle e_{n+1}, x_{n+1} - e_{n+1} \rangle} (x_{n+1} - e_{n+1})$$

so that

$$zf = \sum_{j=1}^n a_j \lambda_j^{(n)} \left( f_j^{(n)} + \frac{\langle f_j^{(n)}, x_{n+1} - e_{n+1} \rangle}{\langle e_{n+1}, x_{n+1} - e_{n+1} \rangle} (x_{n+1} - e_{n+1}) \right) + b x_{n+1}.$$

Now we decompose

$$x_{n+1} = \sum_{k=1}^n \mu_k^{(n)} f_k^{(n)} + \nu_n e_{n+1}$$

and

$$\begin{aligned} zf &= \sum_{j=1}^{n} a_j \lambda_j^{(n)} \left( f_j^{(n)} + \frac{\overline{\mu_j^{(n)}}}{\overline{\nu_n} - 1}(x_{n+1} - e_{n+1}) \right) + bx_{n+1} \\ &= \sum_{j=1}^{n} a_j \lambda_j^{(n)} f_j^{(n)} + \left( \sum_{\ell=1}^{n} a_\ell \lambda_\ell^{(n)} \frac{\overline{\mu_\ell^{(n)}}}{\overline{\nu_n} - 1} \right) (x_{n+1} - e_{n+1}) + bx_{n+1}. \end{aligned}$$

Because $f_1^{(n)}, ..., f_n^{(n)}, e_{n+1}$ is a basis for $\mathbb{C}^{n+1}$, we deduce the system of $n+1$ equations

$$za_j = a_j \lambda_j^{(n)} + \mu_j^{(n)} \sum_{\ell=1}^{n} a_\ell \lambda_\ell^{(n)} \frac{\overline{\mu_\ell^{(n)}}}{\overline{\nu_n} - 1} + b\mu_j^{(n)}$$

for $j = 1, \ldots, n$ and

$$zb = b + (\nu_n - 1) \sum_{\ell=1}^{n} a_\ell \lambda_\ell^{(n)} \frac{\overline{\mu_\ell^{(n)}}}{\overline{\nu_n} - 1} + b(\nu_n - 1).$$

For $z \notin \{\lambda_1^{(n)}, \ldots, \lambda_n^{(n)}, 1\}$, let us consider the linear transform $Q : \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$ whose matrix representation in the basis $f_1^{(n)}, \ldots, f_n^{(n)}, e_{n+1}$ is

$$Q = I + wv^t,$$

where

$$w = \begin{pmatrix} \frac{\mu_1^{(n)}}{\lambda_1^{(n)} - z} \\ \vdots \\ \frac{\mu_n^{(n)}}{\lambda_n^{(n)} - z} \\ \frac{\nu_n - 1}{1 - z} \end{pmatrix}$$

and

$$v^t = \left( \lambda_1^{(n)} \frac{\overline{\mu_1^{(n)}}}{\overline{\nu_n} - 1}, \cdots, \lambda_n^{(n)} \frac{\overline{\mu_n^{(n)}}}{\overline{\nu_n} - 1}, 1 \right).$$

Then, the above system can be written

$$Qf = 0.$$

Clearly, rank $Q \in \{n, n+1\}$. If it has full rank then $f = 0$, but we assume a priori that $z$ is an eigenvalue for $u_{n+1}$ and so has a non-trivial eigenspace. Thus we must have rank $Q = n$ and

$$0 = Qf = f + w(v^t f).$$

The right hand side can only vanish if $f$ is proportional to $w$, so $f = \alpha w$ for some complex constant $\alpha \in \mathbb{C} \setminus \{0\}$ and $v^t w = -1$. In particular,

$$\sum_{j=1}^{n} \frac{\lambda_j^{(n)} |\mu_j^{(n)}|^2}{\lambda_j^{(n)} - z} + \frac{|\nu_n - 1|^2}{1 - z} = 1 - \overline{\nu_n},$$

as required.

Conversely, if $z \notin \{\lambda_1^{(n)}, \ldots, \lambda_n^{(n)}, 1\}$ satisfies this equation, then

$$Qw = w + w(v^t w) = w + w(-1) = 0,$$

which implies that $w$ is an eigenvector of $u_{n+1}$ for the eigenvalue $z$.

Let us now show that the eigenvalues $z \notin \{\lambda_1^{(n)}, \ldots, \lambda_n^{(n)}, 1\}$ of $u_{n+1}$ strictly interlace between 1 and the eigenvalues of $u_n$: since $u_{n+1}$ has at most $n+1$ eigenvalues, this implies that $\lambda_1^{(n)}, \ldots, \lambda_n^{(n)}, 1$ are not eigenvalues of $u_{n+1}$.

Define the rational function $\Phi : S^1 \to \mathbb{C} \cup \{\infty\}$ by

$$\Phi(z) = \sum_{j=1}^{n} \frac{\lambda_j^{(n)} |\mu_j^{(n)}|^2}{\lambda_j^{(n)} - z} + \frac{|\nu_n - 1|^2}{1 - z} - (1 - \overline{\nu_n})$$

Note that $\Phi$ vanishes precisely on the eigenvalues of $u_{n+1}$ which are different from $\lambda_1^{(n)}, \dots, \lambda_n^{(n)}, 1$. Recalling that $|\mu_1^{(n)}|^2 + \cdots + |\mu_n^{(n)}|^2 + |\nu_n|^2 = 1$, we can rearrange the expression of $\Phi$ to the equivalent form

$$\Phi(z) = \frac{1}{2} \left( \sum_{j=1}^{n} |\mu_j^{(n)}|^2 \frac{\lambda_j^{(n)} + z}{\lambda_j^{(n)} - z} + |1 - \nu_n|^2 \frac{1 + z}{1 - z} - \nu_n + \overline{\nu_n} \right).$$

Hence, $\Phi$ takes values only in $i\mathbb{R} \cup \{\infty\}$, since for all $z \neq z' \in S^1$, $(z + z')/(z - z')$ is purely imaginary (the triangle $(-z', z, z')$ has a right angle at $z$). Note that for $z \in \{\lambda_1^{(n)}, \dots, \lambda_n^{(n)}, 1\}$, a unique term of the sum defining $\Phi$ is infinite, since by assumption, $\mu_1^{(n)}, \dots, \mu_n^{(n)}, 1 - \nu_n$ are nonzero and $\lambda_1^{(n)}, \dots, \lambda_n^{(n)}, 1$ are distinct: $\Phi(z) = \infty$.

Next, we consider $t \mapsto \Phi(e^{it})$ in a short interval $(\theta_j^{(n)} - \delta, \theta_j^{(n)} + \delta)$. Then, for $t = \theta_j^{(n)} + u$ in this interval,

$$\frac{\lambda_j^{(n)} + \lambda_j^{(n)} e^{iu}}{\lambda_j^{(n)} - \lambda_j^{(n)} e^{iu}} = \frac{1 + e^{iu}}{1 - e^{iu}} = 2iu^{-1} + O(1)$$

while the other terms in $\Phi(e^{it})$ are uniformly bounded as $\delta \to 0$; likewise for the interval $(-\delta, \delta)$. In particular, $\Phi \to i\infty$ as $u \to 0$ from the right and $\Phi \to -i\infty$ as $u \to 0$ from the left. We therefore conclude, as $\Phi$ is continuous, that on each interval of the partition

$$(0, \theta_1^{(n)}) \cup (\theta_1^{(n)}, \theta_2^{(n)}) \cup \cdots \cup (\theta_n^{(n)}, 2\pi)$$

of the unit circle into $n + 1$ parts, $t \mapsto \Phi(e^{it})$ must assume every value on the line $i\mathbb{R}$, and in particular must have at least one root. But we know that $\Phi$ has only $n + 1$ roots on the circle so there must be exactly one root in each part of the partition, which proves the interlacing property.

It remains to check the expression of the eigenvectors $(f_k^{(n+1)})_{1 \leq k \leq n+1}$ given in the theorem, but this expression is immediately deduced from the expression of the vector $w$ involved in the operator $Q$ defined above, and the fact that $\|f_k^{(n+1)}\| = 1$.

□

In order to prove the convergence of the normalized eigenangles of $u_n$ when $n$ goes to infinity, we need the following lemma. It should also be noted that all the strong convergence results that follow rely on a priori estimates given in Appendix 7.

LEMMA 4.2. *Let $\epsilon > 0$. Then, almost surely under the Haar measure on $U^\infty$, for $n \geq 1$ and $0 < k \leq n^{1/4}$, we have*

$$\frac{\theta_k^{(n+1)} |\mu_k^{(n)}|^2}{\theta_k^{(n)} - \theta_k^{(n+1)}} = 1 + O(kn^{-\frac{1}{3}+\epsilon})$$

*and for $n \geq 1$ and $-n^{1/4} \leq k \leq 0$,*

$$\frac{\theta_k^{(n+1)} |\mu_k^{(n)}|^2}{\theta_k^{(n)} - \theta_k^{(n+1)}} = \frac{\theta_k^{(n+1)} |\mu_{k+n}^{(n)}|^2}{\theta_k^{(n)} - \theta_k^{(n+1)}} = 1 + O((1 + |k|)n^{-\frac{1}{3}+\epsilon}),$$

REMARK 4.3. *The implied constant in the $O(\cdot)$ notation depends on $(u_m)_{m \geq 1}$ and $\epsilon$: in particular, it is a random variable. However, for given $(u_m)_{m \geq 1}$ and $\epsilon$, it does not depend on k and n.*

PROOF. By symmetry of the situation, we can assume $k > 0$. Moreover, let us fix $\epsilon \in (0, 0.01)$. We will suppose that the event $E := E_0 \cap E_1 \cap E_2, \cap E_3$ holds ( see Appendix 7) , where

$$\begin{aligned}
E_0 &= \{\theta_0^{(1)} \neq 0\} \cap \{\forall n \geq 1, \nu_n \neq 0\} \cap \{\forall n \geq 1, 1 \leq k \leq n, \mu_k^{(n)} \neq 0\} \\
E_1 &= \{\exists n_0 \geq 1, \forall n \geq n_0, |\nu_n| \leq n^{-\frac{1}{2}+\epsilon}\} \\
E_2 &= \{\exists n_0 \geq 1, \forall n \geq n_0, 1 \leq k \leq n, |\mu_k^{(n)}| \leq n^{-\frac{1}{2}+\epsilon}\} \\
E_3 &= \{\exists n_0 \geq 1, \forall n \geq n_0, k \geq 1, n^{-\frac{5}{3}-\epsilon} \leq \theta_{k+1}^{(n)} - \theta_k^{(n)} \leq n^{-1+\epsilon}\}.
\end{aligned}$$

It is possible to do this assumption, since by the result proven in Appendix 7, the event $E$ occurs almost surely. As we will see now, this *a priori* information on the distribution of the eigenvalues of the random virtual isometry implies strong quantitative bounds on the change in eigenvalues of successive unitary matrices.

Recall from Theorem 4.1 that

$$\sum_{j=1}^{n} \frac{\lambda_j^{(n)} |\mu_j^{(n)}|^2}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} + \frac{|1-\nu_n|^2}{1-\lambda_k^{(n+1)}} = 1 - \bar{\nu}_n.$$

By using the $n$-periodictiy of $\lambda_j^{(n)}, \mu_j^{(n)}, f_j^{(n)}$ with respect to $j$, we can write

$$\sum_{j \in J} \frac{\lambda_j^{(n)} |\mu_j^{(n)}|^2}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} + \frac{|1-\nu_n|^2}{1-\lambda_k^{(n+1)}} = 1 - \bar{\nu}_n, \tag{10}$$

where $J$ is the random set of $n$ consecutive integers, such that $\theta_k^{(n+1)} - \pi < \theta_j^{(n)} \leq \theta_k^{(n+1)} + \pi$. Iterating the lower bound on the distance between adjacent eigenvalues, given by the definition of the event $E_3$, we get, for $j \in J \backslash \{k-1, k\}$,

$$|\theta_j^{(n)} - \theta_k^{(n+1)}| \gtrsim |k-j| n^{-\frac{5}{3}-\epsilon},$$

and then

$$|\lambda_j^{(n)} - \lambda_k^{(n+1)}| \gtrsim |k-j| n^{-\frac{5}{3}-\epsilon},$$

since $|\theta_j^{(n)} - \theta_k^{(n+1)}| \leq \pi$.

Likewise, we have by $E_3$, $1 - \lambda_k^{(n+1)} = O(kn^{-1+\epsilon})$, and by $E_2$, $|\mu_j^{(n)}|^2 = O(n^{-1+2\epsilon})$, which gives, for $j \in J \backslash \{k-1, k\}$,

$$\frac{\lambda_j^{(n)}(1-\lambda_k^{(n+1)})|\mu_j^{(n)}|^2}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} \lesssim \frac{k}{|k-j|} n^{-\frac{1}{3}+4\epsilon}.$$

Summing for $j$ in $J \backslash \{k-1, k\}$, which is included in the interval $[k-1-n, k+n]$, gives

$$\sum_{j \in J \backslash \{k-1,k\}} \frac{\lambda_j^{(n)}(1-\lambda_k^{(n+1)})|\mu_j^{(n)}|^2}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} = O(kn^{-\frac{1}{3}+4\epsilon} \log n) = O(kn^{-\frac{1}{3}+5\epsilon}).$$

Now, subtracting this equation from the product of (10) by $1 - \lambda_k^{(n+1)}$, and bounding $\nu_n = O(n^{-\frac{1}{2}+\epsilon})$ (by the property $E_1$) gives us the resulting equation

$$\frac{\lambda_k^{(n)}(1-\lambda_k^{(n+1)})|\mu_k^{(n)}|^2}{\lambda_k^{(n)} - \lambda_k^{(n+1)}} \mathbf{1}_{k \in J} + \frac{\lambda_{k-1}^{(n)}(1-\lambda_k^{(n+1)})|\mu_{k-1}^{(n)}|^2}{\lambda_{k-1}^{(n)} - \lambda_k^{(n+1)}} \mathbf{1}_{k-1 \in J} = -1 + O(kn^{-\frac{1}{3}+5\epsilon}).$$

Next we estimate the first two terms in terms of the eigenangles. We find

$$1 - \lambda_k^{(n+1)} = -i\theta_k^{(n+1)} + O((\theta_k^{(n+1)})^2)$$

and

$$\lambda_j^{(n)} - \lambda_k^{(n+1)} = i(\theta_j^{(n)} - \theta_k^{(n+1)})\lambda_j^{(n)} + O((\theta_j^{(n)} - \theta_k^{(n+1)})^2)$$

for $j = k-1, k$. Collecting terms and using the trivial bounds gives

$$\text{(11)}\quad \frac{\theta_k^{(n+1)}|\mu_k^{(n)}|^2}{\theta_k^{(n)}-\theta_k^{(n+1)}}\left(1+O(kn^{-1+\epsilon})\right)\mathbf{1}_{k\in J} + \frac{\theta_k^{(n+1)}|\mu_{k-1}^{(n)}|^2}{\theta_{k-1}^{(n)}-\theta_k^{(n+1)}}\left(1+O(kn^{-1+\epsilon})\right)\mathbf{1}_{k-1\in J} = 1+O(kn^{-\frac{1}{3}+5\epsilon}).$$

From Theorem 4.1, the eigenvalues of $u_n$ and $u_{n+1}$ interlace, so for $n$ sufficiently large the real part of the first term is positive and the real part of the second term is negative. The real part of the right hand side tends to 1 as $n$ grows with $k$ fixed, so the first term has real part bounded below for $n$ sufficiently large. In particular,

$$\frac{\theta_k^{(n+1)}|\mu_k^{(n)}|^2}{\theta_k^{(n)}-\theta_k^{(n+1)}} \gtrsim 1.$$

Using the a priori bounds for $\theta_k^{(n+1)}$ and $|\mu_k^{(n)}|^2$, we find

$$\theta_k^{(n)}-\theta_k^{(n+1)} \lesssim kn^{-2+3\epsilon}.$$

Hence,

$$\theta_k^{(n+1)}-\theta_{k-1}^{(n)} = (\theta_k^{(n)}-\theta_{k-1}^{(n)})-(\theta_k^{(n)}-\theta_k^{(n+1)}) \gtrsim n^{-\frac{5}{3}-\epsilon}-O(kn^{-2+3\epsilon}) \gtrsim n^{-\frac{5}{3}-\epsilon},$$

since $kn^{-2+3\epsilon}/n^{-\frac{5}{3}-\epsilon} = O(n^{1/4-2+0.03+5/3+0.01}) = o(1)$. We deduce that the second term of (11) is dominated by $kn^{-1/3+4\epsilon}$, and then

$$\frac{\theta_k^{(n+1)}|\mu_k^{(n)}|^2}{\theta_k^{(n)}-\theta_k^{(n+1)}} = 1+O(kn^{-\frac{1}{3}+5\epsilon}).$$

Changing the value of $\epsilon$ appropriately gives the desired result. □

This lemma is enough for us to estimate the change in $\theta_k^{(n)}$ as $n$ grows, and in particular to find a limit for the renormalized angle.

THEOREM 4.4. *There is a sine-kernel point process $(y_k)_{k\in\mathbb{Z}}$ such that almost surely,*

$$\frac{n}{2\pi}\theta_k^{(n)} = y_k + O((1+k^2)n^{-\frac{1}{3}+\epsilon}),$$

*for all $n \geq 1$, $|k| \leq n^{1/4}$ and $\epsilon > 0$, where the implied constant may depend on $(u_m)_{m\geq 1}$ and $\epsilon$, but not on $n$ and $k$.*

PROOF. The proof proceeds exactly as in [4]. It is sufficient to prove the result for $\epsilon$ equal to the inverse of an integer: hence, it is enough to show the estimate for fixed $\epsilon$. By symmetry, one can take $k > 0$. We rearrange the equation in Lemma 4.2 to find

$$|\mu_k^{(n)}|^2 = \left(\frac{\theta_k^{(n)}}{\theta_k^{(n+1)}}-1\right)(1+O(kn^{-\frac{1}{3}+\epsilon}))$$

Because almost surely, $|\mu_k^{(n)}|^2 = O(n^{-1+2\epsilon})$, we get

$$|\mu_k^{(n)}|^2 = \frac{\theta_k^{(n)}}{\theta_k^{(n+1)}}-1+O(kn^{-\frac{4}{3}+3\epsilon}).$$

Using the asymptotic $\log(1-\delta) = -\delta+O(\delta^2)$ for $\delta = o(1)$, we conclude, if $\epsilon$ is small enough,

$$\log\frac{\theta_k^{(n)}}{\theta_k^{(n+1)}} = |\mu_k^{(n)}|^2+O(kn^{-\frac{4}{3}+3\epsilon}).$$

Define the random variable $L_k^{(n)} = \log \theta_k^{(n)} + \sum_{j=1}^{n-1} |\mu_k^{(j)}|^2$; we have just shown $L_k^{(n+1)} - L_k^{(n)} = O(kn^{-\frac{4}{3}+3\epsilon})$ so for $k$ fixed, $L_k^{(n)}$ converges to a limit $L_k^{(\infty)}$ almost surely as $n \to \infty$, with $|L_k^{(n)} - L_k^{(\infty)}| = O(kn^{-\frac{1}{3}+3\epsilon})$. Now,

$$\begin{aligned}
\exp L_k^{(n)} &= \theta_k^{(n)} \exp \sum_{j=k}^{n-1} |\mu_k^{(j)}|^2 \\
&= n\theta_k^{(n)} \exp \left( -\log n + \sum_{j=1}^{n-1} \frac{1}{j} + \sum_{j=1}^{n-1} (|\mu_k^{(j)}|^2 - \frac{1}{j}) \right)
\end{aligned}$$

Recall $-\log n + \sum_{j=1}^{n-1} \frac{1}{j} = \gamma + O(n^{-1})$ where $\gamma$ is the Euler-Mascheroni constant. Next we define

$$M_k^{(n)} := \sum_{j=1}^{n-1} \left( |\mu_k^{(j)}|^2 - \frac{1}{j} \right)$$

and observe that each term of the sum is an independent mean-zero random variable. Therefore, for $k$ fixed, $(M_k^{(n)})_{n \geq k}$ is a martingale. We claim that $M_k^{(n)}$ is bounded in $L^2$; in fact,

$$\mathbb{E}(|\mu_k^{(n)}|^2 - \frac{1}{n})^2 = O(n^{-2}).$$

so that

$$\mathbb{E}((M_k^{(\infty)} - M_k^{(n)})^2) = \sum_{j \geq n} \mathbb{E}(|\mu_k^{(n)}|^2 - \frac{1}{n})^2 = O(n^{-1}),$$

where $M_k^{(\infty)}$ is the claimed limit of $M_k^{(n)}$ (this limit exists since $M_k^{(n)}$ is a sum of centered and independent random variables with summable variances). To see this, we write

$$|\mu_k^{(n)}|^2 = \frac{e_1}{e_1 + \cdots + e_n}$$

where the variables $e_r$ are independent standard exponential random variables. Then we compute

$$\mathbb{E}(|\mu_k^{(n)}|^2 - \frac{1}{n})^2 = \mathbb{E} \left( \frac{(n-1)e_1 - e_2 - \cdots - e_n}{n(e_1 + \cdots + e_n)} \right)^2.$$

As shown before in this paper, $\mathbb{P}(e_1 + \cdots + e_n \leq \frac{n}{2}) = O(n^{-C})$ for all $C \geq 2$ so that

$$\begin{aligned}
\mathbb{E}(|\mu_k^{(n)}|^2 - \frac{1}{n})^2 &\leq O(n^{-C}) + \frac{4}{n^4} \mathbb{E}(((n-1)e_1 - e_2 - \cdots - e_n)^2) \\
&\leq O(n^{-2})
\end{aligned}$$

Now, by the triangle inequality and Doob's maximal inequality, for $q$ positive integer, $k \leq 2^q$,

$$\begin{aligned}
\mathbb{E}(\sup_{n \geq 2^q} (M_k^{(\infty)} - M_k^{(n)})^2) &\lesssim \mathbb{E}((M_k^{(\infty)} - M_k^{(2^q)})^2) + \mathbb{E}(\sup_{n \geq 2^q} (M_k^{(n)} - M_k^{(2^q)})^2) \\
&\lesssim \mathbb{E}(M_k^{(\infty)} - M_k^{(2^q)})^2 \\
&= O(2^{-q}).
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E} \left[ \sup_{2^q \leq n \leq 2^{q+1}} \sup_{k \leq n^{1/4}} (M_k^{(\infty)} - M_k^{(n)})^2 \right] &\leq \mathbb{E} \left[ \sup_{k \leq 2^{(q+1)/4}} \sup_{2^q \leq n \leq 2^{q+1}} (M_k^{(\infty)} - M_k^{(n)})^2 \right] \\
&\leq \sum_{k \leq 2^{(q+1)/4}} \mathbb{E} \left[ \sup_{2^q \leq n \leq 2^{q+1}} (M_k^{(\infty)} - M_k^{(n)})^2 \right] \\
&\lesssim 2^{(q+1)/4} 2^{-q} = O(2^{-3q/4})
\end{aligned}$$

and

$$\mathbb{E}\left[\sup_{n\geq 2^q}\sup_{k\leq n^{1/4}}(M_k^{(\infty)}-M_k^{(n)})^2\right] \leq \sum_{r\geq q}\mathbb{E}\left[\sup_{2^r\leq n\leq 2^{r+1}}\sup_{k\leq n^{1/4}}(M_k^{(\infty)}-M_k^{(n)})^2\right]$$
$$\lesssim \sum_{r\geq q} 2^{-3r/4} = O(2^{-3q/4}).$$

By Markov's inequality, we get

$$\mathbb{P}(\sup_{n\geq 2^q}\sup_{k\leq n^{1/4}}|M_k^{(\infty)}-M_k^{(n)}|\geq 2^{-q/3}) \leq 2^{2q/3}\mathbb{E}(\sup_{n\geq 2^q}\sup_{k\leq n^{1/4}}(M_k^{(\infty)}-M_k^{(n)})^2) = O(2^{-q/12}),$$

which, by Borel-Cantelli lemma, shows that almost surely for some $q_0 \geq 1$, all $q \geq q_0$, $n \geq 2^q$ and $k \leq n^{1/4}$ satisfy $|M_k^{(\infty)}-M_k^{(n)}| \leq 2^{-q/3}$. Hence,

$$|M_k^{(\infty)}-M_k^{(n)}| = O(n^{-\frac{1}{3}})$$

almost surely. Collecting these estimates and applying them to the equation

$$\exp L_k^{(n)} = n\theta_k^{(n)}\exp(\gamma + O(n^{-1}) + M_k^{(n)})$$

gives us

$$\exp\left(L_k^{(\infty)} + O(kn^{-\frac{1}{3}+3\epsilon})\right) = n\theta_k^{(n)}\exp(\gamma + M_k^{(\infty)} + O(n^{-\frac{1}{3}}))$$

Rearranging,

$$n\theta_k^{(n)} = \exp(L_k^{(\infty)} - M_k^{(\infty)} - \gamma)(1 + O(kn^{-\frac{1}{3}+3\epsilon})) =: 2\pi y_k(1 + O(kn^{-\frac{1}{3}+3\epsilon}).$$

Now, by **[4]**, $(y_k)_{k\in\mathbb{Z}}$ is a determinantal sine-kernel process, so we have almost surely the estimate $y_k = O(1+|k|)$, which proves Theorem 4.4. □

CHAPTER 3

# The Ramachandra conjecture and local limit theorems

## 1. A general local limit theorem

In this chapter, we take a closer look at the Ramachandra conjecture, which is the statement that the set $\zeta(1/2+it)$, when $t \in \mathbb{R}$, is dense in $\mathbb{C}$, following **[8]**. In **[20]**, we proved that a quantitative version of the moments conjecture implies the Ramachandra conjecture and we proved the analogue for function fields. Here we wish to show that a much weaker statement than the moments conjecture implies the Ramachandra conjecture. We first start by noting that while Selberg's limit theorem can be viewed as a central limit theorem, Ramachandra's conjecture is more a statement of the local limit theorem type. We hence try to work in a framework in which we have a sequence of random vectors which satisfy a central limit theorem and from which we want to deduce a local limit theorem.

Our results heavily rely on Fourier analysis so we first mention the conventions we shall take. We define the Fourier transform as is usually done in probability theory, namely

$$\hat{f}(t) = \int_{\mathbb{R}^d} \exp(i\, t \cdot x) f(x)\, dx.$$

The inversion formula is, at least when $\hat{f} \in L^1(\mathbb{R}^d)$, given by

$$f(x) = \left(\frac{1}{2\pi}\right)^d \int_{\mathbb{R}^d} \exp(-it \cdot x)\hat{f}(t)\, dt.$$

In particular, when $\mu$ is a probability measure with an integrable characteristic function $\varphi$, we get that $\mu$ is absolutely continuous with respect to Lebesgue measure $m$, and its density is given by

$$\frac{d\mu}{dm}(x) = \left(\frac{1}{2\pi}\right)^d \int \exp(-it \cdot x)\varphi(t)\, dt,$$

which is therefore continuous.

We fix $d \geq 1$ and a probability measure $\mu$ on $\mathbb{R}^d$. We then assume given a sequence $(X_n)$ of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $\mathbb{R}^d$. We define $\varphi_n$ to be the characteristic function of $X_n$. We now consider the following properties:

- **H1.** The characteristic function $\varphi$ of the probability measure $\mu$ is integrable; in particular, $\mu$ has a density $d\mu/dm$, with respect to Lebesgue measure $m$.
- **H2.** There exists a sequence of linear automorphisms $A_n \in \mathrm{GL}_d(\mathbb{R})$, with inverses $\Sigma_n = A_n^{-1}$, such that $\Sigma_n$ converges to 0 and $\varphi_n(\Sigma_n^* t)$ converges continuously at 0 (or what is equivalent: uniformly on compact sets) to $\varphi(t)$. In other words, the renormalized random variables $\Sigma_n(X_n)$ converge in law to $\mu$. (Recall that $\Sigma_n^*$ is the transpose of $\Sigma_n$.)
- **H3.** For all $k \geq 0$, the sequence

  $$f_{n,k} = \varphi_n(\Sigma_n^* t)\mathbf{1}_{|\Sigma_n^* t| \leq k}$$

  is *uniformly integrable* on $\mathbb{R}^d$; since $f_{n,k}$ are uniformly bounded in $L^1$ and $L^\infty$ (for fixed $k$), this is equivalent to the statement that, for all $k \geq 0$, we have

$$\lim_{a \to +\infty} \sup_{n \geq 1} \int_{|t| \geq a} |\varphi_n(\Sigma_n^* t)| \mathbf{1}_{|\Sigma_n^* t| \leq k} dt = 0. \tag{12}$$

DEFINITION 1.1. *[Mod-$\varphi$ convergence] If $\mu$ is a probability measure on $\mathbb{R}^d$ with characteristic function $\varphi$, $X_n$ is a sequence of $\mathbb{R}^d$-valued random variables with characteristic functions $\varphi_n$, and if the properties **H1**, **H2**, **H3** hold, we say that there is* mod$-\varphi$ convergence *for the sequence $X_n$.*

We observe that mod-$\varphi$ convergence will hold when **H1** is true and we have

- **H2′** There exists a sequence of linear automorphisms $A_n \in \mathrm{GL}_d(\mathbb{R})$, with inverses $\Sigma_n = A_n^{-1}$, such that $\Sigma_n$ converges to 0, and there exists a continuous function $\Phi : \mathbb{R}^d \to \mathbb{C}$ such that for arbitrary $k > 0$

$$\varphi_n(t) = \Phi(t)\varphi(A_n^* t)(1 + o(1)) \tag{13}$$

uniformly for $t$ such that $|\Sigma_n^* t| \leq k$.

THEOREM 1.2. *[Local limit theorem for mod-$\varphi$ convergence] Suppose that mod$-\varphi$ convergence holds for the sequence $X_n$. Then we have*

$$|\det(A_n)|\, \mathbb{E}[f(X_n)] \to \frac{d\mu}{dm}(0) \int_{\mathbb{R}^d} f(x)dx,$$

*for all continuous functions with compact support. Consequently we also have*

$$|\det(A_n)|\, \mathbb{P}[X_n \in B] \to \frac{d\mu}{dm}(0)\, m(B). \tag{14}$$

*for relatively compact Borel sets $B \subset \mathbb{R}^d$ with $m(\partial B) = 0$, or in other words for bounded Jordan-measurable sets $B \subset \mathbb{R}^d$.*

The proof relies on the following result from harmonic analysis:

THEOREM 1.3 (see **[8]**). *Suppose $f : \mathbb{R}^d \to \mathbb{R}$ is a continuous function with compact support. Then for each $\eta > 0$ we can find two integrable functions $g_1, g_2 \colon \mathbb{R}^d \to \mathbb{R}$ such that*

(1) *$\widehat{g_1}, \widehat{g_2}$ have compact support,*
(2) *$g_2 \leq f \leq g_1$,*
(3) *$\int_{\mathbb{R}^d}(g_1 - g_2)(t)\, dt \leq \eta$.*

PROOF. We first assume that $f$ is continuous, bounded, integrable and that $\hat{f}$ has compact support; using Theorem 1.3, the case of a general continuous function with compact support will follow easily. We write

$$[f(X_n)] = \int_{\mathbb{R}^d} f(x) d\mu_n(x)$$

where $\mu_n$ is the law of $X_n$. Applying the Parseval formula transforms this into

$$[f(X_n)] = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \varphi_n(t)\hat{f}(-t)\, dt.$$

By the linear change of variable $t = \Sigma_n^* s$, we get

$$\mathbb{E}[f(X_n)] = (2\pi)^{-d} |\det(\Sigma_n)| \int_{\mathbb{R}^d} \varphi_n(\Sigma_n^* s)\hat{f}(-\Sigma_n^* s)\, ds.$$

Now fix $k$ so that the support of $\hat{f}$ is contained in the ball of radius $k$; we then have

$$\mathbb{E}[f(X_n)] = (2\pi)^{-d} |\det(\Sigma_n)| \int_{|\Sigma_n^* s| \leq k} \varphi_n(\Sigma_n^* s)\hat{f}(-\Sigma_n^* s)\, ds.$$

The integrand converges pointwise to $\varphi(s)\hat{f}(0)$ according to the assumption **H2**. The condition **H3** of uniform integrability then implies the convergence in $L^1$. One can see this quickly in this case: for any $\epsilon > 0$, and for any $a > 0$ large enough, we have

$$\int_{|s|>a} |\varphi_n(\Sigma_n^* s)\mathbf{1}_{|\Sigma_n^* s| \leq k}\hat{f}(-\Sigma_n^* s)|\, ds \leq \|\hat{f}\|_\infty \int_{|s|>a} |\varphi_n(\Sigma_n^* s)\mathbf{1}_{|\Sigma_n^* s| \leq k}|\, ds < \epsilon$$

for all $n$ by (12). On $|s| \leq a$, the pointwise convergence is dominated by $\|\hat{f}\|_\infty \mathbf{1}_{|s| \leq a}$, hence

$$\int_{|s| \leq a} \varphi_n(\Sigma_n^* s)\mathbf{1}_{|\Sigma_n^* s| \leq k}\hat{f}(-\Sigma_n^* s)\, ds \to \hat{f}(0) \int_{|s| \leq a} \varphi(s) ds.$$

For $a$ large enough, this is $\hat{f}(0) \int \varphi$, up to error $\epsilon$, hence we get the convergence

$$\int_{|\Sigma_n^* s| \leq k} \varphi_n(\Sigma_n^* s)\hat{f}(-\Sigma_n^* s)\, ds \to \hat{f}(0) \int_{\mathbb{R}^d} \varphi(s) ds.$$

Finally, this leads to

$$|\det(A_n)| [f(X_n)] \to (2\pi)^{-d} \hat{f}(0) \int_{\mathbb{R}^d} \varphi(s) ds = \frac{d\mu}{dm}(0) \int_{\mathbb{R}^d} f(s) ds,$$

which concludes the proof for $f$ integrable and with $\hat{f}$ with compact support. Now if $f$ is continuous with compact support, we use Theorem 1.3: by linearity, we can assume $f$ to be real-valued, and then, given $\eta > 0$ and $g_2 \leq f \leq g_1$ as in the approximation theorem, we have

$$|\det(A_n)|\mathbb{E}[g_2(X_n)] \leq |\det(A_n)|\mathbb{E}[f(X_n)] \leq |\det(A_n)|\mathbb{E}[g_1(X_n)],$$

and hence

$$|\det(A_n)|\mathbb{E}[g_2(X_n)] - \frac{d\mu}{dm}(0)\int g_2(x)dx \\ - \frac{d\mu}{dm}(0)\int (g_1 - g_2)(x)dx \leq |\det(A_n)|\mathbb{E}[f(X_n)] - \frac{d\mu}{dm}(0)\int f(x)dx$$

and

$$|\det(A_n)|\mathbb{E}[f(X_n)] - \frac{d\mu}{dm}(0)\int f(x)dx \leq \\ |\det(A_n)|\mathbb{E}[g_1(X_n)] - \frac{d\mu}{dm}(0)\int g_1(x)dx + \frac{d\mu}{dm}(0)\int (g_1 - g_2)(x)dx$$

and hence

$$\limsup_n \Big| |\det(A_n)|\mathbb{E}[f(X_n)] - \frac{d\mu}{dm}(0)\int f(x)dx \Big| \leq \eta$$

which proves the result since $\eta > 0$ is arbitrary. The proof of (14) is performed in standard ways. □

It can happen that $d\mu/dm(0) = 0$. We can overcome this by the following:

PROPOSITION 1.4. *[Mod-$\varphi$ convergence and shift of the mean] Let $d \geq 1$ be an integer, and let $(X_n)$ be a sequence of $\mathbb{R}^d$-valued random variables such that there is mod$-\varphi$ convergence with respect to the linear maps $A_n$. Let $\alpha \in \mathbb{R}^d$ be arbitrary, and let $\alpha_n \in \mathbb{R}^d$ be a sequence of vectors such that*

$$\lim_{n\to+\infty} \Sigma_n \alpha_n = \alpha, \tag{15}$$

*for instance $\alpha_n = A_n\alpha$. Then the sequence $Y_n = X_n - \alpha_n$ satisfies mod-$\psi$ convergence with parameters $A_n$ for the characteristic function*

$$\psi(t) = \varphi(t)e^{-it\cdot\alpha}.$$

*In particular, for any continuous function $f$ on $\mathbb{R}^d$ with compact support, we have*

$$\lim_{n\to+\infty} |\det(A_n)|\mathbb{E}[f(X_n - \alpha_n)] = \frac{d\mu}{dm}(\alpha)\int_{\mathbb{R}^d} f(x)dx,$$

*where $\mu$ is the probability measure with characteristic function $\varphi$, and for any bounded Jordan-measurable subset $B \subset \mathbb{R}^d$, we have*

$$\lim_{n\to+\infty} |\det(A_n)|\mathbb{P}[X_n - \alpha_n \in B] = \frac{d\mu}{dm}(\alpha)m(B). \tag{16}$$

PROOF. This is elementary: $\psi$ is of course integrable and since

$$\mathbb{E}[e^{itY_n}] = \varphi_n(t)e^{-it\cdot\alpha_n},$$

we have $\mathbb{E}[e^{i\Sigma_n^* t\cdot Y_n}] = \varphi_n(\Sigma_n^* t)e^{-it\cdot\Sigma_n\alpha_n}$, which converges locally uniformly to $\psi(t)$ by our assumption (15). Since the modulus of the characteristic function of $Y_n$ is the same, at any point, as that of $X_n$, Property **H3** holds for $(Y_n)$ exactly when it does for $(X_n)$, and hence mod-$\psi$ convergence holds. If $h = d\mu/dm$, the density of the measure with characteristic function $\psi$ is $g(x) = h(x + \alpha)$, and therefore the last two limits hold by Theorem 1.2. □

## 2. Application to random matrices and the Riemann zeta function

This general local limit theorem surprisingly applies to large variety of situations such as the classical Stone-Feller local limit theorem for sums of independent and identically distributed random variables in the domain of attraction of a stable distribution, the winding number for the planar Brownian motion, as well as many other situations (see **[8]** and **[20]** for more arithmetic examples). The next examples deal with the objects studied in this lecture.

We first start with the characteristic polynomial of random unitary matrices, which corresponds to the case $d = 2$. The results of the previous chapter tell us that $A_n(t) = (\frac{\log n}{2})^{1/2} diag(t_1, t_2)$; $\varphi$ is the characteristic function of the standard Gaussian distribution. Last we are (this is the result of Keating and Sanith) in the situation of **H2′** with

$$\varphi_n(t) = \varphi(A_n^* t)\Phi(t)(1 + o(1))$$

for any fixed $t$, as $n$ goes to infinity, where $\Phi(t) = \frac{G(1+\frac{it_1-t_2}{2})G(1+\frac{it_1+t_2}{2})}{G(1+it_1)}$ . With a little more analysis similar to the one used in Chapter 1 for the Barnes function (**[11]**), we can obtain uniform estimates (see **[20]**)

$$|\varphi_n(t)| \leq C|\Phi(t)\varphi(A_n^* t)|$$

for all $t$ such that $|t| \leq n^{1/6}$, where $C$ is an absolute constant. This immediately gives the uniform integrability for $\varphi_n(\Sigma_n^* t)\mathbf{1}_{|\Sigma_n t| \leq k}$ since $|\Sigma_n^* t|$ is only of logarithmic size with respect to $n$. In other words, we have checked **H3**, and hence there is mod-$\varphi$ convergence.

We can thus state for random unitary matrices:

PROPOSITION 2.1. *Let us note $\mu_n$ for the Haar probability measure on the unitary group $U_n$. Then we have*

$$\lim_{n\to+\infty} |\det(A_n)|\mu_n(g \in U_n;\ \log\det(1-g) \in B) = \frac{m(B)}{(2\pi)^{d/2}}$$

*for any bounded Jordan-measurable set $B \subset \mathbb{C}$*

This can be viewed as a random matrix analogue of Ramachandra. Next we take a look at a similar problem but for the stochastic zeta function. The computations are very similar to those done in the first chapter with the difference that here we do not restrict ourselves to the modulus (hence we work with $d = 2$ but the computations are similar so we do not repeat them and refer the reader to **[8]** for more details).

We recall the finite products finite products $Y_n = \prod_{p\leq n}(1 - \frac{X_p}{\sqrt{p}})$ where the $X_p$ are i.i.d random variables uniformly distributed on the unit circle. The random variables $Z_n$ are then defined as minus the logarithm of $Y_n$, (taken along its principal branch defined as $\log(1) = 0$). So

$$Z_n = -\sum_{p\leq n} \log\left(1 - \frac{Y_p}{\sqrt{p}}\right) = \sum_{p\leq n}\sum_k \frac{1}{k}\left(\frac{X_p}{\sqrt{p}}\right)^k.$$

Then we can show that

$$\varphi_n(t) = \mathbb{E}[\exp(it\cdot Z_n)] = \prod_{p\leq n} {}_2F_1\left(\frac{1}{2}(it_1+t_2), \frac{1}{2}(it_1-t_2); 1; \frac{1}{p}\right),$$

where $t = (t_1, t_2) \in \mathbb{R}^2$, $t \cdot x = t_1x_1 + t_2x_2$ is the inner product in $\mathbb{R}^2$ and ${}_2F_1$ denotes the Gauss hypergeometric function. Straightforward estimates (see **[20]** for details) then give

(1) $|\varphi_n(t)| \leq c(t)\exp(-\frac{1}{16}(\log\log n)|t|^2)$, where $c$ is a non-decreasing function (in fact one can take a constant);
(2) $\varphi_n(\sqrt{\frac{2}{\log\log n}}t) \to \exp(-\frac{1}{2}|t|^2)$.

The conditions of Theorem 1.2 are fulfilled and hence we have

$$\frac{\log\log n}{2}\mathbb{P}\left[Z_n \in B\right] \to \frac{1}{2\pi}m(B).$$

for any bounded Jordan measurable set $B \subset \mathbb{C}$. This is the analogue of the Ramachandra conjecture for the stochastic zeta function.

Based on the random matrix analogy, we can make the following conjecture:

CONJECTURE 2.2. *[Quantitative density of values of $\zeta(1/2+it)$] For any bounded Borel subset $B \subset \mathbb{C}$ with boundary of measure* 0*, we have*

$$\lim_{T\to+\infty} \frac{\frac{1}{2}\log\log T}{T} m\big(u \in [0,T] \mid \log\zeta(1/2+iu) \in B\big) = \frac{m(B)}{2\pi},$$

*where $m(\cdot)$ denotes the Lebesgue measure on $\mathbb{C}$.*

The point is that the following much weaker estimate of the characteristic function of $\log\zeta(1/2+it)$ suffices to prove this:

THEOREM 2.3. *If for all $k > 0$ there exists $C_k \geq 0$ such that*

$$\Big|\frac{1}{T}\int_0^T \exp(it\cdot\log\zeta(1/2+iu))du\Big| \leq \frac{C_k}{1+|t|^4(\log\log T)^2} \tag{17}$$

*for all $T \geq 1$ and $t$ with $|t| \leq k$, then Conjecture 2.2 holds.*

These statements illustrate that the method developed in this chapter can be considered as a new method of potential interest in any number theoretical problem dealing with sets of zero density.

CHAPTER 4

# Appendix

## 1. Classical Techniques

In this appendix, we present the results the way they are usually presented in random matrix theory textbooks on the CUE. The notation might differ from those in the main text. As usual we assume that Weyl's integration formula is known.

We introduce some extra notation to be consistent with the literature.

$$S(x) = \frac{\sin(\pi x)}{\pi x}, \quad S_N(x) = \frac{\sin(Nx/2)}{\sin(x/2)},$$

as well as

$$K_N(x,y) = S_N(x-y), \quad K(x,y) = S(x-y).$$

We recall the following basic facts about Vandermonde determinants

$$\begin{aligned}\Delta(x_1, \cdots, x_N) &= \det\,(x_k^{j-1})_{1\leqslant j,k\leqslant N} \\ &= \prod_{1\leqslant j,k\leqslant N} (x_k - x_j).\end{aligned}$$

Then we observe that

$$\prod_{1\leqslant j,k\leqslant N} \left|e^{i\theta_j} - e^{i\theta_k}\right|^2 = \left|\Delta(l^{i\theta_1}, \cdots, l^{i\theta_N})\right|^2.$$

LEMMA 1.1 (Transposing). *Let $\varphi_j$ and $\psi_K$ be measurable functions. Then the following holds*

$$\begin{aligned}&\det\,(\phi_{j-1}(x_k))_{1\leqslant j,k\leqslant N} \det\,(\psi_{j-1}(x_k))_{1\leqslant j,k\leqslant N} \\ &= \det\left(\sum_{n=1}^{N} \phi_{n-1}(x_j)\psi_{n-1}(x_k)\right)_{1\leqslant j,k\leqslant N}.\end{aligned}$$

PROOF. Follows from: $\forall A, B$ complex matrices of size $N \times N$ we havedet$(A) = \det(A^T)$ and $\det(A)\det(B) = \det(AB)$. □

Now we give an alternative formula for the Haar measure

$$\prod_{1\leqslant j<k\leqslant N} \left|e^{i\theta_k} - e^{i\theta_j}\right|^2 = \det\,(S_N(\theta_k - \theta_j))_{1\leqslant j<k\leqslant N}.$$

To see it, we apply the transposing lemma to

$$\phi_j(\theta_k) = e^{ij\theta_k}, \quad \psi_j(\theta_k) = e^{-ij\theta_k}.$$

We note that

$$\begin{aligned}\sum_{n=1}^{N} e^{i(n-1)\theta} &= \frac{e^{iN\theta} - 1}{e^{i\theta} - 1} = \frac{e^{iN\theta/2}}{e^{i\theta/2}} \frac{\sin(N\theta/2)}{\sin(\theta/2)} \\ &= e^{i(N-1)\theta/2} S_N(\theta).\end{aligned}$$

This implies that

$$\begin{aligned}\left|\Delta(e^{i\theta_1}, \cdots, e^{i\theta_N})\right|^2 &= \det(e^{iN(\theta_1+\cdots+\theta_N)2} S_N(\theta_1 \cdots \theta_N)) \\ &= \det(S_N(\theta_j - \theta_k)).\end{aligned}$$

(Factor $e^{iN\theta_j/2}$ from the $j$th line and $e^{-iN\theta_k/2}$ from the $k$th column and note that the product is 1).

If we formally write the Haar measure $dX_N$ (for functions which only depend on conjugacy classes), then

$$dX_N = \det\left(S_N(\theta_j - \theta_k)\right)_{1\leqslant j<k\leqslant N} \frac{d\theta_1 \cdots d\theta_N}{(2\pi)^N N!}.$$

Another useful lemma is the following result.

LEMMA 1.2 (Andreirf's lemma). *For any interval $J$, and integrable functions $\phi_j$ and $\psi_j$, we have*

$$\begin{aligned}&\frac{1}{N!}\int_{J^N}\cdots\int_{J^N} \det(\phi_j(\theta_k))\det(\psi_j(\theta_k))d\theta_1\cdots d\theta_N\\ &= \det\left(\int_J \phi_j(\theta)\psi_j(\theta)d\theta\right).\end{aligned}$$

Note how $n$ integrals in the LHS become just one integral on the RHS.

PROOF. Recall that if $X = (x_{jk})$ then

$$\det X = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{j=1}^N x_{j,\sigma(j)}.$$

Thus, we may write

$$\begin{aligned}&\int_{J^N} \det(\phi_j(\theta_k))\det(\psi_j(\theta_k))d\vec{\theta}\\
&= \int_{J^N} \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{j=1}^N \phi_j(\theta_{\sigma(j)}) \sum_{\tau\in S_n}\operatorname{sgn}(\tau)\prod_{k=1}^N \psi_k(\theta_{\tau(k)})d\vec{\theta}\\
&\overset{(a)}{=} \int_{J^N} \sum_{\sigma,\tau}\operatorname{sgn}(\tau)\prod_{j=1}^N \phi_j(\theta_{\sigma(j)})\prod_{k=1}^N \psi_k(\theta_{\sigma(\tau(k))})d\vec{\theta}\\
&\overset{(b)}{=} \int_{J^N} \sum_{\sigma,\tau}\operatorname{sgn}(\tau)\prod_{j=1}^N \phi_j(\theta_{\sigma(j)})\prod_{k=1}^N \psi_{\tau^{-1}(k)}(\theta_{\sigma(k)})d\vec{\theta}\\
&= \int_{J^N} \sum_{\sigma,\tau}\operatorname{sgn}(\tau)\prod_{j=1}^N \phi_j(\theta_{\sigma(j)})\psi_{\tau^{-1}(j)}(\theta_{\sigma(j)})d\vec{\theta}\\
&\overset{(c)}{=} \int_{J^N} \sum_{\sigma,\tau}\operatorname{sgn}(\tau)\prod_{j=1}^N \phi_j(\theta_{\sigma(j)})\psi_{\tau(j)}(\theta_{\sigma(j)})d\vec{\theta}\\
&= \sum_{\sigma,\tau}\operatorname{sgn}(\tau)\prod_{j=1}^N \int_J \phi_j(\theta)\psi_{\tau(j)}(\theta)d\theta\\
&= N!\sum_{\tau}\operatorname{sgn}(\tau)\prod_{j=1}^N \int_J \phi_j(\theta)\psi_{\tau(j)}(\theta)d\theta\\
&= N!\det\left(\int_J \phi_j(\theta)\psi_{\tau(j)}(\theta)d\theta\right),\end{aligned}$$

where (a) follows by letting $\tau \to \sigma^2$ with $\sigma$ fixed, (b) since $a \in G$ (a group), so $Ta : G \to G$ such that $x \mapsto ax$ and $k \to \tau^{-1}k$ is a bijection, and (c) because $\tau \to \tau^{-1}$ is a bijection. □

We note that the same proof leads to the more general result

$$\frac{1}{N!}\int_{J^N}\prod_{i=1}^N f(\theta_i)\det(\phi_j(\theta_k))\det(\psi_j(\theta_k))d\theta_1\cdots d\theta_N = \det\left(\int_J f(\theta)\phi_j(\theta)\psi_{\tau(j)}(\theta)d\theta\right).$$

LEMMA 1.3 (Gaudin's lemma). *Let $f$ be a measurable map and $J$ an interval such that*

$$\forall x,y \quad \int_J f(x,\theta)f(\theta,y)d\theta = Cf(x,y)$$

*where C is a constant. Assume further that*

$$\int_J f(x,x)dx = D$$

*then*

$$\int_J \det_{M\times M}(f(\theta_j,\theta_k))d\theta_M = (D-(M-I)C) \det_{(M-1)\times(M-1)} f(\theta_j,\theta_k).$$

EXAMPLE 1.4. *Take $f(x,y) = S_N(x,y)$ and interval $J = [0,2\pi]$. Then*

$$D = \int_J S_N(0)d\theta = 2\pi N.$$

*We have the following*

$$\int_0^{2\pi} S_N(\theta_j-\theta)S_N(\theta-\theta_k)d\theta = 2\pi S_N(\theta_k-\theta_j) \quad \Rightarrow \quad C = 2\pi,$$

*as well as*

$$\int_0^{2\pi} \det_{N\times N}(S_N(\theta_j-\theta_k))d\theta_N = 2\pi \det_{(N-1)\times(N-1)} S_N(\theta_k-\theta_j).$$

*Assuming successively Gaudin's lemma we have*

$$\int_{[0,2\pi]^{N-n}} \det_{N\times N}(S_N(\theta_j-\theta_k))d\theta_{n+1}\cdots d\theta_N = (N-n)!(2\pi)^{N-n}\det_{n\times n}(S_N(\theta_j-\theta_k)).$$

We note the following remark. Let $f:\mathbb{R}^n\to\mathbb{R}$ be a suitable test function. Then

$$\int_{U(N)} \sum_{J\subset\{1,\cdots,N\}} f(\theta_{j_1},\cdots,\theta_{j_n})dX_{U(N)} = \mathbb{E}\left[\sum_{J\subset\{1,\cdots,N\},J=\{j_1,\cdots,j_n\}} f(\theta_{j_1},\cdots,\theta_{j_n})\right].$$

By Gaudin's lemma we have that this is equal to

$$= \frac{1}{(2\pi)^n n!}\int_{[0,2\pi]^{N-n}} f(\theta_1,\cdots,\theta_n)\det_{n\times n}(S_N(\theta_j-\theta_k))d\theta_{n+1}\cdots d\theta_N.$$

PROOF OF GAUDIN'S LEMMA.

$$\det_{M\times M} f(\theta_j,\theta_k) = \sum_{\sigma\in S_M}\operatorname{sgn}(\sigma)\prod_{j=1}^M f(\theta_j,\theta_{\sigma(j)}).$$

We consider two cases: $\sigma(M)\neq M$ and $\sigma(M)=M$. In the former we have

$$\begin{aligned}\int_J\prod_{j=1}^M f(\theta_j,\theta_{\sigma(j)})d\theta_M &= \prod_{j=1,\,\sigma(j)\neq M}^{M-1} f(\theta_j,\theta_{\sigma(j)})\int_J f(\theta_{\sigma^{-1}(M)},\theta_M)f(\theta_M,\theta_{\sigma(M)})d\theta \\ &= \prod_{j=1,\,\sigma(j)\neq M}^{M-1} f(\theta_j,\theta_{\sigma(j)})CF(\theta_{\sigma^{-1}(M)},\theta_M)\end{aligned}$$

For $\sigma\in S_M$ such that $\sigma M\neq M$ we can construct $\sigma'\in S_{M-1}$ as follows

$$\sigma'(j) = \begin{cases}\sigma(j) & \text{if } \sigma(j)\neq M\\ \sigma M & \text{if } \sigma(j) = M\end{cases}$$

thus the above is equal to

$$= C\prod_{j=1}^{M-1} f(\theta_j,\theta_{\sigma'(j)}).$$

Clearly each permutation $\sigma'$ can be obtained from $M-1$ permutations of $S_M$ and $\operatorname{sgn}(\sigma') = -\operatorname{sgn}(\theta)$. Thus we have

$$\begin{aligned}\int_J\sum_{\sigma\in S_M,\,\sigma(M)\neq M}\operatorname{sgn}(\sigma)\prod_{j=1}^M f(\theta_j,\theta_{\sigma(j)})d\theta_M &= -(M-1)C\sum_{\sigma'\in S_{M-1}}(\operatorname{sgn}(\sigma'))\prod_{j=1}^M f(\theta_j,\theta_{\sigma'(j)}) \\ &= -(M-1)C\det_{(M-1)\times(m-1)} f(\theta_j,\theta_k).\end{aligned}$$

Now we go case $\sigma(M) = M$ which yields

$$\begin{aligned}\int_J \prod_{j=1}^{M} f(\theta_j, \theta_{\sigma(j)}) d\theta_M &= \prod_{j=1,\, \sigma(j)\neq M}^{M-1} f(\theta_j, \theta_{\sigma(j)}) \int_J f(\theta_{\sigma^{-1}(M)}, \theta_M) f(\theta_M, \theta_{\sigma(M)}) d\theta \\ &= D \prod_{j=1,\, \sigma(j)\neq M}^{M-1} f(\theta_j, \theta_{\sigma(j)}) F(\theta_{\sigma^{-1}(M)}, \theta_M).\end{aligned}$$

This implies that

$$\begin{aligned}\int_J \prod_{j=1}^{M} f(\theta_j, \theta_{\sigma(j)}) d\theta_M &= D \sum_{\sigma' \in S_{M-1}} \operatorname{sgn}(\sigma') \prod_{j=1}^{M-1} f(\theta_j, \theta_{\sigma'(j)}) \\ &= D \det_{(M-1)\times(M-1)} f(\theta_j, \theta_k)\end{aligned}$$

Putting both cases together yields the lemma. □

As an application we give the classical pair correlation computation which historically played an important role. As before

$$\tilde{\theta}_j = \frac{N}{2\pi} \theta_j.$$

We want to compute the following expression for suitable test functions (to be determined by the reader):

$$Q_N(f) = \int_{U(N)} \sum_{j<k} f(\tilde{\theta}_j - \tilde{\theta}_k) dX_{U(N)}.$$

Gaudin's lemma implies that

$$\begin{aligned}Q_N(f) &= \int_{[0,2\pi]^2} f(\tilde{\theta}_1 - \tilde{\theta}_2) \det_{2\times 2} \begin{pmatrix} N & S_N(\theta_1-\theta_2) \\ S_N(\theta_1-\theta_2) & N \end{pmatrix} \frac{d\tilde{\theta}_1 d\tilde{\theta}_2}{(2\pi)^2 2!} \\ &= \frac{1}{2} \int_{[0,2\pi]^2} f(\theta_1 - \theta_2) \det_{2\times 2} \left( \frac{1}{N} S_N\left( \frac{2\pi(\theta_1-\theta_2)}{N} \right) \right) d\theta_1 d\theta_2 \\ &= \frac{1}{2} \int_{[0,N]^2} f(\theta_1 - \theta_2) \left[ 1 - \left( \frac{S_N(2\pi(\theta_1-\theta_2)/N)}{N} \right)^2 \right] d\theta_1 d\theta_2.\end{aligned}$$

We make the change of variables $u = \theta_1$ and $v = \theta_1 - \theta_2$ so that

$$Q_N(f) = \int_{[-N,N]} f(v) \left[ 1 - \left( \frac{S_N(2\pi(\theta_1-\theta_2)/N)}{N} \right)^2 \right] (N - |v|) dv,$$

where the last term comes from $\max(v, 0) \leq u \leq \min(N, N - v)$. We have already seen that

$$\lim_{N\to\infty} \frac{1}{N} S_N\left( \frac{2\pi v}{N} \right) = S(v)$$

and this implies that

$$\frac{1}{N} Q_N(f) \underset{N\to\infty}{\to} \frac{1}{2} \int_{-\infty}^{\infty} f(v) \left[ 1 - \left( \frac{\sin(\pi v)}{\pi v} \right)^2 \right] dv.$$

Now, it follows that

$$\frac{1}{N} \int_{U(N)} \sum_{j\neq k} f(\tilde{\theta}_j - \tilde{\theta}_k) dX_{U(N)} \underset{N\to\infty}{\to} \int_{-\infty}^{\infty} f(v) \left[ 1 - \left( \frac{\sin(\pi v)}{\pi v} \right)^2 \right] dv.$$

It is the same as what one expect to hold for the zeros of $\zeta(s)$ on the critical line if $RH$ is true (Montgomery's conjecture).

## 2. Beta-Gamma algebra and uniform vectors

First recall that the Gamma function is defined by

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$$

for $\alpha > 0$ (or $\mathrm{Re}(\alpha) > 0$). An integration by parts shows that

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$$

and since $\Gamma(1) = 1$ we then have $\Gamma(n) = (n-1)!$ when $n$ is an integer. We also have Stirling's formula

$$n! \sim \sqrt{2\pi}e^{-n}n^{n+1/2}$$

as $n \to \infty$.

DEFINITION 2.1. *A Gamma random variable with parameter $\alpha > 0$ noted $\gamma_\alpha$ has density given by*

$$\mathbb{P}(\gamma_a \in dt) = \frac{t^{\alpha-1}e^{-t}}{\Gamma(\alpha)}\mathbf{1}_{\mathbb{R}^+}(t)dt.$$

*A Beta random variable with parameters $a > 0$ and $b > 0$ and noted* $\mathrm{B_{a,b}}$ *has density*

$$\mathbb{P}(\mathrm{B}_{a,b} \in dt) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}t^{a-1}(1-t)^{b-1}\mathbf{1}_{[0,1]}(t)dt$$

LEMMA 2.2. *The following statements hold.*

(1) *One has*

$$\mathbb{E}[e^{in\gamma_a}] = \frac{1}{(1-in)^a}.$$

(2) *If $X$ is exponential with mean $1/\gamma$ with $\gamma > 0$ then*

$$\mathbb{E}[e^{inX}] = \frac{1}{\lambda - in}.$$

Let $S^n$ be the unit sphere in $\mathbb{R}^n$, i.e. $S^n = \{x \in \mathbb{R}^n | \, \|x\| = 1\}$. The mapping $P$ such that

$$R^n \backslash \{0\} \to ]0,\infty[\times S^n$$

with

$$u \mapsto \left(\|u\|, \frac{u}{\|u\|}\right)$$

is a homeomorphism. For $A$ measurable, we define $\sigma_n(A) = n\lambda_n(\tilde{A})$ where $\lambda_n$ is the Lebesgue measure in $\mathbb{R}^n$ and $\tilde{A}$ is

$$\tilde{A} = \{ru | r \in [0,1], \ u \in A\}$$

$\sigma_n$ defines the surface/area measure (or uniform measure) on $S^n$. For $g$ nonnegative Borel measurable on $S^n$,

$$\int_{S^n} g d\sigma_n = n\int_{\|x\|\leqslant 1} g\left(\frac{x}{\|x\|}\right)d\lambda_n(x).$$

In particular,

$$\sigma_n(S_n) = nV_n = \frac{n\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})},$$

where $V_n$ is the volume of the unit ball. Moreover, for any Borel measurable $\geq 0$ and integrable function $f$ on $\mathbb{R}^n$ we have

$$\int_{\mathbb{R}^n} f d\lambda_n = \int_{\mathbb{R}_+}\int_{S^n} r^{n-1}f(ru)drd\sigma_n(u).$$

In spherical coordinates $u = (u_1, \cdots, u_n)$

$$\begin{aligned} u_1 &= \sin\phi_1 \\ u_2 &= \cos\phi_1\sin\phi_2 \\ &\vdots \\ u_n &= \cos\phi_1\cos\phi_2\cdots\sin\phi_{n-1} \end{aligned}$$

with $\phi_j \in [-\pi/2, \pi/2]$ and $\phi_{n-1} \in (\pi, \pi)$. The differential element is

$$d\sigma_n = \cos^{n-2}\phi_1 \cdots \cos\phi_{n-2} d\phi_1 \cdots d\phi_{n-1}.$$

COROLLARY 2.3. *Let $X_1, \cdots, X_n$ be iid $\mathcal{N}(0,1)$ random variables. Then*

$$R = \sqrt{x_1^2 + \cdots + x_n^2}$$

*and*

$$U = \frac{X}{R}, \quad X = (X_1, \cdots, X_n)$$

*are independent and $U$ is uniformly distributed on the sphere $S^n$.*

PROOF. Use the above integral

$$\mathbb{E}[f(R,U)] = \int_{\mathbb{R}^n} f\left(r, \frac{x}{r}\right) \exp\left(-\frac{1}{2}\sqrt{x_1^2 + \cdots + x_n^2}\right) \frac{dx_1 \cdots dx_n}{\sqrt{(2\pi)^n}}$$

$$= \int_{\mathbb{R}_+} \int_{S^n} f(r,u) \frac{e^{-r^2/2}}{(2\pi)^{n/2}} r^{n-1} dr d\sigma_n(u)$$

Since the densities split up, we have the independence. □

PROPOSITION 2.4. *Let $Z_a$ be a Gamma variable with parameter $a > 0$ and let $Z_{a,b}$ be a Beta variable with parameters $a > 0$ and $b > 0$.*

(1) *The following identity holds in law*

$$(Z_a, Z_b) \overset{\text{law}}{=} (Z_{a,b}Z_{a+b}, (1 - Z_{a,b})Z_{a+b})$$

*all variables on the left hand side are independent as well as tose on the RHS.*

(2) *We also have the identity in law*

$$Z_{a,b+c} \overset{\text{law}}{=} Z_{a,b}Z_{a+b,c}$$

(3) *Let $h \in \mathbb{N}$ and $X_1, \cdots, X_h$ be $h$ independent $\mathcal{N}(0,1)$ variables. We note*

$$R_h := \left(\sum_{j=1}^{h} X_j^2\right)^{1/2}$$

*Then*

$$R_h^2 \overset{\text{law}}{=} 2Z_{h/2}.$$

PROOF. Let $f$ be any bounded Borel function defined on $\mathbb{R}_+$.

$$\begin{aligned}
&\mathbb{E}[f(Z_{a,b}Z_{a+b}, (1 - Z_{a,b})Z_{a+b})] \\
&= \int_0^1 \int_0^\infty dxdyf(xy, (1-x)y) \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{x^{a-1}(1-x)^{b-1}y^{a+b-1}e^{-y}}{\Gamma(a+b)} \\
&= \frac{1}{\Gamma(a)\Gamma(b)} \int_0^\infty \int_0^\infty f(z,t) z^{a-1}e^{-z}t^{b-1}e^{-t} dzdt \\
&= \mathbb{E}[f(Z_a, Z_b)]
\end{aligned}$$

by the change of variables $z = xy$ and $t = (1-x)y$. This proves the first part. The second part is equivalent to showing

$$Z_{a+b+c}Z_{a,b+c} \overset{\text{law}}{=} Z_{a+b+c}Z_{a,b}Z_{a+b,c},$$

where $Z_{a+b+c}$ is independent of the other variables. (Technically this (i.e. multiplying with $Z_{a+b+c}$) is illegal. Ex: $\Sigma_1\Sigma_2 \overset{\text{law}}{=} \Sigma_1$ does not imply $\Sigma_2 = 1$.) But we know from part 1 that

$$Z_{a+b+c}Z_{a,b+c} \overset{\text{law}}{=} Z_a, \quad \text{and} \quad Z_{a+b+c}Z_{a+b,c} \overset{\text{law}}{=} Z_{a+b}.$$

Hence it is equivalent to showing that

$$Z_a \overset{\text{law}}{=} Z_{a+b}Z_{a,b},$$

which is true (from part 1). This proves part 2. What is $X_1^2$ in law? Let $f$ be Borel and bounded

$$\begin{aligned}\mathbb{E}[f(X_1^2)] &= \int_{-\infty}^{\infty} f(x^2)e^{-x^2/2}\frac{dx}{\sqrt{2\pi}} \\ &= 2\int_0^{\infty} f(x^2)e^{-x^2/2}\frac{dx}{\sqrt{2\pi}} \\ &\underset{u=x^2}{=} 2\int_0^{\infty} f(u)e^{-u/2}\frac{du}{\sqrt{2\pi}2x} \\ &= \mathbb{E}[f(2Z_{1/2})]\end{aligned}$$

Hence

$$E[e^{inR_n^2}] = E[e^{inX_i^2}]^h = E[e^{2inZ_{1/2}}]^h = \frac{1}{(1-2in)^{h/2}}$$

but $1/(1-2in)^{h/2}$ is the characteristic function of $2Z_{h/2}$, this implies that $R_n^2 \overset{\text{law}}{=} 2Z_{h/2}$ and therefore part 3 is proved. □

It follows from the proposition that $X_1^2 + X_2^2 \overset{\text{law}}{=} \exp(1/2)$, density $\frac{1}{2}e^{-1/2}dx$. Furthermore, let $(Z_h)_{h\leq n}$ be independent standard complex Gaussian random variables then

$$\frac{Z}{\|Z\|} = \left(\frac{Z_1}{\sqrt{Z_1^2+\cdots+Z_n^2}},\cdots,\frac{Z_n}{\sqrt{Z_1^2+\cdots+Z_n^2}}\right)$$

is uniformly distributed on $\mathscr{S}_{\mathbb{C}}^n$.

THEOREM 2.5 (Poincaré's theorem). *Fix $h \in \mathbb{N}$ and let $n > 0$. Let $Y = (Y_1^{(n)},\cdots,Y_h^{(n)})$ be uniformly distributed on $S^n$. Then*

$$\sqrt{n}(Y_1^{(n)},\cdots,Y_h^{(n)}) \underset{n\to\infty}{\overset{\text{law}}{\to}} (X_1,\cdots,X_h),$$

*where $X_1,\cdots,X_h$ are independent $\mathcal{N}(0,1)$ variables.*

PROOF. We know from the previous Corollary (1.9) that

$$\begin{aligned}(Y_1^{(n)},\cdots,Y_h^{(n)}) &\overset{\text{law}}{=} \frac{1}{\sqrt{2Z_{h/2}}}(X_1,\cdots,X_h), \\ \sqrt{n}(Y_1^{(n)},\cdots,Y_h^{(n)}) &\overset{\text{law}}{=} \frac{\sqrt{n}}{\sqrt{2Z_{h/2}}}(X_1,\cdots,X_h).\end{aligned}$$

Now, from the expression of the characteristic function of the Gamma variables of parameter $\alpha > 0$, it follows that

$$2Z_{h/2} \overset{\text{law}}{=} 2\tilde{Z}_{1/2}^{(1)} + \cdots + 2\tilde{Z}_{1/2}^{(h)}$$

where $\tilde{Z}_{1/2}^{(i)}$ are independent Gamma variables with parameter $1/2$. Moreover,

$$\mathbb{E}[\tilde{Z}_{1/2}^{(i)}] = \frac{1}{2}$$

The law of large numbers gives

$$\frac{2Z_{h/2}}{h} \to 1, \quad \text{a.s.}$$

and therefore

$$\frac{\sqrt{n}}{\sqrt{2Z_{n/2}}} \to 1, \quad \text{a.s.}$$

□

### 3. Distribution of elements of random unitary matrices $U \in U(n)$

Clearly the entries of Haar distributed random matrices are not independent. We already know that the first column is uniformly distributed on the unit sphere $\mathbb{S}^n_{\mathbb{C}}$. The Haar measure is invariant under left and right translations, and since permutation matrices are in $U(n)$ we deduce that each column and row has the same distribution (in fact, each element).
Let $U = (U_{ij})_{1\leq i,j\leq n}$ then the $U_{ij}$ all have the same distribution. If $u_1$ is the first column of $U_1$

$$U_1 \overset{\text{law}}{=} \left(\frac{X_1}{|X|}, \cdots, \frac{X_n}{|X|}\right)$$

where $X_1, \cdots, X_n$ are independent standard complex Gaussian random variables and

$$|X| = \sqrt{|X_1|^2 + \cdots + |X_n|^2}.$$

THEOREM 3.1. *Each entry of $U$ is distributed like $e^{i\theta}\sqrt{\mathrm{B}_{1,n-1}}$ where $\sigma$ is uniformly distributed on $[0, 2\pi]$ and $\mathrm{B}_{1,n-1}$ is a Beta variable with parameters $(1, n-1)$ and which is independent of $\theta$. The convention is $\mathrm{B}_{1,0} = 1$ almost surely.*

In other words, each entry has density

$$\frac{n-1}{\pi}(1-n^2)^{n-2} r dr d\theta, \quad r \in [0,1],\ \theta \in [0, 2\pi]$$

PROOF. We know that each element has the same distribution as

$$\frac{X_1}{|X|}, \quad X_1 \overset{\text{law}}{=} e^{i\theta}\sqrt{2R}.$$

Let $f : [0,1] \to \mathbb{R}$ be bounded. One has

$$\begin{aligned}
\mathbb{E}\left[f\left(\frac{X_1}{|X|}\right)\right] &= \mathbb{E}\left[f\left(\frac{\sqrt{2Z_1}}{\sqrt{2Z_1 + 2Z_{n-1}}}\right)\right] \\
&= \mathbb{E}\left[f\left(\frac{\sqrt{2Z_{1,n-1}Z_n}}{\sqrt{2Z_{1,n-1}Z_n + (1 - Z_{1,n-1})Z_n}}\right)\right] \\
&= \mathbb{E}[f\sqrt{Z_{1,n-1}}].
\end{aligned}$$

We also have

$$\begin{aligned}
\mathbb{P}[|U_{ij}| \leqslant r] &= \mathbb{P}[\sqrt{\mathrm{B}_{1,m-1}} \leqslant r] \\
&= \mathbb{P}[\mathrm{B}_{1,m-1} \leqslant r^2] \\
&= \int_0^{r^2} \frac{\Gamma(n)}{\Gamma(n-1)}(1-x)^{n-2} dx \\
&= 1 - (1-r^2)^{n-1}.
\end{aligned}$$

The density of $|U_{ij}|$ is $2r(1-r^2)^{n-2}$. □

COROLLARY 3.2. *One has*

$$\sqrt{n}U_{ij} \underset{n\to\infty}{\overset{\text{law}}{\to}} X + iY$$

*where $X$ and $Y$ are independent $\mathcal{N}(0, 1/2)$ variables.*

PROOF. We compute $U_{ij} = e^{i\theta}\sqrt{\mathrm{B}_{1,\mathrm{n}-1}}$. We note that

$$\mathbb{P}[\sqrt{n}|U_{ij}|^2 \leqslant x] = 1 - \left(1 - \frac{x}{n}\right)^{n-1} \to 1 - e^{-x}$$

so

$$\sqrt{n}U_{ij} \underset{n\to\infty}{\overset{\text{law}}{\to}} e^{i\theta}\sqrt{R}$$

where $R$ is a standard exponential variable. But

$$e^{i\theta}\sqrt{R} = \frac{1}{\sqrt{2}}e^{i\theta}\sqrt{2R}.$$

□

We remark that one can also compute the moments of $|U_{ij}|$. For $n \geq 1$, we have

$$\mathbb{E}[|U_{ij}|^{2k}] = \mathbb{E}[\mathrm{B}_{1,n-1}^k] = \frac{\Gamma(n)}{\Gamma(n-1)} \int_0^1 x^k (1-x)^{n-2} dx$$

and we know that

$$\int_0^1 t^{a-1}(1-t)^{b-1} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

which imples that

$$\begin{aligned}
\mathbb{E}[|U_{ij}|^{2k}] &= \frac{\Gamma(n)}{\Gamma(n-1)} \frac{\Gamma(k+1)\Gamma(n-1)}{\Gamma(k+n)} \\
&= \frac{\Gamma(n)\Gamma(k+1)}{\Gamma(k+n)} \\
&= \begin{pmatrix} n+k-1 \\ n-1 \end{pmatrix}^{-1}
\end{aligned}$$

DEFINITION 3.3. *For $n \geq 2$ and $a_1, \cdots, a_n > 0$, let* $\mathrm{Dir}(a_1, \cdots, a_n)$ *be the Dirichlet distribution on the simplex*

$$\left\{ (x_1, \cdots, x_n) \in [0,1]^n : \sum_{k=1}^n x_k = 1 \right\}$$

*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 $a_1 = a_2 = \cdots = a_n = a$, it is called the Dirichlet distribution with parameter $a$ and is noted* $\mathrm{Dir}(a)$*. If $a_1 = a_2 = \cdots = a_n = 1$, then this is the uniform distribution on the simplex.*

LEMMA 3.4. *Let $(Z_{a_i})$ with $1 \leq i \leq k$ be $k$ independent gamma variables with respective parameters $a_i$. Then*

$$\sum_{i=1}^k Z_{a_i}$$

*is independent of*

$$\left( Z'_{a_j} \equiv \frac{Z_{a_j}}{\sum_{i=1}^k Z_{a_i}} \right)_{1 \leqslant j \leqslant k},$$

*and this vector above is uniformly distributed on the simplex.*

PROOF. Since

$$\sum_{j=1}^k \frac{Z_{a_j}}{\sum_{i=1}^k Z_{a_i}} = 1$$

it suffices to show that $\sum_{i=1}^k Za_i$ is independent of

$$\left( \frac{Z_{a_j}}{\sum_{i=1}^k Z_{a_i}} \right)_{1 \leqslant j \leqslant k-1}.$$

Let $f : \mathbb{R}_+ \to \mathbb{R}$, $g : \mathbb{R}_+^{k-1} \to \mathbb{R}$ be bounded Borel functions. We want to compute

$$\mathbb{E}\left[ f\left( \sum_{i=1}^k Z_{a_i} \right) g(Z'_{a_1}, \cdots, Z'_{a_{k-1}}) \right].$$

We make a change of variables

$$\left\{ \begin{array}{c} y_1 = \frac{x_1}{y_k} \\ \vdots \\ y_{k-1} = \frac{x_{k-1}}{y_k} \\ y_k = x_1 + \cdots + x_k \end{array} \right. \Leftrightarrow \left\{ \begin{array}{c} x_1 = y_1 y_k \\ \vdots \\ x_{k-1} = y_{k-1} y_k \\ x_k = y_k - \sum_{i=1}^{k-1} y_i y_k \end{array} \right.$$

Thus we have

$$\mathbb{E}\left[f\left(\sum_{i=1}^{k} Z_{a_i}\right) g(Z'_{a_1},\cdots,Za_{k-1}')\right] = \int_{\mathbb{R}_+^k} f(x_1+\cdots+x_k) g\left(\frac{x_1}{x_1+\cdots+x_k},\cdots,\frac{x_{k-1}}{x_1+\cdots+x_k}\right)$$

$$\times \frac{x_1^{a_1-1}\cdots x_k^{a_k-1}}{\Gamma(a_1)\cdots\Gamma(a_k)} e^{-(x_1+\cdots+x_n)} dx_1\cdots dx_k$$

$$= \left(\int_{\mathbb{R}_+} dy_k f(y_k) \frac{y^{a_1+\cdots+a_k-1}}{\Gamma(a_1+\cdots+a_k)}\right)\left(\int_{\mathbb{R}_+^{k-1}} g(y_1,\cdots,y_{k-1}) y_1^{a_1-1}\cdots y_{k-1}^{a_{k-1}-1}\left(1-\sum_{i=1}^{k-1} y_i\right)^{a_k-1}\right)$$

$$\times \frac{\Gamma(a_1+\cdots+a_k)}{\Gamma(a_1)\cdots\Gamma(a_k)} \mathbf{1}_{\{y_j\in[0,1],\, \sum_{j=1}^{k-1} y_j \leqslant 1\}} dy_1\cdots dy_{k-1}$$

$$= \mathbb{E}\left[f\left(\sum_{i=1}^{k} Z_a{}^i\right)\right] \mathbb{E}[g(Z'_{a_1},\cdots,Z'_{a_{k-1}})],$$

and this ends the proof. □

COROLLARY 3.5. *The joint distribution* $(|U_{1,1}|^2,\cdots,|U_{n,1}|^2)$ *is uniform on the unit simplex.*

PROOF. The proof comes from

$$|U_{i,j}|^2 \overset{\text{law}}{=} \frac{Z_1^{(j)}}{\sum_{k=1}^{n} Z_1^{(k)}},$$

and the result follows. □

## 4. Hypergeomertic functions

See *Special functions* by Andrews, Askey and Rog. Let $a \in \mathbb{C}$ for $n > 0$ we have $(a)_n = a(a+1)\cdots(a+n-1)$ and $(a)_0 = 1$. One notes that

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}.$$

DEFINITION 4.1. *We formally define a hypergeometric function as*

$$ {}_pF_q\left(\begin{array}{c} a_1,\cdots,a_p \\ b_1,\cdots,b_q \end{array}, x\right) = \sum_{n=0}^{\infty} \frac{(a_1)_n\cdots(a_p)_n}{(b_1)_n\cdots(b_q)_n}\frac{x^n}{n!},$$

*where* $a_1,\cdots,a_p,b_1,\cdots,b_q \in \mathbb{C}$ *but not in* $\{0,1,\cdots\}$.

THEOREM 4.2. *The series* ${}_pF_q$ *converges absolutely for all* $x$ *if* $p \leq q$; *for* $|x| < 1$ *if* $p = q+1$; *and it diverges for all* $x \neq 0$ *if* $p > q+1$.

PROOF. See *Special functions* by Andrews, Askey and Roy. □

As a special case of interest to us:

DEFINITION 4.3. *The hypergeometric function* ${}_2F_1(a,b,c;x)$ *is defined by the series*

$$\sum_{n=0}^{\infty} \frac{(a)_n(b)_n}{n!(c)_n}\frac{x^n}{n!},$$

*for* $|x| < 1$, *and by analytic continuation elsewhere.*

THEOREM 4.4. *The series* ${}_{q+1}F_q$ *with* $|x| < 1$ *converges absolutely if* $\mathrm{Re}(\Sigma b_i - \Sigma a_i) > 0$. *It converges conditionally, except for* $x = 1$, *if* $-1 < \mathrm{Re}(\Sigma b_i - \Sigma a_i) \leq 0$ *and diverges otherwise.*

PROOF. It follows from

$$\frac{(a_1)_n\cdots(a_{q+1})_n}{(b_1)_n\cdots(b_q)_n} \sim \frac{\prod\Gamma(b_i)}{\prod\Gamma(a_i)} n^{\Sigma a_i - \Sigma b_i - 1}.$$

The estimate follows from

$$\frac{\Gamma(x)}{\Gamma(y)} \lim_{n\to\infty} \frac{(x)_n}{(y)_n} n^{y-x} = 1.$$

□

We remark that one may also use Stirling's formula for the proof

$$\Gamma(x) \sim \sqrt{2\pi} x^{x-1/2} e^{-x}$$

where $\mathrm{Re}(x) \to \infty$.

THEOREM 4.5. *If* $\mathrm{Re}(c-a-b) < 0$ *then*

$$\lim_{x \to 1^-} \frac{{}_2F_1(a,b,c;x)}{(1-x)^{c-a-b}} = \frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)},$$

*for* $c = a+b$

$$\lim_{x \to 1^-} \frac{{}_2F_1(a,b,c;x)}{\log(\frac{1}{1-x})} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}.$$

THEOREM 4.6. *Let $S_n$ denote the nth partial sum of ${}_2F_1(a,b,c,1)$. For* $\mathrm{Re}(c-a-b) < 0$

$$S_n \sim \frac{\Gamma(c) n^{a+b-c}}{\Gamma(a)\Gamma(b)\Gamma(a+b-c)},$$

*and for the case* $c = a+b$

$$S_n \sim \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \log n.$$

THEOREM 4.7 (Gauss). *For* $\mathrm{Re}(c-a-b) > 0$ *we have*

$$\sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{n!(c)_n} = {}_2F_1(a,b,c,1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$$

Let us give a few examples which show that many elementary functions have representations as hypergeometric functions.

- $\log(1+x) = x {}_2F_1(1,1,2,-x)$ for $|x| < 1$
- $(1-x)^{-a} = {}_1F_0(a,x)$, for $|x| < 1$
- $\arctan(x) = x {}_2F_1(\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, x^2)$,
- $\sin(x) = x {}_0F_1(\frac{3}{2}, -\frac{x^2}{4})$
- $\cos(x) = x {}_0F_1(\frac{1}{2}, -\frac{x^2}{4})$
- $e^x = {}_0F_0(x)$.

In many situations one can analytically continue these functions.

## 5. Central limit theorem and Berry-Esseen bounds

For the CLT, we refer to Feller *An introduction to probability theory and its applications*, volume 2.

THEOREM 5.1. *Let $X_1, X_2, \cdots$ be independent random variables such that $\mathbb{E}[X_k] = 0$, $\mathrm{var}(X_k) = \sigma_k^2$ and put $s_n^2 = \sigma_1^2 + \cdots \sigma_n^2$. Assume that for each $t > 0$ we have*

$$\frac{1}{s_n^2} \sum_{ki=1}^{n} \mathbb{E}[|X_k|^2 \mathbf{1}_{|X_k| > tS_n}] \to 0$$

*as $n \to \infty$. Then*

$$\frac{X_1 + \cdots + X_n}{n} \overset{\text{law}}{\to} \mathcal{N}(0,1)$$

*as $n \to \infty$.*

It is often easier to check Lyapunov condition for some $\delta > 0$

$$\lim_{n \to \infty} \frac{1}{s_n^{2+\delta}} \sum_{k=1}^{n} E[|X_k|^{2+\delta}] = 0.$$

Check

$$\frac{1}{s_n^2} \sum_{k=1}^{n} E[|X_k|^2 1_{|X_k| > tS_n}] \leqslant \frac{1}{t^\delta s_n^{2+\delta}} \sum_{k=1}^{n} E[|X_k|^{2+\delta}].$$

Now we present a multidimensional version of Lindberg's CLT. We refer to *Probability Theory, an analytic view* by D. Stroock page 85. We assume that $(X_n)$ is a sequence of independent random vectors with values in $\mathbb{R}^l$ such that

$$\mathbb{E}[|X_n|^2] < \infty, \quad X_n = \begin{pmatrix} X_n^1 \\ \vdots \\ X_n^l \end{pmatrix}, \quad |X_n|^2 = |X_n^1|^2 + \cdots + |X_n^l|^2.$$

Further, we assume that

$$E[X_n] = \begin{pmatrix} E[X_n^1] \\ \vdots \\ E[X_n^l] \end{pmatrix} = 0,$$

and the covariance matrix is stricly positive

$$\operatorname{cov}(X_n) = (E[X_n^i X_n^j])_{1 \leqslant i,j \leqslant l}$$

that is,

$$\forall U = \begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix}, \quad U^t(\operatorname{cov} X_n)U > 0 \quad U \neq 0.$$

We set $S_n = \sum_{k=1}^n X_k$ and $C_n := \operatorname{cov}(S_n) = \sum_{k=1}^n \operatorname{cov}(X_n)$.

$$\Sigma_n = (\det C_n)^{1/(2\ell)}, \quad S_n^1 = \frac{S_n}{\Sigma_n}.$$

THEOREM 5.2. *Assume that $A := \lim_{n\to\infty} C_n/\Sigma_n^2$ exists, and assume that $\forall \epsilon > 0$:*

$$\lim_{n\to\infty} \frac{1}{\Sigma_n^2} \sum_{k=1}^n \mathbb{E}[|X_k|^2 \mathbf{1}_{|X_k| \geq \varepsilon \Sigma_n}] = 0.$$

*Then the vector $S_n$ converges in law to a Gaussian vector with mean zero and covariance matrix $A$.*

Question: what is the speed or rate of convergence of CLT $(d = 1)$?

THEOREM 5.3. *Let $X_1, \cdots, X_n$ be independent such that $\mathbb{E}[X)j] = 0$, $\mathbb{E}|X_j|^2 < \infty$. Put $\sigma_j^2 = \mathbb{E}[X_j^2]$ and*

$$s_n^2 = \sum_{j=1}^n \sigma_j^2, \quad F_n(x) = \mathbb{P}\left(s_n^{-1} \sum_{j=1}^n X_j \leqslant x\right).$$

*Define*

$$L_n = \frac{1}{s_n^3} \sum_{k=1}^n E[|X_j|^3].$$

*Then there exist two constants $A$ and $C$, not depending on $n$, such that the following uniform and non-uniform estimates holds*

$$\sup_x \left| F_n(x) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt \right| \leqslant A L_n,$$

*and*

$$\left| F_n(x) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt \right| \leqslant \frac{C L_n}{(1 + |x|)^3}.$$

These are called the Berry-Essen inequalities. As a reference, check Petrov's *Limit theorems of Probability*.

## 6. Rank one matrices

Let $(e_1, \cdots, e_n)$ be the canonical basis of $K^n$, where $K = \mathbb{R}$ or $\mathbb{C}$. Let $A$ be a matrix such that $\text{rank}(A) = 1$. When can it be diagonalized?
$\text{rank}(A) = 1$: all columns $c_1, \cdots, c_n$ of $A$ are linearly dependent: thus $A$ can be written as $A = (v_1 U, \cdots, v_n U)$, where $U$ is a column vector, $U \neq 0$, and $V = (v_1, \cdots, v_n)^T \in K^n$, $U = (u_1, \cdots, u_n)^T$. Then we see $A = UV^T$, and if we note $A = (a_{ij})$ where $a_{ij} = u_i v_j$

$$A^2 = UV^T UV^T = \underbrace{\sum_{i=1}^n u_i v_i}_{\text{tr}(A)} UV^T = \text{tr}(A)A.$$

Let $P(X) = X^2 - \text{tr}(A)X$, $P(A) = 0$.

(1) $\text{tr}(A) \neq 0$. In this case, $P$ is a polynomial of degree 2 with simple roots, hence it is the minimal polynomial of $A$, $P(X) = X(X - \text{tr}(A))$ and since $A \neq \text{tr}(A)\mathbb{1}_n$, $P$ is indeed the minimal polynomial. Thus $A$ can be diagonalized.
The eigenspace associated with the eigenvalue 0 is $H = \ker(A)$

$$AX = 0 \Leftrightarrow UV^T X = 0 \Leftrightarrow \left(\sum_{i=1}^n v_i x_i\right) U = 0,$$

hence

$$0 = \sum_{i=1}^n v_i x_i$$

is the equation of the hyperplane $H$. The eigenspace associated with $\text{Tr}(A)$ is spanned by $U$:

$$AU = UV^T U = \text{tr}(A)U.$$

Let $(f_2, \cdots, f_n)$ be a basis of $H$, then in the basis $(U, f_2, \cdots, f_n)$ the matrix $A$ has the form

$$\begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix},$$

with $a = \text{Tr}(A)$.

(2) $\text{Tr}(A) = 0$. In this case $A^2 = 0$, 0 is the only eigenvalue of $A$ and since $A \neq 0$, we have that $A$ cannot be diagonalized. $\ker(A) = H$ is a hyperplane but $\text{Tr}(A) = 0$ implies that $U \in \ker(A)$

$$AV = UV^T T = \left(\sum_{i=1}^n v_i^2\right) U$$

from which we consider the cases

- $K = \mathbb{R}$ then $AV = \|v\|^2 U$, $v \neq 0$ implies $\|v\| \neq 0$ which then gives us $v \notin \ker(A)$. We choose a basis $(U, f_3', \cdots, f_n', V)$. In this basis $A$ has the form

$$\begin{pmatrix} 0 & 0 & \cdots & a \\ 0 & 0 & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}$$

where $a = \|v\|^2$.

- $K = \mathbb{C}$: here $\sum v_i^2$ can be zero. Let us keep $(U, f_3', \cdots, f_n')$ as a basis of $\ker(A)$. We know that $Ae_j = c_j = v_j U$ for all $j$. Since $A \neq 0$, then there exists $j_0 \in \{1, \cdots, n\}$ such that $Ae_{j_0} \neq 0$. Now consider the basis $(U, f_3', \cdots, f_n', e_{j_0})$. In this basis the matrix $A$ has the form

$$\begin{pmatrix} 0 & 0 & \cdots & a \\ 0 & 0 & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}$$

where $a = v_{j_0}$.

EXAMPLE 6.1. *Let $e, m$ with $\|e\| = \|m\| = 1$ be two vectors of $\mathbb{C}^n$ such that $e \neq m$. There exists a unique reflection $r$ such that $r(e) = m$. Next,*

$$\operatorname{rank}(\mathrm{id} - r) = 1$$

*and one can easily check that the matrix of $(\mathrm{id} - r)$ is given by*

$$\frac{1}{1-\bar{\gamma}}(m-e)(\overline{m-e})^T, \quad \gamma = \langle m, e \rangle .$$

LEMMA 6.2. *Let $A$ be a matrix of rank one. Then $\det(\mathbf{1} + A) = 1 + \operatorname{Tr}(A)$.*

PROOF. Use the simpler forms of $A$ above. □

## 7. A priori estimates for unitary matrices

Let us fix $\epsilon > 0$. The goal of this section is the proof that the event $E := E_0 \cap E_1 \cap E_2 \cap E_3$ holds almost surely under the Haar measure on the space of virtual isometries, for

$$\begin{aligned}
E_0 &= \{\theta_0^{(1)} \neq 0\} \cap \{\forall n \geq 1, \nu_n \neq 0\} \cap \{\forall n \geq 1, 1 \leq k \leq n, \mu_k^{(n)} \neq 0\} \\
E_1 &= \{\exists n_0 \geq 1, \forall n \geq n_0, |\nu_n| \leq n^{-\frac{1}{2}+\epsilon}\} \\
E_2 &= \{\exists n_0 \geq 1, \forall n \geq n_0, 1 \leq k \leq n, |\mu_k^{(n)}| \leq n^{-\frac{1}{2}+\epsilon}\} \\
E_3 &= \{\exists n_0 \geq 1, \forall n \geq n_0, k \geq 1, n^{-\frac{5}{3}-\epsilon} \leq \theta_{k+1}^{(n)} - \theta_k^{(n)} \leq n^{-1+\epsilon}\}.
\end{aligned}$$

REMARK 7.1. *In* [**4**] *an analogous event was defined to avoid small values of the dimension $n$ where the behavior of the eigenvalues can be more erratic. This event will serve the same purpose, although we have chosen the exponents more carefully to sharpen our results.*

We begin by showing that for any fixed basis of $\mathbb{C}^n$, the coefficients of a uniform random vector on the unit sphere are almost surely $O(n^{-\frac{1}{2}+\epsilon})$ for any $\epsilon > 0$.

LEMMA 7.2. *Suppose $v_1, ..., v_n \in \mathbb{C}^n$ is an orthonormal basis and $x \in \mathbb{C}^n$, $\|x\| = 1$ is chosen uniformly from the unit sphere. Then if we write $x = x_1 v_1 + \cdots + x_n v_n$, we have bound*

$$\mathbb{P}(|x_j|^2 > \delta) = O(\exp(-\delta n/6)),$$

*for all $\delta > 0$ and $j = 1, ..., n$.*

REMARK 7.3. *We will prove this statement for deterministic vectors $v_1, \dots, v_n \in \mathbb{C}^n$. However, by conditioning, one deduces that the result remains true if the vectors $v_1, \dots, v_n$ are random, as soon as they are independent of $x$.*

PROOF. Let us assume $\delta < 1$, since the probability is zero otherwise. Let $\phi_1, ..., \phi_n$ be iid uniform random phases in $S^1$ and let $e_1, ..., e_n$ be independent standard exponential random variable. Then, it is well-known that $x$ has the same distribution as the random vector $y = y_1 v_1 + \cdots + y_n v_n$ where

$$y_j = \phi_j \sqrt{\frac{e_j}{e_1 + \cdots + e_n}}.$$

Now,

$$\mathbb{P}(e_1 + \cdots + e_n < \frac{n}{2}) \leq e^{n/2} \mathbb{E}[\exp(-e_1 - \cdots - e_n)] \leq e^{n/2} 2^{-n} \leq e^{-n/6},$$

so

$$\begin{aligned}
\mathbb{P}(|x_j|^2 > \delta) &\leq \mathbb{P}(e_j > \frac{\delta n}{2}) + O(\exp(-n/6)) \\
&= O(\exp(-\delta n/6)),
\end{aligned}$$

since $\delta \in (0, 1)$. □

From this estimate, we deduce the following bound on the coordinates of the eigenvectors $f_k^{(n)}$.

LEMMA 7.4. *Let $\epsilon > 0$. Then, almost surely, we have*

$$\sup_{1 \leq j, \ell \leq n} |\langle f_j^{(n)}, e_\ell \rangle|^2 = O(n^{-1+\epsilon})$$

*and*

$$\sup_{1 \leq j, \ell \leq n} \mathbb{E}\left[ |\langle f_j^{(n)}, e_\ell \rangle|^2 | \mathcal{A} \right] = O(n^{-1+\epsilon}),$$

*where the implied constant may depend on $\epsilon$ and $(u_m)_{m \geq 1}$.*

PROOF. Consider the vector $f_j^{(n)}$ for each fixed $j$ and $n$. By the invariance by conjugation of the Haar measure on $U(n)$, this eigenvector is, up to multiplication by a complex of modulus 1, a uniform vector on the unit sphere of $\mathbb{C}^n$. More precisely, if $\xi \in \mathbb{C}$ is uniform on the unit circle, and independent of $f_j^{(n)}$, then $\xi f_j^{(n)}$ is uniform on the unit sphere. One deduces that for all $n, j, \ell$,

$$\mathbb{P}(|\langle f_j^{(n)}, e_\ell \rangle|^2 > n^{-1+\epsilon}) = O(\exp(-n^\epsilon/6)).$$

Using the Borel-Cantelli lemma gives the first result. Moreover,

$$\begin{aligned}
\mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\epsilon}] &= \int_0^\infty \mathbb{P}[|\langle f_j^{(n)}, e_\ell \rangle|^2 \geq \delta^{\epsilon/4}] d\delta \\
&\lesssim \int_0^\infty e^{-n\delta^{\epsilon/4}/6} d\delta \\
&= \int_0^\infty e^{-z^{\epsilon/4}/6} d(z/n^{4/\epsilon}) = O(n^{-4/\epsilon}).
\end{aligned}$$

We deduce:

$$\begin{aligned}
\mathbb{P}\left( \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\epsilon} | \mathcal{A}] \geq n^{4-\frac{4}{\epsilon}} \right) &\leq n^{\frac{4}{\epsilon}-4} \mathbb{E}\left[ \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\epsilon} | \mathcal{A}] \right] \\
&= n^{\frac{4}{\epsilon}-4} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\epsilon}] = O(n^{-4}).
\end{aligned}$$

By the Borel-Cantelli lemma, for all but finitely many $n \geq 1$, $1 \leq j, \ell \leq n$,

$$\mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\epsilon} | \mathcal{A}] \leq n^{4-\frac{4}{\epsilon}}.$$

By the Hölder inequality applied to the conditional expectation, for $\epsilon$ sufficiently small,

$$\mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^2 | \mathcal{A}] \leq \left( \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\epsilon} | \mathcal{A}] \right)^{\epsilon/4} \leq n^{-1+\epsilon}.$$ □

Another consequence of Lemma 7.2 is the following:

PROPOSITION 7.5. *The events $E_0$, $E_1$, $E_2$ all hold almost surely.*

PROOF. We apply Lemma 7.2 to the decomposition

$$x_{n+1} = \sum_{j=1}^n \mu_j^{(n)} f_j^{(n)} + \nu_n e_{n+1},$$

which gives

$$\mathbb{P}(|\mu_k^{(n)}|^2 > n^{-1+\epsilon}) = O(\exp(-n^\epsilon/6))$$

so, in particular,

$$\sum_{n \geq 1} \sum_{1 \leq k \leq n} \mathbb{P}(|\mu_k^{(n)}|^2 > n^{-1+\epsilon}) = O(1).$$

Therefore, by the Borel-Cantelli lemma, almost surely only a finite number of the events $\{|\mu_k^{(n)}|^2 > n^{-1+\epsilon}\}$ hold simultaneously. A similar argument controls the coefficients $\nu_n$. □

Before we can control $E_3$, we require some estimates on the eigenvalues of a Haar unitary random matrix. Recall that if $u_n$ is distributed according to the Haar measure, then one can define, for $1 \leq p \leq n$, the $p$-point correlation function $\rho_p^{(n)}$ of the eigenangles, as follows: for any bounded, measurable function $\phi$ from $\mathbb{R}^p$ to $\mathbb{R}$,

$$\mathbb{E}\left[\sum_{1 \leq j_1 \neq \cdots \neq j_p \leq n} \phi(\theta_{j_1}^{(n)}, \ldots, \theta_{j_p}^{(n)})\right]$$
$$= \int_{[0,2\pi)^p} \rho_p^{(n)}(t_1, \ldots, t_p)\phi(t_1, \ldots, t_p)dt_1 \ldots dt_p.$$

Moreover, if the kernel $K$ is defined by

$$K(t) := \frac{\sin(nt/2)}{2\pi \sin(t/2)}$$

then the $p$-point correlation function can be given by

$$\rho_p^{(n)}(t_1, ..., t_n) = \det\left(K(t_j - t_k)\right)_{j,k=1}^{p}.$$

Let us first show that the gaps between eigenvalues cannot be asymptotically much larger than average.

LEMMA 7.6. *Let $I \subseteq [0, 2\pi)$ be Lebesgue measurable. Then*

$$\mathbb{P}(\textit{all of the eigenvalues of } u_n \textit{ are in } I) \leq \exp(-\frac{|I^c|}{2\pi}n).$$

PROOF. We recall the Andreiev-Heine identity **[26]**, which says that

$$\mathbb{P}(\text{all of the eigenvalues of } u_n \text{ are in } I) = \det M^I$$

where $M^I$ is an $n \times n$ matrix with entries

$$M^I_{j,k} = \int_I \exp(i(j-k)t)\frac{dt}{2\pi}$$

for $j, k$ between 1 and $n$. Note that the matrix $\left(\exp(i(j-k)t)\right)_{j,k=1}^n$ is hermitian and positive, $M^I$ is also; likewise $M^{I^c}$. Moreover, by computing the entries of $M^I + M^{I^c}$, one checks that this sum is the identity matrix: hence, $M^I$, $M^{I^c}$ have the same eigenvectors and, if we denote by $(\tau_j)_{1 \leq j \leq n}$ the eigenvalues of $M^{I^c}$, then $(1 - \tau_j)_{1 \leq j \leq n}$ are the eigenvalues of $M^I$. The eigenvalues of each matrix must lie in the interval $[0, 1]$, as otherwise one of the eigenvalues of the other matrix would be negative. Now,

$$\det M^I = \prod_{j=1}^n (1 - \tau_j) \leq \exp(-\sum_{j=1}^n \tau_j) = \exp(-\operatorname{Tr} M^{I^c}) = \exp(-\frac{I^c}{2\pi}n)$$

as was to be shown. □

Note that the previous lemma applies to all measurable subsets, although we will only need to apply it to intervals.

Next we control the gaps between eigenvalues from below.

LEMMA 7.7. *Suppose $t_1, ..., t_p \in I$ lie in an interval of length $|I| = \delta \leq 1/n$. Then we have the estimate*

$$\rho_p^{(n)}(t_1, ..., t_p) = O_p(\delta^{2p-2}n^{3p-2}).$$

PROOF. We have

$$\rho_p^{(n)}(t_1, \ldots, t_p) = \det\left(K(t_i - t_j)\right)_{i,j=1}^p.$$

The Taylor series for the sine function shows that for $|t| \leq 1/n$,

$$K(t) = \frac{n}{2\pi}\left(1 - \frac{1}{24}(n^2 - 1)t^2 + O(n^4t^4)\right).$$

Thus, we have:

$$\rho_p^{(n)}(t_1, \ldots, t_p) = \frac{n^p}{(2\pi)^p} \det\left(1 - \tfrac{1}{24}(n^2 - 1)(t_i - t_j)^2 + O(n^4(t_i - t_j)^4)\right)_{i,j=1}^p$$

Let $A$ denote the $p \times p$ matrix in the last display, let $1$ denote the column vector of all ones and let $w_j$ denote the column vector whose $i$th entry is

$$(w_j)_i = 1 - A_{ij} = \frac{1}{24}(n^2-1)(t_i - t_j)^2 + O(n^4(t_i-t_j)^4).$$

Then by multilinearity and the inclusion-exclusion principle,

$$\det A = \sum_{\sigma \subset [p]} (-1)^{|\sigma|} \det \begin{pmatrix} v_1 & \cdots & v_p \end{pmatrix}, \qquad \text{where } v_j = \begin{cases} w_j, & j \in \sigma \\ 1, & \text{otherwise} \end{cases}$$

Clearly each term is zero if more than one of the columns is equal to 1, so we get

$$\det A = (-1)^{p-1} \sum_{j=1}^{p} \det M_j + (-1)^p \det M$$

where $M$ is the matrix with columns $w_1, ..., w_p$ and $M_j$ is $M$ with the $j$th column replaced with 1. Then in the expansion of each determinant we can bound each term by $O_p((n^2\delta^2)^{p-1})$, and the conclusion follows. □

PROPOSITION 7.8. *The event $E_3$ holds almost surely.*

PROOF. Fix $n \geq 1$. The probability that two adjacent eigenvalues of $u_n$ differ by at least $2\delta$ is bounded above by the probability that one of the parts of the partition

$$(0,\delta) \cup (\delta, 2\delta) \cup \cdots \cup (\lfloor 2\pi\delta^{-1} \rfloor \delta, \lfloor 2\pi\delta^{-1} + 1 \rfloor \delta)$$

contains no eigenvalue. This, by Lemma 7.6, is bounded by

$$\lfloor 2\pi\delta^{-1} + 1 \rfloor \exp(-\delta n / 2\pi).$$

Now we let $\delta = n^{-1+\epsilon}$ and apply the Borel-Cantelli lemma to show that at most a finite number of the $u_n$ have gaps larger than $n^{-1+\epsilon}$.

Next, we see by Lemma 7.7 that the probability that two adjacent eigenvalues of $u_n$ differ by at most $\delta \leq 1/n$ is bounded by

$$\iint_{|t_1-t_2|<\delta} \rho_2^{(n)}(t_1,t_2)\, dt_1\, dt_2 = O(n^4\delta^3)$$

which, when we specialize $\delta = n^{-\frac{5}{3}-\epsilon}$, is $O(n^{-1-\epsilon})$; summing over $n$ and applying the Borel-Cantelli lemma shows that these events occur at must a finite number of times as well. □

# Bibliography

[1] R. Arratia, A.D. Barbour and S. Tavaré: *Logarithmic combinatorial structures: a probabilistic approach*, E.M.S Monographs in Mathematics. European Mathematical Society, 2003.

[2] A. Barbour, E. Kowalski and A. Nikeghbali. Mod discrete expansions. *Probability Theory and Related Fields*, 158, no. 3-4, 859–893 (2014).

[3] P. Bourgade, C.-P. Hughes, A. Nikeghbali, M. Yor, The characteristic polynomial of a random unitary matrix: a probabilistic approach. Duke Math. J., 145 (2008), no. 1, 45-69.

[4] P. Bourgade, J. Najnudel, and A. Nikeghbali. A unitary extension of virtual permutations. *International Mathematics Research Notices*, 2012.

[5] P. Bourgade, A. Nikeghbali and A. Rouault Circular Jacobi ensembles and deformed Verblunsky coefficients. *International Mathematics Research Notices*, 23, 4357–4394, 2009.

[6] R. Chhaibi, F. Delbaen, P-L. Méliot and A. Nikeghbali *Mod-phi convergence: approximation of discrete measures and harmonic analysis on the torus*, preprint, 2017.

[7] R. Chhaibi, J. Najnudel, and A. Nikeghbali. The Circular Unitary Ensemble and the Riemann Zeta Function: the microscopic landscape and a new approach to ratios. *Inventiones Mathematicae*, 207, no.1, 23–117, 2017.

[8] F. Delbaen, E. Kowalski and A. Nikeghbali. Mod-phi convergence. *International Mathematics Research Notices*, no.11, 3445–3485 (2015).

[9] P. Diaconis and M. Shahshahani. The subgroup algorithm for generating random variables. *Prob. Eng. Inf. Sc.*, 1:15–32, 1987.

[10] P. Diaconis and M. Shahshahani. On the eigenvalues of random matrices. *J. Appl. Probab.*, 31A:49–62, 1994. Studies in applied probability.

[11] T. Ehrhardt and B. Silbermann. Toeplitz determinants with one Fisher-Hartwig singularity. *Journal of FUnctional Analysis*, 148, 229–256, 1997.

[12] V. Féray, P-L. Méliot and A. Nikeghbali: *Mod-$\phi$ convergence: normality zones and precise deviations*, Springer Briefs in Probability and Mathematical Statistics. Springer, 2016.

[13] N.M. Katz and P. Sarnak: *Random matrices, Frobenius eigenvalues, and monodromy*, A.M.S Colloquium Publ. 45, A.M.S, 1999.

[14] J.P. Keating and N.C. Snaith: *Random matrix theory and $\zeta(1/2+it)$*, Commun. Math. Phys. 214, (2000), 57–89.

[15] J.P. Keating and N.C. Snaith: *Random matrix theory and L-functions at $s = 1/2$*, Comm. Math. Phys. 214 (2000), 91–110.

[16] S.-V. Kerov, G.-I. Olshanski, A.-M. Vershik, Harmonic analysis on the infinite symmetric group, Comptes Rendus de l'Académie des sciences de Paris, 316 (1993), 773-778.

[17] R. Killip and M. Stoiciu Eigenvalue statistics for CMV matrices: from Poisson to clock via random matrix ensembles *Duke Math. J.*, 146 (3): 361–399, 2009.

[18] E. Kowalski Arithmetic Randonnée: an introduction to probabilistic number theory *https://people.math.ethz.ch/~kowalski/probabilistic-number-theory.pdf*.

[19] E. Kowalski and A. Nikeghbali. Mod-Poisson convergence in probability and number theory. *International Mathematics Research Notices*, 18, pp. 3549-3587, (2010).

[20] E. Kowalski and A. Nikeghbali. Mod-Gaussian convergence and the value distribution of $\zeta(1/2+it)$ and related quantities. *Journal of the London Mathematical Society*, (2), 86, no. 1, 291-319 (2012).

[21] F. Mezzadri, N.C. Snaith (editors), Recent Perspectives in Random Matrix Theory and Number Theory, London Mathematical Society Lecture Note Series 322 (CUP), (2005).

[22] F. Mezzadri How to generate random matrices from the classical compact groups. *Notices of the AMS*, 54: 592–604, 2007.

[23] J. Najnudel and A. Nikeghbali. On a flow of operators associated to virtual permutations. preprint.

[24] Y.-A. Neretin, Hua type integrals over unitary groups and over projective limits of unitary groups, Duke Math. J., 114 (2002), 239-266.

[25] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.

[26] A. Soshnikov. Determinantal random point fields. *Uspekhi Mat. Nauk*, 55(5(335)):107–160, 2000.

[27] G. Tenenbaum: *Introduction to analytic and probabilistic number theory*, volume 46, Cambridge Univ. Press, 1995.

[28] E.C. Titchmarsh: *The theory of the Riemann zeta function*, 2nd edition, Oxford Univ. Press, 1986.

[29] B. Valkó and B. Virág. Continuum limits of random matrices and the Brownian carousel. *Invent. Math.*, 177 (2009), no. 3, 463–508.