24. RANDOM MATRICES IN NUMBER THEORY - SURVEY

Mathematically speaking the Random Matrix Theory handles the eigenvalue distributions in statistical aspects. The RMT emerged from the study of quantum systems (an analogue of chaotic systems), while its first relevance to number theory appeared in the publication of "Pair correlation of zeros of the Riemann zeta function" due to H.L. Montgomery of 1972. In these notes we concentrate on the harmonic analysis on four classical groups which offer models for the distribution of zeros of familiar L-functions.

1. MATRIX PRIMER

Let $M_n(\mathbb{C})$ denote the set of $n \times n$ matrices

$$A = (a_{ij}), \quad 1 \le i,j \le n$$

with the entries $a_{ij} \in \mathbb{C}$. We add matrices entry-wise and we multiply matrices by the scalar product of rows against the columns

$$AB = (a_{ij})(b_{ij}) = (c_{ij}) \quad \text{with } c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

The trace of A is the sum of its diagonal terms

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}.$$

The multiplication is not commutative, but the trace is

$$\operatorname{tr} AB = \operatorname{tr} BA.$$

The determinant of A is defined by

$$|A| = \det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

where σ runs over the symmetric group S_n, i.e. over all permutations of the numbers $1, 2, \ldots, n$. The determinant is multiplicative

$$|AB| = |A||B|.$$

A is non-singular if $|A| \neq 0$. This implies that A is invertible under the matrix product

$$A \cdot A^{-1} = I = (\delta_{ij}) \quad \text{the identity matrix.}$$

and the determinant of the inverse A^{-1} is $|A^{-1}| = 1/|A|$.

The minor M_{kl} of $A = (a_{ij})$ is the determinant of the $(n-1)\times(n-1)$ matrix which comes out of A after removing k-th row and l-th column. For every k we have the Laplace expansion

$$|A| = \sum_{l=1}^{n} (-1)^{k+l} a_{kl} M_{kl}.$$

The adjoint matrix to A consists of the minor entries; precisely

$$\operatorname{adj} A = \left((-1)^{i+j} M_{ij} \right).$$

Hence

$$A^{-1} = \operatorname{adj} A \,/\, |A|.$$

The transpose matrix of $A = (a_{ij})$ is given by

$$A^t = (a_{ji}), \qquad 1 \le j, i \le n.$$

Obviously $\operatorname{tr} A^t = \operatorname{tr} A$ and $\det A^t = \det A$.

If V is an n-dimensional vector space over $\mathbb{C}$, then a non-singular matrix $A = (a_{ij})$ maps V onto itself as a linear transformation T. Any given basis $\{v_1, \dots, v_n\} \subset V$ goes to another basis $\{w_1, \dots, w_n\} \subset V$ with

$$w_i = \sum_{j=1}^{n} a_{ij} v_j, \qquad i = 1, \dots, n.$$

We say that A is a matrix representation of T.

A non-zero vector $v \in V$ is called an eigenvector of T if

$$Tv = \lambda v$$

for some $\lambda \in \mathbb{C}$, called the eigenvalue of v. In the matrix notation we have

$$Au = \lambda u$$

where $u = (u_1, \dots, u_n) \in \mathbb{C}^n$ are the coordinates of v. Equivalently we write

$$(A - \lambda I)u = 0$$

where 0 is the zero vector. Since $u \neq 0$ we must have

$$|A - \lambda I| = 0,$$

which means that the eigenvalue λ is one of the roots of the polynomial (the characteristic polynomial of A)

$$|A - XI| = (\lambda_1 - X) \dots (\lambda_n - X).$$

A basis of $\mathbb{C}^n$ can be formed from n linearly independent eigenvectors of A. The multiplicity of an eigenvalue λ is the dimension of the linear space of eigenvectors u satisfying $(A - \lambda I)u = 0$.

PROPERTIES OF EIGENVALUES. Let A be non-singular. We have

$$\operatorname{tr} A = \lambda_1 + \dots + \lambda_n$$

$$|A| = \lambda_1 \dots \lambda_n .$$

$$|A^{-1} - XI| = \left(\frac{1}{\lambda_1} - X\right) \dots \left(\frac{1}{\lambda_n} - X\right).$$

The eigenvalues of A^k, $k \geq 1$, are $\lambda_1^k, \dots, \lambda_n^k$, so the eigenvalues of $P(A)$ are $P(\lambda_1), \dots, P(\lambda_k)$ for any polynomial $P(X) \in \mathbb{C}[X]$.

We say that B is similar to A if

$$B = U^{-1} A U \quad \text{with } U \text{ non-singular.}$$

Obviously the eigenvalues of A and B are the same ones with the corresponding eigenfunctions u_j changed into Uu_j. Putting

$$\exp A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

we have

$$\det(\exp A) = \exp(\operatorname{tr} A).$$

NEWTON'S IDENTITY

$$\det A = (-1)^n \sum_{a_1+2a_2+\dots+na_n=n} (a_1!\dots a_n!)^{-1} \left(-\frac{\operatorname{tr} A}{1}\right)^{a_1} \dots \left(-\frac{\operatorname{tr} A^n}{n}\right)^{a_n}$$

where $a_1,\dots,a_n$ run over the non-negative integers.

VANDERMONDE DETERMINANT is the polynomial in the variables $X_1,\dots,X_n$ given by the determinant

$$\Delta(X_1,\dots,X_n) = \det\left(X_i^{j-1}\right), \qquad 1 \le i,j \le n.$$

$$= \prod_{1\le i<j\le n} (X_j - X_i).$$

Let A^* denote the complex conjugate of the transpose of A;

$$A^* = (\bar{a}_{ji}).$$

Then A is said to be normal if it commutes with A^*;

$$AA^* = A^*A.$$

and A is said to be Hermitian if

$$A^* = A, \quad \text{i.e.} \quad a_{ij} = \bar{a}_{ji}.$$

A complex matrix $A = (a_{ij})$ such that

$$A^{-1} = A^{*}$$

is called unitary. A real unitary matrix is called orthogonal, it has $\det A = \pm 1$. If A is orthogonal with $\det A = 1$, then A is said to be special orthogonal, its columns and rows are orthonormal vectors in $\mathbb{R}^n$. As a linear transformation of $\mathbb{R}^n$ an orthogonal matrix A is a rotation of the Eucleadian space, actually it acts as an isometry (it preserves the inner products of vectors).

A real matrix A which is Hermitian is just symmetric, $A^t = A$, $a_{ij} = a_{ji}$. If $A^t = -A$, then A is called skew-symmetric. In this case the diagonal of A vanishes.

A real $2n \times 2n$ matrix A is said to be symplectic if

$$A^t Z A = Z$$

for some $2n \times 2n$ non-singular skew-symmetric matrix Z. We choose

$$Z = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

which has $Z^{-1} = Z^t = Z$ and $\det Z = 1$. Obviously $(\det A)^2 = 1$, but it turns out that $\det A = 1$ always.

For example, if $n=1$, then any $Z \in SL_2(\mathbb{R})$ is symplectic. In general the product of symplectic matrices is symplectic.

If A is unitary, then every eigenvalue is on the unit circle, so its eigenvalues can be written as

$$e^{i\theta_1}, \dots, e^{i\theta_n} \quad \text{with} \quad 0 \le \theta_1 \le \dots \le \theta_n < 2\pi.$$

If A is orthogonal (= unitary real), then its complex conjugates of eigenvalues are also eigenvalues, so we can write them as

$$e^{\pm i\theta_1}, \dots, e^{\pm i\theta_N} \quad \text{if } n = 2N$$

$$1, e^{\pm\theta_1}, \dots, e^{\pm i\theta_N} \quad \text{if } n = 2N+1$$

with $0 \le \theta_1 \le \dots \le \theta_N \le \pi$.

If A is unitary symplectic, then its eigenvalues can be arranged as

$$e^{\pm i\theta_1}, \dots, e^{\pm i\theta_n} \quad \text{with} \quad 0 \le \theta_1 \le \dots \le \theta_n \le \pi.$$

We shall consider the following groups of $n \times n$ matrices

$U(n)$ – the unitary matrices
$SO(2N)$ – the orthogonal matrices with determinant 1
$SO(2N+1)$ – the orthogonal matrices with determinant 1
$USp(2N)$ – the unitary symplectic matrices.

2. HAAR MEASURE

The embedding of $GL_n(\mathbb{R})$ into $\mathbb{R}^{n^2}$ induces a locally compact Hausdorff topology. Let G be a subgroup of $GL_n(\mathbb{R})$ and μ be a measure on G defined on Borel sets (the sets generated by the open subsets of σ-algebra). Of course, if $S \subset G$ is a Borel set then

$$gS = \{ gs \; ; \; s \in S \}$$

is a Borel set for every $g \in G$. We say that the measure μ is left-invariant if

$$\mu(gS) = \mu(S)$$

for every Borel set S and every $g \in G$. Similarly a measure μ is called right-invariant if $\mu(Sg) = \mu(S)$.

THEOREM (Definition of Haar measure). There exists unique (up to a positive constant factor) countable, additive, non-trivial measure μ on the Borel sets $S \subset G$ having the following properties;

- μ is left-invariant
- $\mu(K) < \infty$ for every $K \subset G$, K compact
- $\mu(S) = \inf\{ \mu(U) \; ; \; S \subset U \subset G, \; U \text{ open} \}$
- $\mu(U) = \sup\{ \mu(K) \; ; \; K \subseteq U, \; K \text{ compact} \}$.

The existence of Haar measure allows to do integration on the locally compact groups. Hence one can perform probability theory, ergodic theory and harmonic analysis on G. Constructing μ is not easy, but its properties are well exposed by the mean-values of functions (Lebesgue integrals)

$$\int_G f(gx)\, d\mu x = \int f(x)\, d\mu(x).$$

Similarly we obtain the right-invariant measure on G. By the uniquess (up to a positive scalar) it follows that the right-invariant measure is a multiple of the left-invariant measure by a factor which is a continuous homomorphism δ of G onto the multiplicative group of positive real numbers $\mathbb{R}^*$. Hence

$$\int_G f(xg)\, d\mu(x) = \int_G f(x)\, \delta(x)\, d\mu(x).$$

If $\delta(x) \equiv 1$ (both measures are equal), then the group G is called unimodular. Abelian groups are unimodular. Compact groups are unimodular, because a continuous homomorphism of the multiplicative group $\mathbb{R}^*$ of positive real numbers must be the identity. All groups $U(n)$ $SO(n)$, $USp(2N)$ are compact, so unimodular. Since any non-empty open set has positive measure we can normalize μ on a compact group G by $\mu(G) = 1$.

EXAMPLE. The whole group

$$G = SL_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} ;\ ad - bc = 1 \right\}$$

is unimodular in spite of being non-abelian and not compact. To describe the Haar measure of G we use the Iwasawa decomposition

$$G = NAK$$

where

$$N = \left\{ \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} ;\ x \in \mathbb{R} \right\},$$

$$A = \left\{ \begin{pmatrix} \sqrt{y} & \\ & 1/\sqrt{y} \end{pmatrix} ;\ y \in \mathbb{R}^* \right\},$$

$$K = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} ;\ 0 \leq \theta < 2\pi \right\}$$

Note that N, A are abelian, but the group $P = NA$ of upper triangular matrices is not. Specifically we have the relation

$$a(y)\,n(x) = n(xy)\,a(y) \quad \text{if } x \in \mathbb{R},\ y \in \mathbb{R}^*$$

The invariant measures on N, A, K are

$$dn(x) = dx$$
$$da(y) = y^{-1}\,dy$$
$$dk(\theta) = (2\pi)^{-1}\,d\theta$$

where dx, dy, $d\theta$ are the Lebesgue measures. The left-invariant

measure on P is

$$dp = y^{-1}dxdy \qquad \text{if } p = an$$

and the modular function on P is $\delta(p) = y^{-1}$, so

$$\delta(p)dp = y^{-2}dxdy$$

is the right-invariant measure on P.

3. WEYL'S INTEGRATION FORMULAS

From now on G is one of the four classical compact groups. The Haar measure on G can be expressed conveniently using the Lebesgue measure on specific subgroups of G. In the sequel X denotes the elements of G and dX the Haar measure on G. The goal is to evaluate the integrals

$$\int_G f(X)\,dX$$

for special test functions $f(X)$.

Let $G = U(n)$ be the group of $n \times n$ unitary matrices. Using unitary diagonalization of X, say

$$X = U\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^t,$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of X and U is unitary we can exploit that the measure is unitary invariant. The collection of all matrices X with the same set of n-points $\lambda_1, \ldots, \lambda_n$ on the unit circle $|\lambda_j| = 1$ as the eigenvalues constitutes one conjugacy class of $G = U(n)$. Since we are interested mostly in the statistic of eigenvalues we restrict our integrals for test functions $f(X)$ which are constant on conjugacy classes. Therefore we have

$$f(X) = f(\theta_1, \dots, \theta_n)$$

where $f(\theta_1, \dots, \theta_n)$ is a symmetric function of the angles θ_j of $\lambda_j = e^{i\theta_j}$, $0 \le \theta_j < 2\pi$, $1 \le j \le n$. Conversely, any symmetric function $f(\theta_1, \dots, \theta_n)$ with $0 \le \theta_j < 2\pi$ gives rise to $f(X)$ on $U(n)$ which is constant on conjugacy classes. We call such $f(X)$ a 'class function' on $G = U(n)$.

THEOREM 1 (H. Weyl). If $f(X)$ is a class function, then

$$(24.1) \quad \int_{U(n)} f(X)\,dX = \frac{1}{(2\pi)^n n!} \int_0^{2\pi} \cdots \int_0^{2\pi} f(\theta_1, \dots, \theta_n) \prod_{1 \le j < k \le n} |e^{i\theta_k} - e^{i\theta_j}|^2 \, d\theta_1 \cdots d\theta_n .$$

Next, let $G = SO(2N)$ be the group of orthogonal $2N \times 2N$ matrices (real unitary) of determinant 1. If $X \in G$, then its eigenvalues can be arranged as

$$e^{\pm i\theta_1}, \dots, e^{\pm i\theta_N} \quad \text{with } 0 \le \theta_1 \le \dots \le \theta_N < \pi$$

and any function $f(X)$ on G which is constant on conjugacy classes of G can be written as

$$f(X) = f(\theta_1, \dots, \theta_N)$$

with $f(\theta_1, \dots, \theta_N)$ symmetric. We call such $f(X)$ a class function on $G = SO(2N)$.

THEOREM 2 (H. Weyl). If $f(X)$ is a class function on $SO(2N)$, then

$$(24.2)\quad \int_{SO(2N)} f(X)\,dX = \frac{2^{(N-1)^2}}{\pi^N N!} \int_0^\pi \cdots \int_0^\pi f(\theta_1,\dots,\theta_N) \prod_{1\le j<k\le N} (\cos\theta_k - \cos\theta_j)^2\, d\theta_1 \cdots d\theta_N$$

where dX is the Haar measure.

Note that for $N=1$ we get $\pi^{-1}\int_0^\pi f(\theta)\,d\theta$ which agrees with the measure on the maximal compact subgroup K of $SL_2(\mathbb{R})$.

Next, let $G = SO(2N+1)$ be the group of orthogonal $(2N+1)\times(2N+1)$ matrices (real unitary) of determinant 1. If $X \in G$, then its eigenvalues can be arranged as

$$1,\ e^{\pm i\theta_1},\dots,e^{\pm i\theta_N} \qquad \text{with } 0 \le \theta_1 \le \dots \le \theta_N < \pi$$

and any function $f(X)$ on G which is constant on conjugacy classes of G can be written as

$$f(X) = f(\theta_1,\dots,\theta_N)$$

with $f(\theta_1,\dots,\theta_N)$ symmetric. We call such $f(X)$ a class function on $G = SO(2N+1)$.

THEOREM 3 (H. Weyl). If $f(X)$ is a class function on $SO(2N+1)$, then

$$(24.3)\quad \int_{SO(2N+1)} f(X)\,dX = \frac{2^{N^2}}{\pi^N N!}\int_0^{\pi}\cdots\int_0^{\pi} f(\theta_1,\dots,\theta_N)\prod_{1\le j<k\le N}(\cos\theta_k-\cos\theta_j)^2$$

$$\left(\sin\frac{\theta_1}{2}\cdots\sin\frac{\theta_N}{2}\right)^2 d\theta_1\cdots d\theta_N$$

where dX is the Haar measure.

Finally, let $G = USp(2N)$ be the group of unitary symplectic $2N\times 2N$ matrices (real unitary, so of discriminant 1). If $X\in G$, then its eigenvalues can be arranged as

$$e^{\pm i\theta_1},\dots,e^{\pm i\theta_N} \qquad \text{with } 0\le\theta_1\le\dots\le\theta_N<\pi.$$

and $f(X)=f(\theta_1,\dots,\theta_N)$ symmetric, if $f(X)$ is a class function.

THEOREM 4 (Weyl). If $f(X)$ is a class function on $USp(2N)$, then

$$(24.4)\quad \int_{USp(2N)} f(X)\,dX = \frac{2^{N^2}}{\pi^N N!}\int_0^{\pi}\cdots\int_0^{\pi} f(\theta_1,\dots,\theta_N)\prod_{1\le j<k\le N}(\cos\theta_k-\cos\theta_j)^2$$

$$(\sin\theta_1\cdots\sin\theta_N)^2\,d\theta_1\cdots d\theta_N$$

where dX is the Haar measure.

4. ALTERNATE FORMULAS FOR HAAR MEASURES

The Haar measures given on the set of eigenvalues in the integration formulas of Weyl are proportional to the Lebesgue measure by some factors which are the Vandermonde determinants

(24.5) $$\prod_{1\le j<k\le n} |e^{i\theta_k} - e^{i\theta_j}|^2 = |\Delta(e^{i\theta_1},\dots,e^{i\theta_n})|^2,$$

(24.6) $$\prod_{1\le j<k\le n} (\cos\theta_k - \cos\theta_j)^2 = \Delta(\cos\theta_1,\dots,\cos\theta_n)^2.$$

Using elementary row and column operations one can arrange these determinants in various ways which are more useful for computing the Density Functions and the Correlation Functions Put

(24.7) $$S_n(\theta) = \frac{\sin n\theta/2}{\sin\theta/2}.$$

We get

(24.8) $$\prod_{1\le j<k\le n} |e^{i\theta_k} - e^{i\theta_j}|^2 = \det S_n(\theta_k-\theta_j),$$

(24.9) $$2^{(n-1)^2}\prod_{1\le j<k\le n} (\cos\theta_k - \cos\theta_j)^2 = \det \frac{1}{2}\left(S_{2n-1}(\theta_k-\theta_j) + S_{2n-1}(\theta_k+\theta_j)\right),$$

$$(24.10)\quad 2^{n(n-1)} \prod_{1\le j<k\le n} (\cos\theta_k - \cos\theta_j)^2 \prod_{1\le j\le n} (\sin\theta_j/2)^2$$

$$= \det \frac{1}{2}\left(S_{2n}(\theta_k-\theta_j) - S_{2n}(\theta_k+\theta_j)\right),$$

$$(24.11)\quad 2^{n(n-1)} \prod_{1\le j<k\le n} (\cos\theta_k - \cos\theta_j)^2 \prod_{1\le j\le n} (\sin\theta_j)^2$$

$$= \det \frac{1}{2}\left(S_{2n+1}(\theta_k-\theta_j) - S_{2n+1}(\theta_k+\theta_j)\right).$$

REMARKS. The Haar measures in Weyl's formulas (1)-(4) are normalized so the total mass of the relevant group is always 1.

There are various generalizations of Weyl's integrals. For studying the limiting distribution of eigenvalues we need integrals over a fixed number of eigenvalues. Here is an example for the unitary group

THEOREM 5. Let $F(\theta_1,\dots,\theta_m)$ be a nice function in m variables, $1\le m\le n$. Then the integral

$$(24.12)\quad \int_{U(n)} \sum_{1\le l_1<\dots<l_m\le n} \dots \sum F(\theta_{l_1},\dots,\theta_{l_m})\, dX$$

is equal to

$$(24.13)\qquad \frac{1}{(2\pi)^m\, m!}\int_0^{2\pi}\cdots\int_0^{2\pi} F(\theta_1,\dots,\theta_m)\,\det\left(\frac{\sin n(\theta_k-\theta_j)/2}{\sin(\theta_k-\theta_j)/2}\right)_{m\times m} d\theta_1\cdots d\theta_m .$$

REMARKS. The integration (12) goes over the group $U(n)$ with the Haar measure dX and $\{\theta_{l_1},\dots,\theta_{l_m}\}$ are the subsequences of the eigenangles $\{\theta_1,\dots,\theta_n\}$ with $0\le|\theta_j|<2\pi$. The original formula (1) is the special case of (13) for $m=n$, see (8). A derivation of (13) from (1) goes via repeated applications of Gaudin's Lemma (see B. Conrey [C])

5. THE m-LEVEL DENSITY THEOREMS

Let $f(x_1, \dots, x_m)$ be a nice function in variables $x_\ell \geq 0$. For every $n \geq m$ we consider the integral

$$(24.14) \quad D_n(f) = \int_{U(n)} \sum_{1 \leq \ell_1 < \dots < \ell_m \leq n} \dots \sum f\left(\frac{n}{2\pi}\theta_{\ell_1}, \dots, \frac{n}{2\pi}\theta_{\ell_m}\right) dX.$$

Note that we blowed up θ's into $n\theta/2\pi$ to deal with points which have mean value spacing 1. By Theorem 5 we get

$$D_n(f) = \frac{1}{(2\pi)^m m!} \int_0^{2\pi} \dots \int_0^{2\pi} f\left(\frac{n}{2\pi}\theta_1, \dots, \frac{n}{2\pi}\theta_m\right) \det\left(\frac{\sin n(\theta_k - \theta_j)/2}{\sin(\theta_k - \theta_j)/2}\right)_{m \times m} d\theta$$

where $d\theta = d\theta_1 \dots d\theta_m$ is the Lebesgue measure on the box $0 \leq \theta_1, \dots, \theta_m < 2\pi$. Changing the variables $\theta_j = 2\pi x_j / n$ for $1 \leq j \leq m$ we write

$$(24.15) \quad D_n(f) = \frac{1}{m!} \int_0^n \dots \int_0^n f(x_1, \dots, x_m) \det\left(\frac{\sin \pi(x_k - x_j)}{n \sin \pi(x_k - x_j)/n}\right)_{m \times m} dx_1 \dots dx_m.$$

Now, taking the limit $n \to \infty$ we obtain

THEOREM 6. For a nice function $f(x_1, \dots, x_m)$ we have

$$(24.16) \quad D_n(f) \sim \frac{1}{m!} \int_0^\infty \dots \int_0^\infty f(x_1, \dots, x_m)\, W(x_1, \dots, x_m)\, dx_1 \dots dx_m$$

with

$$(24.17) \qquad W(x_1,\dots,x_m) = \det_{m\times m}\left(\frac{\sin\pi(x_k-x_j)}{\pi(x_k-x_j)}\right).$$

REMARKS. The asymptotic formula (16) is called m-level Density Theorem and the integral kernel (17) is called m-level Density Function for the unitary groups. Similar results hold for the other classical groups. The corresponding Density Functions are:

$$(24.18) \qquad W^{\pm}(x_1,\dots,x_m) = \det_{m\times m}\left(\frac{\sin\pi(x_k-x_j)}{\pi(x_k-x_j)} \pm \frac{\sin\pi(x_k+x_j)}{\pi(x_k+x_j)}\right)$$

if $G = SO(2N)$ and $G = USp(2N)$, respectively. If $G = SO(2N+1)$ the Density Function contains the δ-function which is the consequence of 1 being always an eigenvalue.

6. THE PAIR CORRELATION

We want to understand the statistic of differences between the eigenangles $0 \le \theta_1, \dots, \theta_n < 2\pi$ of the unitary $n \times n$ matrix X as the rank $n \to \infty$. To this end we consider the integral

$$(24.19) \qquad C_n(f) = \frac{1}{n} \int_{U(n)} \sum_{1 \le j < k \le n} f\left(\frac{n}{2\pi}(\theta_k - \theta_j)\right) dX$$

for every nice test function $f(v)$ which vanishes rapidly as $|v| \to \infty$. By Theorem 5 with $m = 2$ we derive

$$C_n(f) = \frac{1}{8\pi^2 n} \int_0^{2\pi} \int_0^{2\pi} f\left(\frac{n}{2\pi}(\theta_1 - \theta_2)\right) \det_{2\times 2}\left(\frac{\sin n(\theta_1-\theta_2)/2}{\sin(\theta_1-\theta_2)/2}\right) d\theta_1 d\theta_2 .$$

Changing the variables $\theta_1 = 2\pi \vartheta_1 / n$, $\theta_2 = 2\pi \vartheta_2 / n$ we get

$$C_n(f) = \frac{1}{2} \int_0^n \int_0^n f(\vartheta_1 - \vartheta_2) \det\left(\frac{\sin \pi(\vartheta_1 - \vartheta_2)}{n \sin \pi(\vartheta_1 - \vartheta_2)/n}\right) d\vartheta_1 d\vartheta_2 .$$

$$= \frac{1}{2} \int_0^n \int_0^n f(\vartheta_1 - \vartheta_2) \left(1 - \left(\frac{\sin \pi(\vartheta_1 - \vartheta_2)}{n \sin \pi(\vartheta_1 - \vartheta_2)/n}\right)^2\right) d\vartheta_1 d\vartheta_2$$

$$= \frac{1}{2} \int_{-n}^{n} f(v) \left(1 - \left(\frac{\sin \pi v}{n \sin \pi v/n}\right)^2\right) \left(1 - \frac{|v|}{n}\right) dv .$$

Taking the limit $n \to \infty$ we obtain

THEOREM 6. For a nice function $f(v)$ we have

$$C_n(f) \sim \frac{1}{2}\int_{-\infty}^{\infty} f(v)\left(1-\left(\frac{\sin \pi v}{\pi v}\right)^2\right) dv. \tag{24.20}$$

Note that H.L. Montgomery predicts that the same pair correlation holds for zeros of the Riemann zeta function.

Recall that for the proof of (20) we appealed to the formula (12)=(13), which follows from Weyl's formula (1) by quite lengthy manipulations with the rows and columns of the Vandermonde determinant (5). Some integration over $\theta_1, \dots, \theta_n$ is also required. Next we are going to derive (20) again by somewhat different integration. The key arguments use exact formulas for power moments of traces

THEOREM 7 (F.J. Dyson, 1962). For any integer $k \neq 0$ we have

$$\int_{U(n)} |\operatorname{Tr} X^k|^2 \, dX = \min(|k|, n). \tag{24.21}$$

Obviously, if $k=0$ we get n^2, because the measure dX is normalized. This formula was generalized for other classical compact groups and for mixed power moments. Here is one example.

THEOREM 8 (P. Diaconis + M. Shahshahani, 1994). Suppose $[a_1, \dots, a_n]$, $[b_1, \dots, b_n]$ are non-negative integers with

$$a_1 + 2a_2 + \dots + na_n \le n, \qquad b_1 + 2b_1 + \dots + nb_n \le n. \tag{24.23}$$

Then

$$\int_{U(n)} \prod_{k=1}^{n} (\operatorname{Tr} X^k)^{a_k} (\operatorname{Tr} X^{-k})^{b_k} = 0 \tag{24.23}$$

if $[a_1, \dots, a_n] \neq [b_1, \dots, b_n]$ and it is equal to

$$1^{a_1} 2^{a_2} \dots n^{a_n} \; a_1! \dots a_n! , \tag{24.24}$$

if $[a_1, \dots, a_n] = [b_1, \dots, b_n]$.

PROOF OF THEOREM 6.

The test function $f(v)$ in (19) decays rapidly to zero so we can assume it has compact support. Therefore only small differences contribute to (19); $\theta_k - \theta_j \ll n^{-1}$. We can replace $f(v)$ by

(24.25) $$f_n(v) = \sum_{k \in \mathbb{Z}} f(v + kn).$$

Note that if n is large, only one term of (25) is essential, namely k being the nearest integer to $-v/n$. The large points $|v+kn|$ contribute to (19) a negligible amount, so that

(24.26) $$C_n^+(f) \sim C_n^+(f_n), \quad \text{as } n \to \infty.$$

where $C_n^+(f)$ denotes (19) without the restriction $j<k$ (for convenience). By Poisson's formula

(24.27) $$f_n(v) = \frac{1}{n} \sum_{k \in \mathbb{Z}} \hat{f}\left(\frac{k}{n}\right) e\left(\frac{vk}{n}\right).$$

Hence

(24.28) $$C_n^+(f_n) = \frac{1}{n^2} \sum_k \hat{f}\left(\frac{k}{n}\right) \int_{U(n)} \sum_{1 \le i,j \le n} \sum e\left(\frac{k}{2\pi}(\theta_j - \theta_i)\right) dX$$

On the other hand

$$\operatorname{Tr} X^k = \lambda_1^k + \dots + \lambda_n^k = e^{ik\theta_1} + \dots + e^{ik\theta_n}$$

and

$$|\mathrm{Tr}\, X^k|^2 = \sum\sum_{1\le i,j\le n} e\left(\frac{k}{2\pi}(\theta_j-\theta_i)\right),$$

so (28) becomes

$$(24.29)\qquad C_n^+(f) = \frac{1}{n^2}\sum_k \hat{f}\left(\frac{k}{n}\right)\int_{U(n)} |\mathrm{Tr}\, X^k|^2\, dX.$$

Then, by Dyson's formula (21) we get

$$C_n^+(f) = \hat{f}(0) + \frac{1}{n^2}\sum_k \hat{f}\left(\frac{k}{n}\right)\min(|k|,n)$$

$$= \hat{f}(0) + \int_{-\infty}^{\infty} \hat{f}(t)\min(|t|,1)\,dt + o(1)$$

$$(24.30)\qquad = \int f(v)\left(1-\left(\frac{\sin \pi v}{\pi v}\right)^2\right) dv + o(1).$$

This agrees with (20) ∎

7. REMARKS ABOUT DYSON'S FORMULA

Using (1) the formula (21) becomes

$$\int_0^{2\pi}\cdots\int_0^{2\pi} |e^{ik\theta_1}+\cdots+e^{ik\theta_n}|^2 \prod_{1\le j<l\le n} |e^{i\theta_j}-e^{i\theta_l}|^2 \, d\theta_1\cdots d\theta_n$$

(24.31) $$= c_n \min(k,n), \quad \text{if } k\ge 1,$$

where $c_n = (2\pi)^n n!$. Here $|\mathrm{Tr}\, X^k|^2$ can be written as

$$|e^{ik\theta_1}+\cdots+e^{ik\theta_n}|^2 = n + \sum_{1\le a<b\le n}\sum \left(e^{ik(\theta_a-\theta_b)}+e^{-ik(\theta_a-\theta_b)}\right).$$

The constant n contributes nc_n, because the total measure is normalized to be 1. Opening the determinant we get a sum of products of at most $n-1$ of exponentials of type $e^{\pm i(\theta_j-\theta_l)}$. Then we can integrate separately with respect to $\theta_1,\dots,\theta_n$ getting a lot of zeros due to

$$\int_0^{2\pi} e^{im\theta}\, d\theta = 0, \quad \text{if } m\neq 0.$$

If $k\ge n$, then the products of exponentials from the determinant cannot neutrilize $e^{\pm ik(\theta_a-\theta_b)}$, so the integration gives zero. This proves Dyson's formula if $k\ge n$. For $1\le k<n$ the arguments are similar, but require a lot of combinatorial arrangements.