Moments of characteristic polynomials enumerate two-rowed lexicographic arrays

E. Strahov

Department of Mathematical Sciences,
Brunel University, Uxbridge, UB8 3PH, United Kingdom
Eugene.Strahov@brunel.ac.uk

Submitted: Nov 21, 2001; Accepted: May 14, 2003; Published: May 29, 2003
MR Subject Classifications: 15A52, 20G05, 11M06

Abstract

A combinatorial interpretation is provided for the moments of characteristic polynomials of random unitary matrices. This leads to a rather unexpected consequence of the Keating and Snaith conjecture: the moments of $|\zeta(1/2+it)|$ turn out to be connected with some increasing subsequence problems (such as the last passage percolation problem).

1 Introduction

Keating and Snaith [1] have proposed to model the limiting distribution of the Riemann zeros using the characteristic polynomials of unitary random matrices U

$$Z(U,\theta) = \det(I - e^{-i\theta}U). \tag{1}$$

In particular they deal with a conjecture for moments of $|\zeta(1/2+it)|$ which states that the limit

$$I_m = \lim_{T\to\infty} \frac{1}{(\log T)^{m^2}} \int\limits_0^T |\zeta(1/2+it)|^{2m} dt \tag{2}$$

exists and is equal to a product of two factors, $f(m)$ and $a(m)$, i. e.

$$I_m = a(m)f(m). \tag{3}$$

The first factor $a(m)$ is the zeta-function specific part,

$$a(m) = \prod_p (1-\frac{1}{p})^{m^2} \sum_{k=0}^{+\infty} \left(\frac{\Gamma(k+m)}{k!\ \Gamma(m)}\right)^2 p^{-k} \tag{4}$$

where the product is taken over prime numbers p. As for the second factor $f(m)$, Keating and Snaith [1] have hypothesized that it is the random matrix (universal) part and may be represented as

$$f(m) = \lim_{N\to\infty} N^{-m^2} \langle |Z(U,\theta)|^{2m} \rangle_{U(N)}. \tag{5}$$

Here the average is over the Circular Unitary Ensemble (CUE) of random unitary matrices $N \times N$.

In this paper I provide a combinatorial interpretation for the moments of characteristic polynomials, $\langle |Z(U,\theta)|^{2m} \rangle_{U(N)}$. I then relate these moments with two-rowed lexicographic arrays which are generalizations of permutations and words in combinatorics. (For basic information about permutations, words and lexicographic arrays see, for example, a book by Fulton [2].) The combinatorial interpretation of moments of the characteristic polynomials enables an explicit formula to be obtained for the total number of two-rowed lexicographic arrays constructed from the letters of an alphabet of m symbols with the increasing subsequences of the length at most N.

This combinatorial interpretation is, in fact, a natural consequence of a profound relation between Random Matrix Theory and Robinson-Schensted-Knuth type problems discovered recently by Gessel [3], Baik, Deift and Johansson [4] and Rains [6] (see also Aldous and Diaconis [5] and the references therein). In particular, it follows that certain expectation values over CUE appear in the theory of last passage percolation. The moments of characteristic polynomials , $\langle |Z(U,\theta)|^{2m} \rangle_{U(N)}$, may also be considered in this context. The reasoning along those lines enables the random matrix part $f(m)$ defined by equations (2)-(4) to be related with the weakly-right/weakly-up lattice model of last passage percolation.

2 Increasing subsequences of lexicographic arrays and $\langle |Z(U,\theta)|^{2m} \rangle_{U(N)}$

Let us recall a generalization of the Cauchy identity (see Gessel [3], Baik and Rains [7], Tracy and Widom [8]):

$$\sum_{N\geq\lambda_1\cdots\geq\lambda_m\geq 0} s_\lambda(\xi_1,\xi_2,\ldots,\xi_m)\, s_\lambda(\eta_1,\eta_2,\ldots,\eta_m) = \left\langle \det\left(\prod_{i,j=1}^{m} (I+\xi_i U)\left(I+\eta_j U^\dagger\right)\right)\right\rangle_{U(N)}. \tag{6}$$

Here $s_\lambda(\xi_1,\xi_2,\ldots,\xi_m)$ denotes the Schur polynomial of m variables. Take $\xi_1 = \xi_2 = \cdots = \xi_m = e^{-i\theta}$ and $\eta_1 = \eta_2 = \cdots = \eta_m = e^{i\theta}$ in equation (6). We will use the formula for the Schur polynomials of identical variables,

$$s_\lambda(x,x,\ldots,x) = x^{\lambda_1+\lambda_2+\cdots+\lambda_m} d_\lambda(m) \tag{7}$$

where the coefficient $d_\lambda(m)$ is calculated by the formula

$$d_\lambda(m) = \frac{\prod\limits_{1\le i<j\le m} (\lambda_i - i - \lambda_j + j)}{0!\ 1!\ 2! \ldots (m-1)!}. \tag{8}$$

From equations (6)-(8) we conclude that the moments of the characteristic polynomials $\langle |Z(U,\theta)|^{2m}\rangle_{U(N)}$, may be represented as sums over partitions. Importantly, these sums should be taken only over partitions of length less than m and with the first row less than N. An explicit expression for the m^{th} moment of the characteristic polynomial is

$$\langle |Z(U,\theta)|^{2m}\rangle_{U(N)} = \sum_{K=0}^{+\infty} \sum_{\lambda\vdash K,\ \lambda_1\le N} d_\lambda^2(m), \tag{9}$$

where $\lambda \vdash K$ means that the set $(\lambda_1, \lambda_2, \ldots, \lambda_m)$ is a partition of K, i. e. $\lambda_1 + \lambda_2 + \cdots + \lambda_m = K$. Representation (9) enables a combinatorial interpretation to be provided for $\langle |Z(U,\theta)|^{2m}\rangle_{U(N)}$.

Consider semi-standard Young tableaux constructed from K boxes with at most m and N boxes in the first column and the first row, respectively. The total number of pairs of such tableaux is then given by the sum $\sum\limits_{\lambda\vdash K,\ \lambda_1\le N} d_\lambda^2(m)$. (It is a well-known fact (see Fulton [2]) that the coefficient $d_\lambda(m)$ defined by equation (8) gives the number of semi-standard Young tableaux of the shape λ whose entries are taken from the set $[1, 2, \ldots, m]$). It is this sum which appears in expression (9) for the moments of the characteristic polynomials. In what follows we will apply the Robinson-Schensted-Knuth correspondence (see Fulton [2]) between two-rowed lexicographic arrays and pairs of semi-standard Young tableaux.

By definition, a two-rowed array of the size K is an object of the form

$$A_{2,K}^m = \begin{pmatrix} u_1 & u_2 & \ldots & u_K \\ v_1 & v_2 & \ldots & v_K \end{pmatrix} \tag{10}$$

where the letters $u_1, u_2, \ldots, u_K$ and $v_1, v_2, \ldots, v_K$ belong to an alphabet (any ordered set) $\aleph_m$ of m different letters. The following two conditions ensure that the array $A_{2,K}^m$ corresponds to a pair of semi-standard Young tableaux (each of K boxes with entries taken from $1, 2, \ldots, m$):

$$u_1 \le u_2 \le \cdots \le u_K \tag{11}$$

$$\text{if } \ u_i = u_{i+1} \Rightarrow v_i \le v_{i+1}. \tag{12}$$

The arrays $A_{2,K}^m$ for which conditions (11) and (12) are satisfied are called lexicographic arrays.

Let us define a (weakly) increasing subsequence of the lexicographic array $A_{2,K}^m$ as follows:

1. The element (column) of the array $A_{2,K}^m$ with the index i_1 is less than or equal to that with the index i_2, $\binom{u_{i_1}}{v_{i_1}} \le \binom{u_{i_2}}{v_{i_2}}$, if $v_{i_1} \le v_{i_2}$ and $u_{i_1} \le u_{i_2}$.

2. A subsequence of s elements of the array $A^m_{2,K}$ such that

$$\begin{pmatrix} u_{i_1} \\ v_{i_1} \end{pmatrix} \leq \begin{pmatrix} u_{i_2} \\ v_{i_2} \end{pmatrix} \leq \cdots \leq \begin{pmatrix} u_{i_s} \\ v_{i_s} \end{pmatrix}, \quad i_1 < i_2 < \cdots < i_k \tag{13}$$

will be called a weakly increasing subsequence of the length s.

Let $l(A^m_{2,K})$ be the length of the maximal (weakly) increasing subsequence of the two-rowed lexicographic array $A^m_{2,K}$. Denote by $R^K_{m,N}$ the number of two-rowed lexicographic arrays of size K constructed from the alphabet of m letters $\aleph_m$ and including weakly increasing subsequences of the length at most N, i. e.

$$R^K_{m,N} = \sharp \text{ of } A^m_{2,K} \ \text{ with } \quad l(A^m_{2,K}) \leq N. \tag{14}$$

The number $R_{m,N}$ defined by the sum over size K,

$$R_{m,N} = \sum_{K=0}^{+\infty} R^K_{m,N} \tag{15}$$

gives the number of arrays of an arbitrary size that include weakly increasing subsequences of N elements or less. By the Robinson-Schensted-Knuth correspondence, the number of ordered pairs of semi-standard Young tableaux with m and N boxes in the first column and the first row respectively is equal to the number $R_{m,N}$ of lexicographic arrays.

This observation enables $R_{m,N}$ to be expressed as an average over CUE

$$R_{m,N} = \sum_{K=0}^{+\infty} \sum_{\lambda \vdash K,\ \lambda_1 \leq N} d^2_\lambda(m) = \langle |Z(U,\theta)|^{2m} \rangle_{U(N)}. \tag{16}$$

Here we have used the representation of the moments $\langle |Z(U,\theta)|^{2m} \rangle_{U(N)}$ as the sums over partitions, see equation (9). We thus conclude that the moments of characteristic polynomials, $\langle |Z(U,\theta)|^{2m} \rangle_{U(N)}$, have a clear combinatorial interpretation. They are equal to the number of two-rowed lexicographic arrays constructed from an alphabet of m symbols with the weakly increasing subsequences of the N elements at most.

It is now possible to obtain an explicit formula for the number $R_{m,N}$ of lexicographic arrays. Once we apply the result of Keating and Snaith [1]:

$$\langle |Z(U,\theta)|^{2m} \rangle_{U(N)} = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+2m)}{[\Gamma(j+m)]^2} \tag{17}$$

equality (16) between the number of arrays $R_{m,N}$ and the moments of characteristic polynomials gives

$$R_{m,N} = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+2m)}{[\Gamma(j+m)]^2}. \tag{18}$$

In order to illustrate how formula (18) works consider the following example. Take an alphabet consisting of two letters a and b, i. e. $\aleph_m = (a,b), \quad m = 2$. It may give rise to both lexicographic and non-lexicographic arrays. A typical lexicographic array compounded from the letters of the alphabet (a, b) is

$$\begin{pmatrix} a & a & b & b \\ a & b & b & b \end{pmatrix}.$$

We observe that conditions (11), (12) are satisfied by this particular array. An example of a two-rowed non-lexicographic array is given below:

$$\begin{pmatrix} a & a & b & b \\ a & b & a & b \end{pmatrix}.$$

In this array the second letter of the word $abab$, b, is larger than the third letter of this word a. Thus, condition (12) is not satisfied resulting in a non-lexicographic array.

Let us list all possible two-rowed lexicographic arrays from the alphabet (a, b) with the weakly increasing subsequences of the length at most $N = 2$:

$$\binom{\emptyset}{\emptyset} \quad \binom{a}{a} \quad \binom{a}{b} \quad \binom{b}{b} \quad \binom{b}{a} \quad \binom{ab}{aa} \quad \binom{aa}{ab} \quad \binom{ab}{ba} \quad \binom{aa}{bb} \quad \binom{aa}{aa}$$
$$\binom{bb}{ab} \quad \binom{bb}{bb} \quad \binom{ab}{bb} \quad \binom{bb}{aa} \quad \binom{ab}{ab} \quad \binom{abb}{bab} \quad \binom{aab}{bba} \quad \binom{abb}{baa} \quad \binom{aab}{aba} \quad \binom{aabb}{bbaa}$$

There are 20 arrays with the required properties. It can be easily verified that formula (18) obtained above gives precisely this number, i. e.

$$R_{m,N} = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+2m)}{[\Gamma(j+m)]^2} = 20 \qquad (m = 2, N = 2). \tag{19}$$

3 Moments of characteristic polynomials in the last passage percolation theory

Let us now turn to an interesting interpretation of lexicographic arrays in the framework of the last passage percolation theory. (See, for example, Baik [9] where different last passage percolation models are described and the relations with different expectation values over CUE are explained.) It is a consequence of the relation between the moments of characteristic polynomials and lexicographic arrays that $\langle |Z(U, \theta)|^{2m} \rangle_{U(N)}$ appear also in the last passage percolation problems.

Consider the lexicographic array $A^m_{2,K}$ of the form (10) where the letters $u_i, v_i; i = 1, 2, \ldots, K$ take values in the set of integers $\aleph_m = 1, 2, \ldots, m$. This lexicographic array corresponds to a matrix with non-negative integer entries. The (i, j) entry of this matrix is the number of times the vector $\binom{i}{j}$ occurs in the array $A^m_{2,K}$. For example, the array

$$\begin{pmatrix} 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 3 \\ 1 & 2 & 2 & 1 & 1 & 1 & 1 & 2 & 3 & 3 \end{pmatrix}$$

corresponds to the matrix

$$\begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 0 \\ 1 & 1 & 2 \end{pmatrix}.$$

Let $X(i,j)$, $i,j \in [1,2,\ldots,m]$ be a planar array of non-negative integers. Consider weakly-up/weakly-right paths $\pi's$, $(i_k,j_k)_{k=1}^{l}$, such that $i_1 \le i_2 \le \cdots \le i_l$ and $j_1 \le j_2 \le \cdots \le j_l$. (This model has discussed at length in Baik [9, 10].) Let $(1,1) \nearrow (m,m)$ denote the set of up/right paths from $(1,1)$ to (m,m) in the planar array $X(i,j)$. The entries of the matrix $X(i,j)$, x_{ij}, may be considered as the passage times at knots (i,j). Each planar array of non-negative integers $X(i,j)$ can be assigned the last passage percolation time $T(m,m)$ for travel from $(1,1)$ to (m,m). $T(m,m)$ is defined explicitly by the following expression:

$$T(m,m) = \max_{\pi \in (1,1) \nearrow (m,m)} \sum_{i,j \in \pi} x_{ij}. \tag{20}$$

So far we have used the Robinson-Schensted-Knuth correspondence between pairs of semi-standard Young diagrams and lexicographic arrays. Turning to the last passage percolation problems Robinson-Schensted-Knuth correspondence is a correspondence between planar arrays $X(i,j)$ of non-negative integers and pairs of semi-standard Young tableaux. In this case the last passage percolation time $T(m,m)$ should be considered instead of the length of the longest increasing subsequence. The number of arrays $R_{m,N}$ defined in the previous section by equations (14) and (15) is replaced by the number of all possible $X(i,j)$ of size $m \times m$ with the last passage percolation time $T(m,m)$ less than N. Applying the results of the previous section we find that

$$\sharp \text{ of } X(i,j) \text{ with } T(m,m) \le N = \langle |Z(U,\theta)|^{2m} \rangle_{U(N)}. \tag{21}$$

With the help of the Keating and Snaith formula (equation (17)) an explicit expression is obtained for the number of planar arrays $X(i,j)$ of the size $m \times m$ with the last passage percolation time bounded by N:

$$\sharp \text{ of } X(i,j) \text{ with } T(m,m) \le N = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+2m)}{[\Gamma(j+m)]^2}. \tag{22}$$

4 Combinatorial consequences of the Keating and Snaith conjecture

The above results lead to an interesting interpretation of the Keating and Snaith conjecture. Once m is a non-negative integer, the random matrix (universal) factor $f(m)$ relates the moments I_m of zeta-function with increasing subsequences. On the assumption that the Keating and Snaith conjecture is valid, that is equation (5) holds, it follows from the results of the previous sections that

$$\frac{I_m}{a(m)} = \lim_{N\to\infty} \frac{\sum_{K=0}^{\infty} \sharp \text{ of } A_{2,K}^m \text{ with } l(A_{2,K}^m) \le N}{N^{m^2}}. \tag{23}$$

In other words, the ratio $\frac{I_m}{a(m)}$ defined by equations (2)-(4) is asymptotically equal to the number of lexicographic arrays constructed from an alphabet of m letters with increasing subsequences of at most N elements divided by N^{m^2}.

As a particular case consider the weakly-up/weakly-right last passage percolation model. Here the ratio $\frac{I_m}{a(m)}$ is equal to the whole number of $m \times m$ planar arrays $X(i,j)$ of non-negative integers with last percolation time at most N divided by N^{m^2}:

$$\frac{I_m}{a(m)} = \lim_{N\to\infty} \frac{\sharp \text{ of } X(i,j) \text{ with } T(m,m) \leq N}{N^{m^2}}. \tag{24}$$

Apart from the combinatorial interpretation of the Keating and Snaith conjecture, the above considerations raise a question of how the moments of the Riemann zeta function are related to the representation theory of finite groups.

Acknowledgements.

For valuable comments and discussions that have contributed to this work I am grateful to Yan Fyodorov. This research was supported by EPSRC grant GR/R13838/01 "Random Matrices close to Unitary or Hermitian."

References

[1] Keating J. P. and Snaith N. C., Random Matrix Theory and $\zeta(1/2+it)$ 2000 *Comm. Math. Phys.* **214** 57–89.

[2] Fulton W., *Young tableaux* (Cambridge University Press) 1997.

[3] Gessel I., Symmetric functions and P-recursiveness. 1990 *J. Combin. Theory Ser. A* **53** 257–285.

[4] Baik J., Deift P. and Johansson K., On the distribution of the length of the longest increasing subsequence of random permutations. 1999 *J. Amer. Math. Soc.* **12** 1119–1178.

[5] Aldous D. and Diaconis P., Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem. 1999 *Bull. Amer. Math. Soc. (N.S.)* **36** 413–432

[6] Rains E. M., Increasing subsequences and the classical groups. 1998 *Electron. J. of Combinatorics* **5(1)** R12.

[7] Baik J. and Rains E. M., Algebraic aspects of increasing subsequences. 1999 math. CO/9905084 at xxx.lanl.gov.

[8] Tracy C. A. and Widom H., On the Distributions of the Lengths of the Longest Monotone Subsequences in Random Words. 1999 math. CO/9904042 at xxx.lanl.gov.

[9] Baik J., Riemann-Hilbert problems for last passage percolation. 2001 math. PR/0107079 at xxx.lanl.gov.

[10] Baik J., Random vicious walks and random matrices. 2000 *Comm. Pure Appl. Math.* **53(11)** 1385–1410.