On the eigenvalues of random matrices

PERSI DIACONIS AND MEHRDAD SHAHSHAHANI

Abstract

Let M be a random matrix chosen from Haar measure on the unitary group U_n. Let $Z = X + iY$ be a standard complex normal random variable with X and Y independent, mean 0 and variance $\frac{1}{2}$ normal variables. We show that for $j = 1, 2, \ldots$, $\mathrm{Tr}(M^j)$ are independent and distributed as $\sqrt{j}Z$ asymptotically as $n \to \infty$. This result is used to study the set of eigenvalues of M. Similar results are given for the orthogonal and symplectic and symmetric groups.

HAAR MEASURE; ORTHOGONAL GROUPS; SYMPLECTIC GROUPS; SYMMETRIC GROUPS

AMS 1991 SUBJECT CLASSIFICATION 15A52

0. Introduction

The classical matching problem of elementary probability is elegantly surveyed by Takács (1980). In one incarnation it says that the number of fixed points in a random permutation has an approximate Poisson(1) distribution. In other words, the trace of a randomly chosen permutation matrix has an approximate Poisson(1) distribution.

We study a continuous generalization to the classical compact groups: orthogonal, unitary, and symplectic. Throughout, random means uniformly (Haar) distributed. Our main results show that the trace of a randomly chosen matrix has an approximate Gaussian distribution. We also derive Gaussian approximations for powers of random matrices and so results for the distribution of their eigenvalues. Figure 1 shows the eigenvalues of a single realization on U_n when $n = 100$. The points appear to be very regularly spread out. This is a consequence of the theory developed.

These continuous problems arise in applied work. Mehta (1991) is a book-length survey showing how the eigenvalue distributions appear in a variety of problems from particle physics. Mehta also reviews Odlyzko's remarkable work connecting the eigenvalues with the distribution of the zeros of the zeta function. Diaconis and Shahshahani (1986) show how the distribution of the trace of a random orthogonal matrix occurs naturally in a standard telephone encryption scheme. Larsen (1993) shows how this same distribution appears when studying the error term for the number of solutions of equations over finite fields.

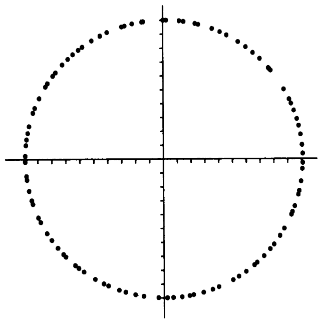

Figure 1. Eigenvalues of a random unitary matrix in U_{100}

Our motivation for studying the present problems comes from another extension of the matching problem. Any permutation can be written as a product of disjoint cycles (so

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 1 & 3 & 9 & 5 & 4 & 8 & 7 & 6 \end{pmatrix}$$

is (12) (3) (496) (78) in cycle notation). Let $a_i(\pi)$ be the number of cycles of length i in the permutation π (so $a_1 = 1$, $a_2 = 2$, $a_3 = 1$ in the example). Thus a_1 is the number of fixed points. Goncharov (1944) and Shepp and Lloyd (1966) showed that if π is uniformly chosen in the permutation group S_n, then $a_i(\pi)$ are asymptotically independent with limiting Poisson $(1/i)$ distributions. Goncharov (1944), Shepp and Lloyd (1966), Erdös and Turan (1967a, b, 1968), Vershik and Kerov (1981) and many others have studied a variety of functionals of the cycle lengths.

The cycle lengths code the conjugacy class of a permutation. Here π and σ are conjugate if $\pi = \eta^{-1}\sigma\eta$ for some $\eta \in S_n$. This is an equivalence relation partitioning S_n into conjugacy classes. Two permutations are conjugate if and only if they have the same cycle lengths. Two matrices in one of the classical groups are conjugate if and only if they have the same eigenvalues. We hoped that some of the richness and elegance of the study of cycles would carry over to eigenvalues.

Sections 1 and 2 present results for the unitary group. The calculations are easiest here. They rely on classical results from symmetric function theory that may be novel in a probability setting. Following this the orthogonal and symplectic cases are studied. The main tool is the method of moments. We show that the various traces have moments which equal the moments of the limiting measures for all large values of n. This high-order contact surprised us, and in the last section we show that it also holds for the matching problem: the first n moments of the number of fixed points of a random permutation in S_n equal the first n moments of Poisson(1).

It is worth remarking here 'why' the results presented hold. Consider the trace of a random unitary matrix. This is a sum of a lot of small, not particularly dependent, random things. The central limit theorem suggests it is approximately normally distributed. It is not hard to construct a proof along these lines. Next consider

$$\int_{U_n} (\mathrm{Tr}(m))^a (\overline{\mathrm{Tr}(m)})^b \, dm.$$

This has a group-theoretic interpretation: $(\mathrm{Tr}(m))^a$ is the character of the ath tensor power of the n-dimensional representation of U_n. The integral is the sum of the multiplicities of the common constituents of the ath and bth tensor powers. In particular it is an integer. By the first remark it converges to $\boldsymbol{E}(Z^a \bar{Z}^b)$ with Z complex normal. These last moments are integers as well. Since integers converging to integers must eventually be equal, we expect equality of moments in all the cases of this paper. It is interesting how rapidly this takes hold.

Remark. The physics literature works with a unitary, orthogonal and symplectic ensemble. While the unitary ensemble is the one considered here, the orthogonal and symplectic ensembles differ. Their orthogonal ensemble consists of the symmetric unitary matrices. This is U_n/O_n. Their symplectic ensemble consists of anti-symmetric unitary matrices. This is U_{2n}/Sp_n. We hope to carry through the distribution of the eigenvalues on these ensembles along the lines of the present paper.

1. The unitary group

A complex normal random variable Z can be represented as $Z = X + iY$ with X and Y independent real normal random variables having mean 0 and variance $\frac{1}{2}$. These variables can be used to represent Haar measure on the unitary group U_n in the following standard fashion. Form an $n \times n$ random matrix with independent identically distributed complex normal coordinates Z_{ij}. Then perform the Gram–Schmidt algorithm. This results in a random unitary matrix M which is Haar distributed on U_n. Invariance of M is easy to see from the invariance of the complex normal vectors under U_n.

This representation suggests that there is a close relationship between the unitary group and the complex normal distribution. For example, Diaconis and Mallows (1986) proved the following result.

Theorem 0. *Let M be Haar distributed on U_n. Let Z be complex normal. Then, for any open ball B,*

$$\lim_{n\to\infty} \boldsymbol{P}\{\mathrm{Tr} M \in B\} = \boldsymbol{P}\{Z \in B\}.$$

The following result generalizes Theorem 0.

Theorem 1. *Fix k in $\{1, 2, 3, \ldots\}$. For every collection of open balls B_i in the complex plane,*

$$\lim_{n\to\infty} \boldsymbol{P}\{\mathrm{Tr}(M) \in B_1, \mathrm{Tr}(M^2) \in B_2, \ldots, \mathrm{Tr}(M^k) \in B_k\} = \prod_{j=1}^{k} \boldsymbol{P}(\sqrt{j} Z \in B_j).$$

https://doi.org/10.1017/S0021900200106989 Published online by Cambridge University Press

Theorem 1 will be proved by the method of moments. To this end, note the following result.

Lemma 1. *Let Z be standard complex normal. Then, for each a, $b \in \{0,1,2,\dots\}$*

$$\boldsymbol{E}(Z^a \bar{Z}^b) = \delta_{ab} a!$$

Proof. By standard properties of normal variables $Z = R \exp(i\Theta)$, with R^2 an exponential random variable and Θ uniform on $[0, 2\pi)$ with Θ independent of R. Now

$$\boldsymbol{E}(Z^a \bar{Z}^b) = \boldsymbol{E}(R^{a+b}) \boldsymbol{E}(\exp(i(a-b)\Theta)).$$

This vanishes unless $a = b$. Then $\boldsymbol{E}(R^{2a}) = a!$

The main result of this section calculates the joint moments of $\mathrm{Tr}(M)$, $\mathrm{Tr}(M^2), \dots, \mathrm{Tr}(M^k)$. The result shows these equal the joint moments of independent complex normal variables.

Theorem 2. *Let M be Haar distributed on U_n. Let $\boldsymbol{a} = (a_1, a_2 \dots, a_k)$; $\boldsymbol{b} = (b_1, b_2, \dots, b_k)$ with a_i, $b_i \in \{0, 1, \dots\}$. Let Z_1, $Z_2, \dots, Z_k$ be independent standard complex normal variables. Then, for all $n \geq \sum_{i=1}^{k} (a_i + b_i)$,*

$$\boldsymbol{E}\left\{\prod_{j=1}^{k} (\mathrm{Tr}\, M^j)^{a_j} \cdot \overline{\mathrm{Tr}\,(M^j)}^{b_j}\right\} = \delta_{\boldsymbol{ab}} \prod_{j=1}^{k} j^{a_j} a_j! = \boldsymbol{E}\left\{\prod_{j=1}^{k} (\sqrt{j} Z_j)^{a_j} \overline{(\sqrt{j} Z_j)}^{b_j}\right\}.$$

Proof. Recall the power sum functions: $P_j(x_1, x_2, \dots, x_n) = \sum_{i=1}^{n} x_i^j$ and for $\lambda = 1^{a_1} 2^{a_2} \cdots k^{a_k}$, $P_\lambda = \prod_{j=1}^{k} P_j^{a_j}$. Here λ is a partition of $a_1 + 2a_2 + \cdots + ka_k = K$. As λ ranges over all partitions of K, the P_λ form a basis for the homogeneous symmetric polynomials in n variables for all $n \geq K$.

A second basis for these polynomials is the Schur functions s_μ. These can be defined by

$$P_\lambda = \sum_{\mu \vdash K} \chi_\lambda(\mu) s_\mu$$

with $\chi_\lambda(\mu)$ the character of the symmetric group S_K associated to the λth irreducible representation on the μth conjugacy class. This and other properties of Schur functions can be found in Macdonald (1979) or Sagan (1991).

A second crucial property of Schur functions is that they give the characters of the unitary group. If $m \in U_n$ has eigenvalues $\exp(i\theta_j)$, $1 \leq j \leq n$, define $s_\mu(m) = s_\mu(\exp(i\theta_1), \dots, \exp(i\theta_n))$. Here, s_μ is defined as zero if μ has more than n parts. As μ varies over partitions of any number with n or fewer parts the s_μ give all characters of U_n. Then, the orthogonality for characters becomes

$$\int_{U_n} s_\lambda(m) \overline{s_\mu(m)} dm = \delta_{\lambda,\mu}.$$

Now $\mathrm{Tr}(m^j) = P_j(m)$, so

$$\prod_{j=1}^{k} \mathrm{Tr}(m^j)^{a_j} \overline{\mathrm{Tr}(m^j)}^{b_j} = P_\lambda(m)\overline{P_{\lambda'}(m)},$$

where $\lambda = 1^{a_1} \cdots k^{a_k}$ and $\lambda' = 1^{b_1} \cdots k^{b_k}$. Expand P_λ and $P_{\lambda'}$ in Schur functions and integrate over U_n. Using the orthonormality gives zero unless λ and λ' are partitions of the same number K. If $|\lambda| = |\lambda'| = K$,

$$E\left\{\prod_{j=1}^{k} (\mathrm{Tr}\, M^j)^{a_j} \cdot (\overline{\mathrm{Tr}\, M^j})^{b_j}\right\} = \sum_{\mu \vdash K} \chi_\lambda(\mu) x_{\lambda'}(\mu) = \delta_{\lambda\lambda'} \prod_{j=1}^{k} j^{a_j} a_j!$$

This last equality follows from the second orthogonality relation for characters, the right side being the size of the conjugacy class corresponding to λ when $\lambda = \lambda'$.

The method of moments for complex-valued random variables now proves Theorem 1 from Theorem 2.

Remarks

1. Let Λ_n^k be the homogeneous polynomials in n variables of degree k. Both the Schur functions and the power sum symmetric functions form a basis for Λ_n^k. There is a well known inner product on this space; the Hall inner product (Macdonald (1979), Section I.4). Under this inner product, the Schur functions are orthonormal and the power sum functions satisfy $\langle P_\lambda | P_{\lambda'} \rangle = \delta_{\lambda\lambda'} z_\lambda$. Thus the inner product can be realized by $\langle f | g \rangle = \int_{U_n} f(m)\bar{g}(m)dm$.

2. Theorem 1 allows the determination of the limiting distribution of the trace of a random unitary matrix in any representation. Recall that for any partition of any integer k with at most n parts there is an irreducible representation of U_n with character the Schur function s_λ. With λ fixed, one may inquire about the distribution of $s_\lambda(m)$ where m is chosen uniformly in U_n. For fixed λ and large n, the limiting distribution of $s_\lambda(m)$ can be determined as follows. A classical formula in the subject (see for example, Macdonald (1979)) gives

$$s_\lambda = \sum_\mu \frac{\chi_\mu(\lambda)}{z_\mu} P_\mu$$

with $\chi_\mu(\lambda)$ the μth character of the symmetric group at the λth conjugacy class and $z_\mu = \prod_i i^{a_i} a_i!$ if μ has a_i parts equal to i. Now $P_\mu(m)$ is a polynomial in $\mathrm{Tr}(m)$, $\mathrm{Tr}(m^2), \cdots$. In the large n limit, these traces are independent normally distributed complex normal random variables by Theorem 1. It follows that s_λ has a known limit law.

For example, s_{1^2} is the character of the symmetric tensor representation of U_n. We have $s_{1^2} = \frac{1}{2} P_2 + P_{1^2}$. From Theorem 1, the limiting distribution is the distribution of

$$\tfrac{1}{2}(\sqrt{2} Z_1) + \tfrac{1}{2} Z_2^2$$

with Z_1, Z_2 independent complex standard normal.

https://doi.org/10.1017/S0021900200106989 Published online by Cambridge University Press

We do not know how to elegantly extend these considerations to the orthogonal or symplectic group. The formula for expanding the characters of these groups in power sums is not known in as clean a way. Ram (1991) has given algorithms for computing the coefficients in special cases.

2. Eigenvalues of M

For $m \in U_n$, let $\exp(i\theta_1), \ldots, \exp(i\theta_n)$ be the eigenvalues of m. Put these together as a probability measure on the unit circle S^1:

$$\mu_m = \frac{1}{n}\sum_{j=1}^{n} \delta_{\exp(i\theta_j)}. \tag{2.1}$$

If M is Haar distributed, then μ_M is a random measure and one may inquire about large n limits. One way to study the law of μ_M is through its Fourier transform

$$\hat{\mu}_M(a) = \int_{S^1} \exp(ia\theta)\mu_M(d\theta) = \frac{1}{n}\sum_{j=1}^{n} \exp(ia\theta_j).$$

Thus,

$$\hat{\mu}_M(a) = \frac{1}{n} P_a(M) \qquad \text{if } a = 0, 1, 2, 3, \ldots,$$

$$\hat{\mu}_M(a) = \frac{1}{n} \overline{P_{-a}(M)} \qquad \text{for } a = -1, -2, \ldots.$$

It is straightforward to show that for $\boldsymbol{a}$, $\boldsymbol{b}$ as in Theorem 2, the functions

$$T_{\boldsymbol{a},\boldsymbol{b}}(\mu) = \prod_{i=1}^{k} \hat{\mu}(a_i)\hat{\mu}(-b_i)$$

determine weak star convergence of random probabilities on S^1. Theorem 2 above shows the following.

Corollary 1. *If M is uniform on U_n, and μ_M is defined in* (2.1), *then, for all* $n \geq \sum_{i=1}^{k}(a_i + b_i)$,

$$E\{T_{\boldsymbol{a},\boldsymbol{b}}(\mu_M)\} = \frac{1}{n^{2k}}\delta_{\boldsymbol{ab}} \prod_{j=1}^{k} j^{a_j} a_j!.$$

The uniform distribution ν on S^1 has Fourier transform $\hat{\nu}(0) = 1$, $\hat{\nu}(a) = 0$ for $a \neq 0$ in $\mathbb{Z}$. The above calculations allow us to reach the following conclusion.

Theorem 3. *Let M be uniformly chosen in U_n. Let ν be the uniform distribution on S^1. Then, as n tends to ∞,*

$$\mu_M \Longrightarrow \delta_\nu \qquad \textit{in probability, weak star.}$$

Proof. For any $a \neq 0$ and $n > 2a$, Corollary 1 implies

$$\boldsymbol{E}(\hat{\mu}_M(a)) = 0, \qquad \boldsymbol{E}(|\hat{\mu}_M(a)|^2) = \frac{a}{n^2}.$$

Thus the Fourier transform at a tends to zero in probability. This easily implies the claim.

Remarks

1. Theorem 3 shows that the pattern of eigenvalues on S^1 becomes uniformly distributed while Theorems 0 and 1 show that the sum of the eigenvalues has a standard complex normal distribution with no further norming. This shows that the eigenvalue distribution must be highly regular so a fair amount of cancellation takes place. For contrast, if n points are chosen at random on S^1 the associated empirical measure converges to the uniform measure but their sum converges to complex normal after being divided by $\sqrt{n}$.

Eric Rains (personal communication) has observed that the nth power of the eigenvalues of a random unitary matrix are exactly distributed as n uniform and independent points on the unit circle. This follows easily from Equation (2.2) below. Rains has used this, together with Corollary 1 and Fourier analysis, to show that the number of eigenvalues falling in an interval is n times the interval's length up to an error of order $\log n$. Thus, the eigenvalues are very regularly distributed.

There has been detailed study of the distribution of spacings (difference between consecutive eigenvalues) in the mathematical physics literature. An elegant, readable summary of this work is given by Tracy and Widom (1992). Mehta (1991) is a book-length treatment.

2. The joint density of the eigenvalues of a random unitary matrix is due to Weyl. It has the form

$$(n!)^{-1} \prod_{1 \leq j < k \leq n} |\exp(i\theta_j) - \exp(i\theta_k)|^2. \tag{2.2}$$

Of course, this contains all the eigenvalue information, but unraveling the information needs some work.

3. In a different direction, consider M randomly chosen in U_n. The arguments of Diaconis and Freedman (1987) generalize in a straightforward way to show that $\sqrt{n}M_{11}$ has a limiting standard complex normal distribution. Presumably, the arguments of Diaconis et al. (1992) generalize to show that the joint distribution of a block $\sqrt{n}M_{jk}$, $1 \leq j, k \leq N$ converge to independent standard complex normal variables in variation distance provided $N \ll n^{1/3}$.

4. The remarks in 3 suggest that the variables $\sqrt{n}M_{11}, \sqrt{n}M_{22}, \ldots, \sqrt{n}M_{nn}$ can be strung together to a complex Brownian motion in the large n limit. This meshes with the claims of Theorem 0. Perhaps the same is true of $\sqrt{n}M^j_{kk}$, $1 \leq k \leq n$ and perhaps, in the limit, all of these Brownian motions are independent.

5. The Gram–Schmidt algorithms described in the opening paragraph of this paper are not the fastest way to generate a random element of U_n. Diaconis and Shahshahani (1987) describe a 'subgroup' algorithm which represents a random element of U_n as a product of random complex reflections. Both algorithms use order n^3 operations, but the subgroup algorithm runs considerably faster.

3. The orthogonal group

In this section we derive the analogs of Theorems 1 and 2 for the group of $n \times n$ orthogonal matrices with real entries. Haar measure on O_n can be represented by forming an $n \times n$ random matrix with independent identically distributed standard normal coordinates and then performing the Gram–Schmidt algorithms. This suggests a close relationship between the orthogonal group and the standard normal distribution. The main result of this section calculates the joint moments of $\mathrm{Tr}(M)$, $\mathrm{Tr}(M^2), \ldots, \mathrm{Tr}(M^k)$. These are equal to the joint moments of independent normal variables for n sufficiently large. This extends a result of Diaconis and Mallows (1986).

Theorem 4. *Let M be Haar distributed on O_n. Let $a_1, a_2, \ldots, a_k$ be a vector of non-negative integers. Let $Z_1, Z_2, \ldots, Z_k$ be independent standard normal variables. Let η_j be 1 if j is even and 0 otherwise. Then, for all $n \geq \sum_{i=1}^{k} a_i$,*

$$\boldsymbol{E}\left\{\prod_{j=1}^{k} \mathrm{Tr}(M^j)^{a_j}\right\} = \prod_{j=1}^{k} g_j(a_j) = \prod_{j=1}^{k} \boldsymbol{E}(\sqrt{j}Z_j + \eta_j)^{a_j}.$$

Here,

$$\text{if } j \text{ is odd} \quad g_j(a) = \begin{cases} 0 & \text{if } a \text{ is odd} \\ j^{a/2}(a-1)(a-3)\cdots 1 & \text{if } a \text{ is even,} \end{cases}$$

$$\text{if } j \text{ is even} \quad g_j(a) = \sum_k \binom{a}{2k} j^k (2k-1)(2k-3)\cdots 1.$$

Proof. Let $\exp(i\theta_j)$ denote the eigenvalues of M, $1 \leq j \leq n$. As with Theorem 1

$$\prod_{j=1}^{k} \mathrm{Tr}(M^j)^{a_j} = P_\lambda(\exp(i\theta_1), \ldots, \exp(i\theta_n))$$

with $\lambda = 1^{a_1}2^{a_2}\cdots k^{a_k}$ and P_λ the power sum symmetric function. To continue, we must express these power sums in terms of the characters of O_n. For this, let V be the n-dimensional defining representation and $V^{\otimes k}$ the kth tensor power. The algebra of all linear transformations of $V^{\otimes k}$ which commute with the action of O_n is B_{kn} – the Brauer algebra. Brauer (1937) studied these algebras, giving a remarkable basis for them. This work has been developed by Hanlon and Wales (1989a, b) and Wenzl (1988). The part of this theory needed here was developed by Ram (1991).

Brauer (1937) showed that

$$V^{\otimes k} = \bigoplus_{\mu} M^{\mu} \otimes F^{\mu}$$

for a well-specified set of partitions where F^{μ} is an irreducible representation of O_n and M^{λ} is an irreducible B_{kn} module. Ram (1991) showed

$$P_{\lambda}(x) = \sum_{\mu} \chi^{\mu}_{(k,n)}(\omega) t_{\mu}(x) \tag{3.1}$$

with $t_{\mu}(x)$ the character of O_n corresponding to F^{μ} and $\chi^{\mu}_{(k,n)}$ the character of B_{kn} corresponding to M^{λ}. Here ω may be taken as a permutation with cycle type λ. Integrating over O_n amounts to taking the inner product with the trivial character. This corresponds to the partition $\mu = (0)$. Thus

$$\int_{O_n} P_{\lambda}(M) dM = \chi^{(0)}_{(k,n)}(\omega).$$

Now, Ram (1991) has given a combinatorial rule for evaluating the characters of B_{kn} much like the Murnahan–Nakayama rule for the characters of the symmetric group. He has kindly massaged his rule to produce the formula

$$\chi^{(0)}_{(k,n)}(\omega) = \prod_{j=1}^{k} g_j(a_j) \qquad \text{for } \omega = 1^{a_1} \cdots k^{a_k} \tag{3.2}$$

and $g_j(a)$ given in the statement of Theorem 3.

Remarks

1. At least for n odd, $M^{2j} \in O_n$ must have at least one eigenvalue equal to 1. This accounts for the η_j in the formula. It also appears that with probability 1, M^{2j} has only one eigenvalue equal to 1. For even n, O_n is the disjoint union of matrices with determinant +1 and −1. On the first part, a generic matrix has no real eigenvalue. On the second part, a generic matrix has one eigenvalue 1 and one eigenvalue −1. Thus even powers have two +1 terms in their trace and the second part has total mass 1/2. So $\boldsymbol{E}(\mathrm{Tr}(M^{2j})) = 1$.

2. Diaconis and Mallows (1986) showed that the moments of the trace of M equal the moments of a standard normal variable for all sufficiently large n. This proves the limiting normality by the method of moments. A similar limiting normality follows from Theorem 3: $\mathrm{Tr}(M)$, $\mathrm{Tr}(M^2), \ldots, \mathrm{Tr}(M^j)$ are asymptotically independent normals. When j is odd, the limit of $\mathrm{Tr}(M^j)$ has mean 0 and variance j. When j is even, the limits have mean 1 and variance $j+1$.

3. Weyl (1946), p. 224, gives the joint density of the eigenvalues on O_n. Let $c(\theta) = \exp(i\theta) + \exp(-i\theta)$, let $s(\theta) = \exp(i\theta) - \exp(-i\theta)$. On O_{2n} let the eigenvalues be $\theta_1, \theta_2, \ldots, \theta_n$ and their conjugates. On O_{2n+1} let the eigenvalues be

https://doi.org/10.1017/S0021900200106989 Published online by Cambridge University Press

$\theta_1, \ldots, \theta_n$, their conjugates, and ± 1. Let

$$\Delta^+ = \prod_{j=1}^{n} s(\theta_j/2) \prod_{j<k} (c(\theta_j) - c(\theta_k)).$$

The joint density of $\theta_1, \ldots, \theta_n$ is proportional to $\Delta^+ \bar{\Delta}^+$ on SO_n and to $\Delta^- \bar{\Delta}^-$ on SO_n^C.

It follows as a corollary to Theorem 4 that the eigenvalue distribution on the unit circle tends to the uniform distribution.

Theorem 5. *Let M be uniformly chosen in O_n. Let*

$$\mu_M = \frac{1}{n} \sum_{j=1}^{n} \delta_{\exp(i\theta_j)}$$

be the empirical measure of the eigenvalues of M. Let ν be the uniform distribution on S^1. Then, as n tends to infinity

$$\mu_M \Longrightarrow \delta_\nu \qquad \textit{in probability, weak star.}$$

Proof. Let a be a fixed integer. The Fourier transform of μ_M at a satisfies

$$\hat{\mu}_M(a) = \frac{1}{n} \sum_{j=1}^{n} \exp(ia\theta_j) = \hat{\mu}_m(-a) = \frac{1}{n} P_a(\exp(i\theta_1) \cdots \exp(i\theta_n)).$$

Thus, for $a \neq 0$,

$$E(\mu_M(a)) = \begin{cases} 0 & \text{if } a \text{ is odd} \\ \dfrac{1}{n} & \text{if } a \text{ is even} \end{cases}$$

$$E(|\hat{\mu}_M(a)|^2) = \begin{cases} \dfrac{a}{n^2} & \text{if } a \text{ is odd} \\ \dfrac{a+1}{n^2} & \text{if } a \text{ is even.} \end{cases}$$

Thus, $\hat{\mu}_M(a)$ tends to zero in probability. This easily implies the claimed result.

4. The symplectic group

Let J be the $2n \times 2n$ matrix of form $\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ with all blocks $n \times n$. Let Sp_n be the $2n \times 2n$ unitary matrices m with complex entries such that $mJm^t = J$. Sp_n consists of the matrices preserving an alternating form. It comes up in mechanics and elsewhere. Sp_n is a compact group with eigenvalues occurring in complex conjugate pairs so $\mathrm{Tr}(m^k)$ is real for all k. The following theorem shows that under Haar measure, $\mathrm{Tr}(m), \mathrm{Tr}(m^2), \ldots, \mathrm{Tr}(m^k)$ are asymptotically independent normal random variables.

Theorem 6. *Let M be Haar distributed on Sp_n. With notation as in Theorem* 3, *for all $n \geq \sum_{i=1}^{k} a_i$*

$$E\left\{\prod_{i=1}^{k} \mathrm{Tr}(M^j)^{a_j}\right\} = \prod_{j=1}^{k}(-1)^{(j-1)a_j} g_j(a_j) = \prod_{j=1}^{k} E(\sqrt{j}Z_j - \eta_j)^{a_j}. \tag{4.1}$$

Proof. The left side of (4.1) is the expected value of $P_\lambda(M)$ with $\lambda = 1^{a_1}2^{a_2}\cdots k^{a_k}$. Brauer showed that P_λ can be expressed as a finite sum of characters sp_μ of Sp_n:

$$P_\lambda(x_1, \ldots, x_{2n}) = \prod_{j=1}^{k}(-1)^{(j-1)a_j} \sum_{\mu \vdash f-2i} \chi^{\mu}_{(f,-2n)}(\omega) sp_\mu(x_1, \ldots, x_{2n}).$$

The sum is over $0 \leq i \leq \lfloor f/2 \rfloor$ and partitions μ_n with $f = |\lambda|$, the size of λ. Further, $\chi^{\mu}_{(f,-2n)}$ is the character of the Brauer group, and ω is a permutation of cycle type λ.

Integrating over Sp_n, the only non-zero term in the sum comes from $\mu = 0$. Thus

$$E\left\{\prod_{j=1}^{k} \mathrm{Tr}(M^j)^{a_j}\right\} = \prod_{j=1}^{k}(-1)^{(j-1)a_j} \chi^{0}_{(f,-2n)}(\omega).$$

Now Ram (1991) shows $\chi^{0}_{(f,-2n)}(\omega)$ is given by (3.2) for all $n \geq \sum_{i=1}^{f} a_i$. So the results follow from the symmetry of the normal distribution.

Remarks

1. The uniform distribution of the eigenvalues follows as in Sections 2 and 3.

2. With the notation of Remark 4 of Section 3, Weyl (1946) showed that the joint density of the eigenvalues of a random matrix is proportional to $\Delta\bar{\Delta}$ where

$$\Delta(\theta_1, \theta_2, \ldots, \theta_n) = \prod_{j} s(\theta_j) \prod_{j<k} (c(\theta_j) - c(\theta_k)).$$

5. The symmetric and other groups

For $\lambda > 0$, the Poisson(λ) distribution puts mass $e^{-\lambda}\lambda^k/j!$ on $j = 0, 1, 2, \ldots$. If X has a Poisson(λ) distribution and $x_{(k)} = x(x-1)\cdots(x-k+1)$, $E(X_{(k)}) = \lambda^k$. The following computation shows that if $a_i(\pi)$ is the number of i-cycles in the permutation π, then the joint moments of $a_1, a_2, \ldots, a_k$ equal the joint moments of Poisson variables with parameter 1, $\frac{1}{2}, \ldots, 1/k$. Since the Poisson is determined by its moments, this proves that the $a_i(\pi)$ are asymptotically independent with Poisson $(1/i)$ distributions. For the moments of a_1, see Irwin (1955). Arratia and Tavaré (1992) give bounds for total variation convergence. The corresponding limit theorem was known to Goncharov. Watterson (1974) also studied the moments.

Theorem 7. *Let π be uniformly distributed on the symmetric group S_n. Let $b_1, \dots, b_k \in \{0, 1, \dots, \}$. Let $Z_1, Z_2, \dots, Z_k$ be independent Poisson with parameters $1, \frac{1}{2}, \dots, 1/k$. Then, for all $n \geq \sum_{i=1}^k ia_i$,*

$$\boldsymbol{E}\left(\prod_{i=1}^k a_i(\pi)^{b_i}\right) = \prod_{i=1}^k \boldsymbol{E}(Z_i^{b_i}).$$

Proof. Introduce generating functions:

$$g_n(x_1, \dots, x_n) = \frac{1}{n!} \sum_{\pi \in S_n} \prod_{i=1}^n x_i^{a_i(\pi)}$$
$$(5.1) \qquad g(t; \boldsymbol{x}) = \sum_{n=0}^\infty t^n g_n = \prod_{n=1}^\infty \exp(t^i x_i / i).$$

The last equality is a classical result of Pólya theory, see for example Pólya and Read (1987). The result now follows by differentiating. For notational simplicity let us carry this out for $\boldsymbol{E}(a_1(\pi)_{(b)})$. In (5.1), set $x_2 = x_3 = \cdots = 1$ and $x_1 = x$. Thus

$$g_n(x) = \frac{1}{n!} \sum_{\pi \in S_n} x^{a_1(\pi)}$$

and

$$g(t, x) = \frac{\exp(t(x-1))}{(1-t)}.$$

Differentiate b times in x and set $x = 1$. This becomes

$$\sum_{n=0}^\infty t^n \boldsymbol{E}(a_1(\pi)_{(b)}) = \frac{t^b}{(1-t)}.$$

Thus $\boldsymbol{E}(a_1(\pi)_{(b)}) = 1 = \boldsymbol{E}(Z_1)_{(b)})$ for $n \geq b$. This implies that the first n moments of $a_1(\pi)$ under the uniform distribution on S_n equal the first n moments of Z_1. Computation of joint moments is easy and entirely similar due to the product form of (5.1).

A permutation matrix has all eigenvalues on the unit circle. We are thus faced with cancellation similar to that of previous sections. On the one hand, the trace tends to be a small random number. On the other hand, it is the sum of n points on the unit circle which must somehow cancel. Here, it is easy to make sense of things. Suppose π has a_j j-cycles. The matrix of π is similar to a direct sum of circulants with a_j copies of a j-cycle. Such a j-cycle has eigenvalues $\exp(2\pi ik/j)$, $0 \leq k < j$. These are perfectly regularly distributed so their sum cancels exactly in the trace leaving only the a_i fixed points.

From these considerations it is straightforward to derive the limiting law of the eigenvalues. For $\pi \in S_n$ let $\rho(\pi)$ be the associated permutation matrix with eigenvalues $\exp(i\theta_j)$, $1 \leq j \leq n$. Let

(5.2) $$\mu_\pi = \sum_{j=1}^{n} \delta_{\exp(i\theta_j)}$$

be the associated counting measure. For $j = 1, 2, \ldots$ let μ_j be the uniform (counting) measure on the jth roots of unity. Then

(5.3) $$\mu_\pi = \sum_{j=1}^{\infty} a_j \mu_j$$

with $a_j(\pi)$ the number of cycles of length j in π. Now, the extensive work on the distribution of cycle lengths can be brought to bear. For example, the multiplicity of 1 as an eigenvalue is the number of cycles. This is approximately normal with mean and variance $\log n$. The root of unity closest to this is about $(0.63) \cdot 2\pi/n$, etc. As before, μ_π/n converges to a point mass at the uniform distribution, in probability weak star.

Remark. One can study the distribution of the conjugacy classes on any compact group. It is straightfoward to carry out the results for wreath products (such as the group of symmetries of the cube). Indeed, Pólya theory, as in Pólya and Read (1987), gives a formula like (5.1) and from this the rest follows. There is much interesting work to be done on the classical groups mod p. Stong (1992) followed by Goh and Schmutz (1992) and Hansen and Schmutz (1992) derive results for $GL_n(\mathbb{F}_q)$. Rudvalis and Shinoda (1991) study the analog of the matching problem for all of the classical finite groups (distribution of number of fixed vectors). The present paper aims to show there are lovely results to be discovered in the continuous setting as well.

The eigenvalues of complex representations of general groups can also be studied. Let G be a compact group (perhaps finite) and ρ a representation of G. The character $\chi_\rho = \mathrm{Tr}(\rho(s))$ can be averaged over G. Standard character theory shows that:
$\boldsymbol{E}(\chi_\rho(s))$ is the number of times the trivial representation appears in ρ,
$\boldsymbol{E}(|\chi_\rho(s)|^2)$ is the number of irreducibles in ρ counted with multiplicity.
In particular, if ρ is irreducible (as for the n-dimensional representation of the orthogonal, unitary and symplectic groups considered above) the mean is 0 and the variance is 1. This holds for all groups and any-dimensional representation. For large dimension it shows that there must be a fair amount of cancellation in the eigenvalues. Further, it shows that the empirical distribution of the eigenvalues is approximately uniform as in Section 2.

Acknowledgement

We thank Eric Belsley, Eric Rains, Arun Ram, Richard Stanley, Simon Tavaré and Craig Tracy for their help.

References

ARRATIA, R. AND TAVARÉ, S. (1992) The cycle structure of random permutations. *Ann. Prob.* **20**, 1567–1591.

BRAUER, R. (1937) On algebras which are connected with the semisimple continuous groups. *Ann. Math.* **38**, 854–872.

DIACONIS, P. (1990) Applications of the method of moments in probability and statistics. In *Moments in Mathematics*, ed. H. J. Landau, American Mathematical Society, Providence, RI.

DIACONIS, P. AND FREEDMAN, P. (1987) A dozen de Finetti theorems in search of a theory. *Ann. Inst. H. Poincaré Sup.* **23**, 397–423.

DIACONIS, P. AND MALLOWS, C. (1986) On the trace of random orthogonal matrices. Unpublished manuscript. Results summarized in Diaconis (1990).

DIACONIS, P. AND SHAHSHAHANI, M. (1986) Products of random matrices as they arise in the study of random walks on groups. *Contemp. Math.* **50**, 183–195.

DIACONIS, P. AND SHAHSHAHANI, M. (1987) The subgroup algorithm for generating uniform random variables. *Prob. Eng. Inf. Sci.* **1**, 15–32.

DIACONIS, P., EATON, J. AND LAURITZEN, S. L. (1992) de Finetti style theorems for linear models. *Scand. J. Statist.* **19**, 289–315.

ERDÖS, P. AND TURÁN, P. (1967a, 1967b, 1968) On some problems of statistical group theory II, III, IV, *Acta. Math. Acad. Sci. Hung.* **18**, 151–163; **18**, 309–320; **19**, 413–435.

GOH, W. AND SCHMUTZ, E. (1991) A central limit theorem on $GL_n(\mathbb{F}_q)$. Preprint, Department of Mathematics, Drexel University.

GONCHAROV, V. (1944) On the domain of combinatory analysis. *Amer. Math. Soc. Transl.* **2** (1962), 1–46.

HANLON, P. AND WALES, D. B. (1989a) On the decomposition of Brauer's centralizer algebras. *J. Alg.* **121**, 409–445.

HANLON, P. AND WALES, D. B. (1989b) Eigenvalues connected with Brauer's centralizer algebras. *J. Alg.* **121**, 446–476.

HANLON, P., STANLEY, R. AND STEMBRIDGE, J. (1992) Some combinatorial aspects of the spectra of normally distributed random matrices. *Contemp. Math.* **138**, 151–174.

HANSEN, J. AND SCHMUTZ, E. (1991) How random is the characteristic polynomial of a random matrix. Preprint, Department of Mathematics, Drexel University.

IRWIN, J. O. (1955) A unified derivation of some well-known frequency distribution of interest in biometry and statistics. *J. R. Statist. Soc.* **118**, 389–404.

LARSEN, M. (1993) The normal distribution as a limit of generalized Sato-Take measures. Technical report, Department of Mathematics, University of Pennsylvania.

MACDONALD, I. (1979) *Symmetric Functions and Hall Polynomials.* Oxford University Press.

MEHTA, M. L. (1991) *Random Matrices.* Academic Press, San Diego.

PÓLYA, G. AND READ, R. (1987) *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds.* Springer-Verlag, New York.

RAM, A. (1991) Characters of Brauer's centralizer algebras. Technical report, Department of Mathematics, University of Wisconsin.

RUDVALIS, A. AND SHINODA, K. (1991) An enumeration in finite classical groups. Preprint. Department of Mathematics, University of Massachusetts, Amherst.

SAGAN, B. (1991) *The Symmetric Group.* Wadsworth, Pacific Grove, CA.

SHEPP, L. AND LLOYD, S. (1966) Ordered cycle length in a random permutation. *Trans Amer. Math. Soc.* **121**, 340-357.

STONG, R. (1988) Some asymptotic results on finite vector spaces. *Adv. Appl. Math.* **9**, 167–199.

STONG, R. (1992) The average order of a matrix. Preprint, Department of Mathematics, UCLA.

TAKÁCS, L. (1980) The problem of coincidences. *Arch. Hist. Exact Sci.* **21**, 229–244.

TRACY, C. AND WIDOM, H. (1992) Introduction to random matrices. Technical report, Department of Mathematics, University of California, Davis.

VERSHIK, A. M. AND KEROV, S. V. (1981) Asymptotic theory of characters of the symmetric group. *Funct. Anal.* **15**, 246–255.

WATTERSON, G. (1974) Models for the logarithmic species abundance distribution. *Theoret. Popn Biol.* **6**, 217–250.

WENZL, H. (1988) On the structure of Brauer's centralizer algebras. *Ann. Math.* **128**, 173–193.

WEYL, H. (1946) *The Classical Groups.* Princeton University Press.