Random matrices, genus expansions and the symmetric group

Oisín Flynn-Connolly

August 3, 2018

Supervised by Prof. Neil O'Connell

UCD Summer Research Project 2018

Contents

1 Random Matrices

Definition 1.1. A *random matrix* is a matrix with entries drawn randomly from some distribution.

One can, of course, consider a generic real random matrix with entries drawn independently and randomly from a uniform (or indeed any other) distribution. Such matrices tend to be neglected in practice as we are generally interested in matrices with some restrictions on the spectra. Typically we would like all eigenvalues to be real. This motivates the definition of the simplest commonly studied random matrix distribution - the *Wigner matrices.*

Example 1.2. Consider a Hermitian matrix A:

$$\begin{bmatrix} a_{1,1} & a_{1,2}+ib_{1,2} & \cdots & a_{1,n}+ib_{1,n} \\ a_{2,1}+ib_{2,1} & a_{2,2} & \cdots & a_{2,n}+ib_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1}+ib_{n,1} & a_{n,2}+ib_{n,2} & \cdots & a_{n,n} \end{bmatrix}$$

where each $a_{i,j}$ and $b_{i,j}$ above or on the diagonal is drawn independently from a given Gaussian distribution. All the other entries are then determined by the rule $A = A^\dagger$. As the matrix is Hermitian, the spectra is real. This distribution of matrices is called the *Wigner ensemble.* $\diamondsuit$

It is generally convenient to take the Gaussian distribution we are drawing from to be the one with mean 0 and variance $\frac{1}{2}$, as in this case the probability density for a single variable x is $\frac{1}{\sqrt{\pi}}e^{-x^2}$. In future we will refer to this as *the Gaussian distribution.* Another commonly studied distribution also relies on Hermitian matrices. The key difference is that, in this case, the entries are no longer completely independent.

Example 1.3. Let A be a random matrix with complex entries such that $\mathrm{Re}[A]_{i,j}$ and $\mathrm{Im}[A]_{i,j}$ are drawn independently from the Gaussian distribution. Consider $X = AA^*$. Then X is Hermitian and thus has real eigenvalues. This distribution of matrices is called the *Laguerre ensemble.* $\diamondsuit$

The following is a well known result.

Theorem 1.4. *The joint probability density of the eigenvalues of X is given by $e^{-\sum_i \lambda_i}\prod_{i<j}(\lambda_i - \lambda_j)^2$.*

The next (and last) matrix ensemble to be introduced will be the primary subject of study of this report.

Definition 1.5. Let A be a $2n \times 2n$ matrix with the block structure

$$\begin{bmatrix} X_1 & X_2 \\ X_3 & 0_n \end{bmatrix}$$

where X_1, X_2 and X_3 are generic random $n \times n$ random matrices with independent complex entries drawn from the Gaussian distribution as in the previous example and O_n is the $n \times n$ zero matrix. Then we call the distribution of $X = AA^\dagger$ the *Three Block Ensemble.* In the case where $n = 1$ we call this the *Small Three Block Distribution.*

Lemma 1.6. *The joint eigenvalue distribution of the Small Three Block Ensemble is*

$$e^{-\lambda_1-\lambda_2}(\lambda_1 - \lambda_2)\ln(\frac{\lambda_1}{\lambda_2}) \text{ where } \lambda_1 > \lambda_2.$$

Proof.

$$A = \begin{bmatrix} x_1+ix_2 & x_3+ix_4 \\ x_5+ix_6 & 0 \end{bmatrix} = \begin{bmatrix} z_1 & z_2 \\ z_3 & 0 \end{bmatrix}$$

Then A has probability distribution

$$d\mu(A) = \frac{1}{\pi^3}\prod_{i=1}^{6} e^{-x_i^2}$$

We have

$$X = AA^\dagger = \begin{bmatrix} z_1^2 + z_2^2 & z_1\bar{z_3} \\ z_3\bar{z_1} & z_3^2 \end{bmatrix}$$

Let λ_1 and λ_2 be the eigenvalues of X. Then

$$p(\lambda_1 = a, \lambda_2 = b) = p(\lambda_1 + \lambda_2 = a + b, \lambda_1\lambda_2 = ab)$$

Since

$$\lambda_1 + \lambda_2 = \operatorname{Tr} X$$

$$\lambda_1\lambda_2 = \det X$$

we have

$$d\mu(\lambda_1, \lambda_2) = \int_R d\mu(A)$$

where $R = \{(x_1, x_2, x_3, x_4, x_5, x_6) : x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 = \lambda_1 + \lambda_2, z_2^2 z_3^2 = \lambda_1\lambda_2\}$. This is then equal to

$$\frac{1}{\pi^3} e^{-\lambda_- \lambda_2} \int_R dx_d x_2 dx_3 dx_4 dx_5 dx_6$$

Switch to a triple of polar coordinates $(r_1, \theta_1, r_2, \theta_2, r_3, \theta_3)$. Then

$$\int_R \prod_6^{i=1} dx_i = \int_0^{2\pi} d\theta_1 \int_0^{2\pi} d\theta_2 \int_0^{2\pi} d\theta_3 \int_S r_1 r_2 r_3 dr_1 dr_2 dr_3$$

where $\{(r_1, r_2, r_3) : r_1^2 + r_2^2 + r_3^2 = \lambda_1 + \lambda_2, r_2^2 r_3^2 = \lambda_1\lambda_2\}$. Let $r_i^2 = \tau_i$. Then

$$= \pi^3 \int_T d\tau_1 d\tau_2 d\tau_3$$

with T= $\{(\tau_1, \tau_2, \tau_3) : \tau_1 + \tau_2 + \tau_3 = \lambda_1 + \lambda_2, \tau_2\tau_3 = \lambda_1\lambda_2\}$. So

$$T = \{(\lambda_1 + \lambda_2 - \tau_2 - \frac{\lambda_1\lambda_2}{\tau_2}, \tau_2, \frac{\lambda_1\lambda_2}{\tau_2}) : \lambda_2 \leq \tau_2 \leq \lambda_1\}$$

Now we change coordinates to $(\lambda_1, \tau_2, \lambda_2)$. The Jacobian is $\frac{\lambda_1}{\tau_2} - \frac{\lambda_2}{\tau_2}$. So

$$\int_T d\tau_1 d\tau_2 d\tau_3 = (\lambda_1 - \lambda_2)\int_{\lambda_2}^{\lambda_1} \frac{d\tau_2}{\tau_2} = (\lambda_1 - \lambda_2)(\ln\lambda_1 - \ln\lambda_2)$$

So

$$d\mu(\lambda_1, \lambda_2) = e^{-\lambda_1 - \lambda_2}(\lambda_1 - \lambda_2)\ln(\frac{\lambda_1}{\lambda_2})$$

as required. $\square$

2 Young diagrams

Definition 2.1. A *partition* is any sequence

$$\lambda = (\lambda_1, \lambda_2, \dots \lambda_r \dots)$$

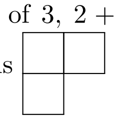
of non-negative integers in decreasing order:

$$l_1 \geq \lambda_2 \geq \cdots \geq \lambda_r \geq \cdots$$

and containing only finitely many non-zero terms. The nonzero λ_i are called the parts of λ, the number of parts is called the length of λ and the sum of parts is called the weight of λ. In particular if the weight of λ is $n \in \mathbb{N}$, λ is called a partition of n.

Definition 2.2. The diagram of a *partition* λ is the set of points $(i, j) \in \mathbb{Z}^2$ such that $1 \leq j \leq \lambda_i$.

It is standard to replace the points in the above definition with 1×1 boxes. This gives us a way to represent partitions in a graphical way. These pictures are called Young diagrams (in the English convention).

Example 2.3. The partition of 3, $2 + 1$ can be written as $(2, 1, 0, 0 \cdots)$. The Young diagram associated with this partition is  ♢

Remark 2.4. Young diagrams with n boxes parametrize irreducible representations of S_n. The correspondence is given by the following algorithm.
Step 1: Write the numbers $\{1, 2, \dots n\}$ in the boxes of λ, left to right then down.
Step 2: Define the following subgroups of S_n:

$$A_\lambda = \{\sigma \in S_n : \sigma \text{ preserves the numbers in each row of } \lambda\}$$

$$B_\lambda = \{\sigma \in S_n : \sigma \text{ preserves the numbers in each column of } \lambda\}$$

Step 3: Working in the group algebra of S_n over $\mathbb{C}$ evaluate the following:

$$a_\lambda = \sum_{\sigma \in A_\lambda} e_\sigma$$

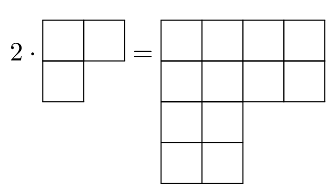
$$b_\lambda = \sum_{\sigma \in B_\lambda} (\operatorname{sgn} \sigma) e_\sigma$$

$$c_\lambda = a_\lambda b_\lambda$$

The space $\mathbb{C}S_n \cdot c_\lambda$ is then an irreducible representation of S_n. ♢

It is possible to define an operation of left multiplication by positive real numbers on the set of Young diagrams. To do this we identify each diagram with its graph. We then extend our understanding of Young diagrams by defining the set of generalized Young diagrams $\mathbb{Y}$ as the set of bounded, non-increasing functions $f : \mathbb{R}^+ \to \mathbb{R}^+$ with a compact support. It is clear that every Young diagram is a generalized Young diagram. If $\lambda \in \mathbb{Y}$ and $c \in \mathbb{R}^+$ we define $c \cdot \lambda(x) = c\lambda(\frac{x}{c})$.

Example 2.5. For a standard Young diagram λ and $n \in \mathbb{N}$, the operation of multiplication corresponds to replacing each 1×1 box of λ with a $n \times n$ box. For instance



♢

3 Kerov polynomials

Definition 3.1. Let $\pi \in S_k$ and let p^λ be an irreducible representation of S_n corresponding to the Young diagram λ. Define the *normalized character* to be

$$\Sigma_\pi^\lambda = \begin{cases} n(n-1)\dots(n-k+1)\frac{\operatorname{Tr} p^\lambda(\pi)}{\dim(p^\lambda)} & \text{if } k \le n \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

We are particularly interested in values of characters on cycles. To that end we introduce the notation $\Sigma^\lambda_{(1,2,\dots k)} = \Sigma^\lambda_k$.

Definition 3.2. Let λ be a generalized Young diagram. Then its free cumulants R_k^λ are given by

$$R_k^\lambda = \lim_{s\to\infty} \frac{1}{s^k}\Sigma^\lambda_{k-1}$$

Remark 3.3. The convergence of free cumulants is nontrivial, so the above should be regarded as both a definition and a lemma.

In the theory of free probability, where the concept originated, free cumulants are defined somewhat differently. The same definition can be used here and provides the normal way to compute these values. To sketch it; one can associate a measure, called the *Kerov transition measure*, to each generalized Young diagram. The moments of this measure can be seen as the coefficients of the power series of a generating function. This function can be formally inverted and, after rescaling, the coefficients of its power series are the free cumulants. $\diamondsuit$

Theorem 3.4. *For each permutation π there is a universal polynomial K_π, with the property that*

$$\Sigma_\pi^\lambda = K_\pi(R_2^\lambda, R_3^\lambda \dots)$$

is true for any generalised Young diagram λ. These polynomials are called the Kerov character polynomials.

Again we are interested in cycles, and define $K_{(1,2\dots,k)} = K_k$.

The mathematician Kerov first announced the above result at a conference in January 2000. The first published proof was given in 2003 by Biane. Subsequently, Kerov polynomials were an active area of research until, in 2011, Dolega, Féray and Śniady [2] gave the following complete combinatorial description of the coefficients of K_k.

Theorem 3.5. *Let $k \ge 1$ and let $s_2, s_3, \dots$ be a sequence of non-negative integers with only finitely many non-zero elements. The coefficient of $R_2^{s_2}R_3^{s_3}\cdots$ in the Kerov polynomial K_k is equal to the number of triples (σ_1, σ_2, q) with the following properties:*

1. *σ_1, σ_2 is a factorization of the cycle i.e. $\sigma_1 \circ \sigma_2 = (1, 2, \dots, k)$;*
2. *the number of cycles of σ_2 is equal to the number of factors in the product $R_2^{s_2}R_3^{s_3}\cdots$;*
3. *the total number of cycles of σ_1 and σ_2 is equal to the degree of the product $R_2^{s_2}R_3^{s_3}\cdots$;*
4. *$q : C(\sigma_2) \to \{2, 3, \dots\}$ is a coloring of the cycles of σ_2 with a property that each color $i \in \{2, 3, \dots\}$ is used exactly s_i times;*
5. *for every set $A \subset C(\sigma_2)$ which is nontrivial there are more than $\sum_{i\in A}(q(i) - 1)$ cycles of σ_1 which intersect $\cup A$;*

Example 3.6. The coefficients of K_k are non-negative integers and the leading term is R_{k+1} $\diamondsuit$

Biane also calculated K_k for small k. The results are given in the following table.

k	K_k
1	R_2
2	R_3
3	$R_4 + R_2$
4	$R_5 + 5R_3$
5	$R_6 + 15R_4 + 5R_2^2 + 8R_2$
6	$R_7 + 35R_5 + 35R_3R_2 + 84R_3$

4 Genus expansions

Definition 4.1. Let X be a $n \times n$ matrix drawn from some ensemble of matrices. Suppose also that $\mathbb{E}(\mathrm{Tr}(X^k)$ can be written as a polynomial in n. Then this polynomial is referred to as the *genus expansion.*

The genus expansion is so-called as the coefficients often have topological interpretations. For example:

Example 4.2. Let X_n be a random $n \times n$ Laguerre matrix. Then $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}(\mathrm{Tr}(X_n^k)) = C_k$, where C_k is the k^{th} Catalan number. C_k can also be interpreted as the number of ways to glue the edges of a $2k$-gon together to produce a genus 0 surface. $\Diamond$

The next theorem of Féray and Śniady [3] is the key result that this report is based on.

Theorem 4.3. *For a given Young diagram λ consider a Gaussian random matrix A^λ with the entries $[A^\lambda]_{i,j}$ independent with $[A^\lambda]_{i,j} = 0$ if box (i,j) does not belong to λ and $\mathrm{Re}[A^\lambda]_{i,j}$ and $\mathrm{Im}[A^\lambda]_{i,j}$ independent Gaussian variables with mean zero and variance $\frac{1}{2}$ if it does. Then*

$$\mathbb{E}[\mathrm{Tr}(A^\lambda(A^\lambda)^\dagger)^n] = -K_n(-R_2^\lambda, -R_3^\lambda, \cdots)$$

where

$$R_i^\lambda = \lim_{s\to\infty} \frac{1}{s^i}\mathbb{E}[\mathrm{Tr}(A^{s\lambda}(A^{s\lambda})^\dagger)^{i-1}]$$

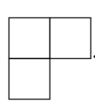
Remark 4.4. Like free cumulants in symmetric group representation theory, it can easily be shown that the R_n^λ are homogeneous of degree n in λ. In other words $R_n^{c\cdot\lambda} = cR_n^\lambda$.

Thus once R_n^λ for some λ is known the genus expansion for the family $[A^{n\lambda}(A^{n\lambda})^\dagger)]_{n>0}$ can be computed by noting

$$\mathbb{E}[\mathrm{Tr}(A^{n\lambda}(A^{n\lambda})^\dagger)^k] = -K_n(-R_2^\lambda n^2, -R_3^\lambda n^3, \cdots).$$

Because R_n^λ is the leading term of the Kerov polynomial, it will be the coefficient of the leading term in the genus expansion. $\Diamond$

5 The three block distribution

We would like to calculate R_i^λ for $\lambda =$ . In our earlier language we are going to study the Small Three Block Distribution.

Lemma 5.1. $\mathbb{E}[\mathrm{Tr}(A^\lambda(A^\lambda)^\dagger)^n] = \Gamma(1+n) + \Gamma(1+n)^2\ {}_2F_1(1+n, 1+n, 2+n, -1)/\Gamma(2+n) - \Gamma(2+n)^2\ {}_2F_1(2+n, 2+n, 3+n, -2)/\Gamma(3+n) - (n+1)\Gamma(1+n)\ {}_3F_2(1,1,2+n,2,2,-1) + \Gamma(3+n)\ {}_3F_2(1,1,3+n,2,2,-1)$

Proof. From Lemma 2.6, we have

$$d\mu(\lambda_1, \lambda_2) = e^{-\lambda_1-\lambda_2}(\lambda_1 - \lambda_2)\ln(\frac{\lambda_1}{\lambda_2}) \text{ where } \lambda_1 > \lambda_2$$

So

$$\mathbb{E}[\mathrm{Tr}(A^\lambda(A^\lambda)^\dagger)^n] = \int_0^\infty \int_{\lambda_2}^\infty (\lambda_1^n + \lambda_2^n)e^{-\lambda_1-\lambda_2}(\lambda_1 - \lambda_2)\ln(\frac{\lambda_1}{\lambda_2})d\lambda_1 d\lambda_2$$

Calculating this integral using Wolfram Mathematica gives the required result. □

Evaluating this expression for small values of n we obtain the following table:

n	$\mathbb{E}[\mathrm{Tr}(A^\lambda(A^\lambda)^*)^n]$
1	3
2	10
3	45
4	248
5	1,610
6	12,024
7	101,556
8	957,312
9	9,964,944
10	113,544,000

We can then use Theorem 5.3 to recursively compute the first values of R_i^λ.

Example 5.2. The computation proceeds as follows:

$$3 = \mathbb{E}[\mathrm{Tr}(A^\lambda(A^\lambda)^\dagger)] = -K_1(-R_2, -R_3, \cdots) = R_2.$$

$$10 = \mathbb{E}[\mathrm{Tr}(A^\lambda(A^\lambda)^\dagger)] = -K_2(-R_2, -R_3, \cdots) = R_3$$

$$45 = \mathbb{E}[\mathrm{Tr}(A^\lambda(A^\lambda)^\dagger)] = -K_3(-R_2, -R_3, \cdots) = R_4 + R_2 \to R_4 = 42$$

$$248 = \mathbb{E}[\mathrm{Tr}(A^\lambda(A^\lambda)^\dagger)] = -K_4(-R_2, -R_3, \cdots) = R_5 + 5R_3 \to R_5 = 198$$

and so on. The results are presented in a table below.

n	R_n
1	1
2	3
3	10
4	42
5	198
6	1001
7	5304

◊

As the leading term of K_k is R_{k+1} it is of mild interest to note that we can continue this procedure indefinitely and thus in principle compute every R_i this way. In practice Kerov polynomials are difficult to compute and this method is of little use in proving anything general about the sequence $\{R_i\}_{i>1}$.

However, if we type the first few values of this sequence into the Online Encyclopedia of Integer Sequences, we obtain a potential match. $R_n = \frac{2}{n}\binom{3n-3}{n-1}$ for $n > 1$. The proof of this is the main result of this report.

Theorem 5.3. *The leading coefficients in the genus expansion of the three block ensemble are* $R_n = \frac{2}{n}\binom{3n-3}{n-1}$ *for* $n > 1$ *and 1 for* $n = 1$.

Proof. For the small three block distribution

$$A^{s\lambda} = \begin{bmatrix} X_1 & X_2 \\ X_3 & 0 \end{bmatrix}$$

where X_i is a $s \times s$ Gaussian random matrix. Now

$$R_{n+1} = \lim_{s\to\infty} \frac{1}{s^{n+1}} \mathbb{E}[\mathrm{Tr}(A^{s\lambda}(A^{s\lambda})^\dagger)^n = \lim_{s\to\infty} \frac{1}{s^{n+1}} \sum_{i_1,i_2,\dots,i_n,j_1,j_2,\dots j_n} \mathbb{E}[a_{i_1j_1}\overline{a_{i_2j_1}}a_{i_2j_2}\overline{a_{i_3j_2}}\cdots a_{i_nj_n}\overline{a_{i_1j_n}}]$$

Each term of this expression corresponds to a walk on a bipartite graph. For the general case - $a_{i_1j_1}\overline{a_{i_2j_1}}a_{i_2j_2}\overline{a_{i_3j_2}}\cdots a_{i_nj_n}\overline{a_{i_1j_n}}$ - we obtain

$$\begin{tikzcd}
i_1 \arrow[d] & i_2 \arrow[d] & i_3 \arrow[d] & \cdots & i_n \arrow[d] \\
j_1 \arrow[ur] & j_2 \arrow[ur] & j_3 \arrow[ur] & \cdots \arrow[ur] & j_n . \arrow[llllu]
\end{tikzcd}$$

For the terms with non-vanishing contribution to the overall sum, each edge must appear at least twice. Thus we have at most n edges in the skeleton of diagram and so at most $n+1$ vertices.

Next we may assume there are $k < n+1$ vertices. The total number of ways to choose k vertices is $\leq C(2s)^k$ where C is a constant not dependent on s. Thus the contribution to the expectation of these terms is $\leq \frac{C'(2s)^k}{s^{n+1}}$ which vanishes as $s \to \infty$.

Lastly, $a_{ij} = 0$ if $i > s$ and $j > s$. So any walk with an edge between i_x and i_y both $> s$ vanishes.

The first two of these conditions tell us we need to look at loops with $n+1$ vertices and n edges. These are exactly the double trees. The last condition gives us a colouring condition on these trees. This gives us the notion of a *three block tree.* These are double trees with $k+1$ vertices equipped with a colouring $h : V \to \{1, 2, \cdots, 2s\}$ such that for each edge (i, j): $h(i) > s \to h(j) \leq s$.

Each three block tree can easily be seen to contribute 1 to the expectation. Thus $\mathbb{E}[\mathrm{Tr}(A^{s\lambda}(A^{s\lambda})^\dagger)^n =$ # three block trees.

We find a map from the set of three block trees to the set of so-called 2-plane trees. A planted plane tree is a rooted tree in which the children of each vertex are ordered. (It is interesting to note the number of planted plane tree on $n+1$ vertices is C_n) A 2-plane tree is a planted plane tree such that each vertex is colored black or white and for each edge at least one of its ends is white. Gu and Prodinger (2009) showed that the total number of such trees is $\frac{2}{n+1}\binom{3n}{n}$.

For a three block tree T, we map it to a tree T' with the same shape. For each $i \in V(T)$:

- if $h(i) > s$ colour the corresponding vertex in T' black.
- if $h(i) \leq s$ colour the corresponding vertex in T' white.

This results in a 2-plane tree. It maps s^{n+1} three block trees to each 2-plane tree. We conclude $R_{n+1} = \frac{2}{n+1}\binom{3n}{n}$ which is our required result. □

6 Acknowledgments

I'd like to thank my supervisor Neil and all the people in the UCD Probability research group, especially Fabio and Antoine, for all their help with the project. You're all marvelous people.

Additional thanks must go to all the staff in the UCD maths office. Lastly thanks to Prof. Robert Osburn for organizing everything.

References

[1] I. G. Macdonald, *Symmetric functions and Hall polynomials*, Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, (1979)

[2] Maciej Dolega, Valentin Féray, Piotr Śniady, *Explicit combinatorial interpretation of Kerov character polynomials as numbers of permutation factorizations*, Adv. Math. 225 **2010**, no. 1, 81-120

[3] Valentin Féray, Piotr Śniady, *Asymptotics of characters of symmetric groups related to Stanley character formula*, Preprint arXiv:math/0701051, **2007**,

[4] N. S. S. Gu and H. Prodinger, *Bijections for 2-plane trees and ternary trees*, European J. of Combinatorics **2009 54** (2004), 969-985.