RANDOM MATRICES, MAGIC SQUARES AND MATCHING POLYNOMIALS

by

Persi Diaconis
Alex Gamburd

Technical Report No. 2003-21
August 2003

Department of Statistics
STANFORD UNIVERSITY
Stanford, California 94305-4065
RANDOM MATRICES, MAGIC SQUARES AND MATCHING POLYNOMIALS

by

Persi Diaconis
Department of Mathematics and Statistics
Stanford University

Alex Gamburd
Department of Mathematics
Stanford University

Technical Report No. 2003-21
August 2003

This research was supported in part by
National Science Foundation grant DMS-0072360

Department of Statistics
STANFORD UNIVERSITY
Stanford, California 94305-4065

http://www-stat.stanford.edu
RANDOM MATRICES, MAGIC SQUARES AND MATCHING POLYNOMIALS

PERSI DIACONIS AND ALEX GAMBURD

ABSTRACT. Characteristic polynomials of random unitary matrices have been intensively studied in recent years: by number theorists in connection with Riemann zeta-function, and by theoretical physicists in connection with Quantum Chaos. In particular, Haake and collaborators have computed the variance of the coefficients of these polynomials and raised the question of computing the higher moments. The answer turns out to be intimately related to counting integer stochastic matrices (magic squares). Similar results are obtained for the moments of secular coefficients of random matrices from orthogonal and symplectic groups. Combinatorial meaning of the moments of the secular coefficients of GUE matrices is also investigated and the connection with matching polynomials is discussed.

1. INTRODUCTION

Two noteworthy developments took place recently in Random Matrix Theory. One is the discovery and exploitation of the connections between eigenvalue statistics and the longest-increasing subsequence problem in enumerative combinatorics [1, 4, 5, 47, 59]; another is the outburst of interest in characteristic polynomials of Random Matrices and associated global statistics, particularly in connection with the moments of the Riemann zeta function and other L-functions [41, 14, 35, 36, 15, 16]. The purpose of this paper is to point out some connections between the distribution of the coefficients of characteristic polynomials of random matrices and some classical problems in enumerative combinatorics.

2. SECULAR COEFFICIENTS OF CUE MATRICES AND MAGIC SQUARES

2.1. Secular coefficients of characteristic polynomial. Let $M$ be a matrix in $U(N)$ chosen uniformly with respect to Haar measure. Denote by $e^{i\theta_1}, \ldots, e^{i\theta_N}$ its eigenvalues and consider the characteristic polynomial of $M$:

\begin{equation}
(1) \quad P_M(z) = \det(M - zI) = \prod_{j=1}^{N} (e^{i\theta_j} - z) = (-1)^N \sum_{j=0}^{N} \text{Sc}_j(M) z^{N-j} (-1)^j,
\end{equation}

The second author was supported in part by the NSF postdoctoral fellowship.
where \( S_{C_j}(M) \) is the \( j \)-th secular coefficient of the characteristic polynomial. Note that

\[
(2) \quad S_{C_1}(M) = \text{Tr}(M),
\]

and

\[
(3) \quad S_{C_N}(M) = \det(M).
\]

The moments of traces were studied by Diaconis and Shahshahani [23] and Diaconis and Evans [21] who proved the following result:

**Theorem 1.** (a) Consider \( \mathbf{a} = (a_1, \ldots, a_l) \) and \( \mathbf{b} = (b_1, \ldots, b_l) \) with \( a_j, \ b_j \) nonnegative natural numbers. Let \( Z_1, \ldots, Z_n \) be independent standard complex normal variables. Then for \( N \geq \max \left( \sum_{j=1}^l ja_j, \sum_{j=1}^l jb_j \right) \) we have

\[
(4) \quad \mathbb{E}_{U_N} \prod_{j=1}^l (\text{Tr}M^j)^{a_j} (\text{Tr}M^j)^{b_j} = \int_{U_N} \prod_{j=1}^l (\text{Tr}M^j)^{a_j} (\text{Tr}M^j)^{b_j} \, dM
\]

\[
= \delta_{ab} \prod_{j=1}^l j^{a_j} a_j! = E \left( \prod_{j=1}^l (\sqrt{j}Z_j)^{a_j} (\sqrt{j}Z_j)^{b_j} \right).
\]

(b) For any \( j, k, \ E_{U_N} \text{Tr}(M^j) \overline{\text{Tr}(M^k)} = \delta_{jk} \min(j, k). \)

Moments of the higher secular coefficients were studied by Haake and collaborators [30, 31] who obtained:

\[
(5) \quad \mathbb{E}_{U(N)} S_{C_j}(M) = 0, \quad \mathbb{E}_{U(N)}|S_{C_j}(M)|^2 = 1;
\]

and posed the question of computing the higher moments. The answer is given by the following result, proved in section 2.3.

**Theorem 2.** (a) Consider \( \mathbf{a} = (a_1, \ldots, a_l) \) and \( \mathbf{b} = (b_1, \ldots, b_l) \) with \( a_j, \ b_j \) nonnegative natural numbers. Then for \( N \geq \max \left( \sum_{j=1}^l ja_j, \sum_{j=1}^l jb_j \right) \) we have

\[
(6) \quad \mathbb{E}_{U_N} \prod_{j=1}^l (S_{C_j}(M))^{a_j} (\overline{S_{C_j}(M)})^{b_j} = N_{\mu \bar{\mu}}.
\]

Here \( \mu \) and \( \bar{\mu} \) are partitions \( \mu = (l_1^{a_1} \ldots l_l^{a_l}), \ \bar{\mu} = (l_1^{b_1} \ldots l_l^{b_l}) \) (see section 2.3 for the partition notations) and \( N_{\mu \bar{\mu}} \) is the number of nonnegative integer matrices \( A \) with \( \text{row}(A) = \mu \) and \( \text{col}(A) = \bar{\mu}. \)

(b) In particular, for \( N \geq jk \) we have

\[
(7) \quad \mathbb{E}_{U(N)}|S_{C_j}(M)|^{2k} = H_k(j),
\]

where \( H_k(j) \) is the number of \( k \times k \) nonnegative integer matrices with each row and column summing up to \( j \) — "magic squares".
2.2. Magic Squares. The reader is likely to have encountered objects, which following Ehrhart [26] are referred to as “historical magic squares”. These are square matrices of order $k$, whose entries are nonnegative integers $(1, \ldots, k^2)$ and whose rows and columns sum up to the same number. The oldest such object,

\[ \begin{bmatrix}
4 & 9 & 2 \\
3 & 5 & 7 \\
8 & 1 & 6
\end{bmatrix} \]

first appeared in ancient Chinese literature under the name Lo Shu in the third millennium BC and repeatedly reappeared in the cabballistic and occult literature in the middle ages. Not knowing ancient Chinese, Latin, or Hebrew it is difficult to understand what is “magic” about Lo Shu; it is quite easy to understand however why it keeps reappearing: there is (modulo reflections) only one historic magic square of order 3.

Following MacMahon [45] and Stanley [52], what is referred to as magic squares in modern combinatorics are square matrices of order $k$, whose entries are nonnegative integers and whose rows and columns sum up to the same number $j$. The number of magic squares of order $k$ with row and column sum $j$, denoted by $H_k(j)$, is of great interest; see [22] and references therein. The first few values are easily obtained:

\[ H_k(1) = k!, \]

(9)

corresponding to all $k$ by $k$ permutation matrices (this is the $k$-th moment of the traces leading in the work of Diaconis and Shahshahani to the result on the asymptotic normality, see section 2.4 below);

\[ H_1(j) = 1, \]

(10)

corresponding to $1 \times 1$ matrix $[j]$. We also easily obtain $H_2(j) = j + 1$,

\[ \begin{bmatrix}
\hat{i} & j - \hat{i} \\
j - \hat{i} & \hat{i}
\end{bmatrix} \]

but the value of $H_3(j)$ is considerably more involved:

\[ H_3(j) = \binom{j+2}{4} + \binom{j+3}{4} + \binom{j+4}{4}. \]

(11)

This expression was first obtained by MacMahon in 1915 [45] and a simple proof was found only a few years ago by M. Bona [7]. The main result on $H_k(j)$ is given by the following theorem, proved by Stanley and Ehrhart (see [25, 26, 52, 53, 54]):

**Theorem 3.** (a) $H_k(j)$ is a polynomial in $j$ of degree $(k-1)^2$.

(b) The following relations hold:

\[ H_k(-1) = H_k(-2) = \cdots = H_k(-k+1) = 0, \]

(12)

and

\[ H_k(-k-j) = (-1)^{k-1} H_k(j). \]

(13)
It can be shown that the two statements above are equivalent to

\[(14) \quad \sum_{j \geq 0} H_k(j)x^j = \frac{h_0 + h_1x + \cdots + h_dx^d}{(1-x)^{(k-1)^2+1}}, \quad d = k^2 - 3k + 2,\]

with \(h_0 + h_1 + \cdots + h_d \neq 0\) and \(h_d = h_{d-1} + \cdots + h_{d-k+1}\).

(c) The leading coefficient of \(H_k(j)\) is the relative volume of \(B_k\) - the \(k\)-th Birkhoff polytope, i.e. leading coefficient is equal to \(\frac{\text{vol}(B_k)}{k^{k-1}}\).

By definition, the \(k\)-th Birkhoff polytope is the convex hull of permutation matrices:

\[(15) \quad B_k = \left\{(x_{ij}) \in \mathbb{R}^{k^2} \mid x_{ij} \geq 0; \quad \sum_{i=1}^{k} x_{ij} = 1; \quad \sum_{j=1}^{k} x_{ij} = 1\right\}.\]

For example,

\[H_3(j) = \frac{1}{8}j^4 + \frac{3}{4}j^3 + \frac{15}{8}j^2 + \frac{9}{4}j + 1,\]

and

\[\sum_{j \geq 0} H_3(j)x^j = \frac{1 + x + x^2}{(1-x)^5}.\]

Further, in the example above,

\[\text{vol}(B_3) = \frac{1}{8} \times 9.\]

Of course, the joint mixed moments in (6) involve counting rectangular arrays with general row and column sums. This subject has an extensive literature; see survey article [22]. The latest results on the complexity of this problem may be found in [19]. We will return to the discussion of computational aspects in section 2.4.

2.3. Proof of Theorem 2. Before proceeding with the proof of Theorem 2 we review some basic notions and notations of symmetric function theory, referring the reader to [44, 50, 55] for more details.

A partition \(\lambda\) of a nonnegative integer \(n\) is a sequence \((\lambda_1, \ldots, \lambda_r) \in \mathbb{N}^r\) satisfying \(\lambda_1 \geq \cdots \geq \lambda_r\) and \(\sum \lambda_i = n\). We call \(|\lambda| = \sum \lambda_i\) the size of \(\lambda\). The number of parts of \(\lambda\) is the length of \(\lambda\), denoted \(l(\lambda)\). Write \(m_i = m_i(\lambda)\) for the number of parts of \(\lambda\) that equal to \(i\), so we have \(\lambda = (1^{m_1}2^{m_2}\ldots)\).

The Young diagram of a partition \(\lambda\) is defined as the set of points \((i, j) \in \mathbb{Z}^2\) such that \(1 \leq j \leq \lambda_j\); it is often convenient to replace the set of points above by squares. The conjugate partition \(\lambda'\) of \(\lambda\) is defined by the condition that the Young diagram of \(\lambda'\) is the transpose of the Young diagram of \(\lambda\); equivalently \(m_i(\lambda') = \lambda_i - \lambda_{i+1}\).

A semi-standard Young tableau (SSYT) of shape \(\lambda\) is a filling of the boxes of \(\lambda\) with positive integers such that the rows are weakly increasing and the columns are strictly increasing.
In the figure we exhibited a partition $\lambda = (5, 5, 3, 2) = (1^0 2^1 3^1 5^2)$, and a SSYT $T$ of shape $\lambda$ (we write $\lambda = \text{sh}(T)$). We say that $T$ has type $\alpha = (\alpha_1, \alpha_2, \ldots)$, denoted $\alpha = \text{type}(T)$, if $T$ has $\alpha_i = \alpha_i(T)$ parts equal to $i$. Thus, the SSYT in the figure has type $(2, 3, 3, 0, 2, 4, 1)$. For any SSYT $T$ of type $\alpha$ write

$$x^T = x_1^{\alpha_1(T)} x_2^{\alpha_2(T)} \ldots$$

In our example we have

$$x^T = x_1^2 x_2^3 x_3^3 x_4^0 x_5^2 x_6^4 x_7^1.$$ 

Let $\lambda$ be a partition. We define a Schur function $s_\lambda$ in the variables $x = (x_1, x_2, \ldots)$ as the formal power series

$$s_\lambda(x) = \sum_T x^T,$$

where the sum is over all SSYT’s $T$ of shape $\lambda$. The number of SSYT of shape $\lambda$ and type $\alpha$ is denoted $K_{\lambda \alpha}$, and is called the Kostka number. We have

$$s_\lambda = \sum_\alpha K_{\lambda \alpha} x^\alpha.$$ 

In the course of this paper, in addition to the combinatorial definition given above, we will make use of (all of) the following equivalent definitions of Schur functions.

The classical definition of Schur functions is as a ratio of two determinants:

$$s_\lambda(x) = \frac{\det (x_i^{\lambda_j+n-j})_{i,j=1}^n}{\det (x_i^{n-j})_{i,j=1}^n}.$$ 

Before proceeding with the next definition of Schur functions we remind the reader that the elementary symmetric functions $e_r(x_1, \ldots, x_n)$ are given by

$$e_r(x_1, \ldots, x_n) = \sum_{i_1 < \cdots < i_r} x_{i_1} \ldots x_{i_r},$$

and for a partition $\lambda$ we denote

$$e_\lambda = \prod e_{\lambda_j}.$$
We now ready to give another definition of Schur functions, known as Jacobi-Trudi identity:

\[(21) \quad s_\lambda = \det \left( e_{\lambda_i-i+j} \right)_{i,j=1}^n.\]

Finally, the Schur functions give the irreducible characters of $U(N)$:

\[(22) \quad \mathbb{E}_U \left( s_\lambda(M) s_\mu(M) \right) = \delta_{\lambda\mu}.\]

We now turn to the proof of Theorem 2. First of all we observe that

\[(23) \quad \text{Sc}_{\lambda_j}(M) = e_j(M),\]

where $e_j$ are the elementary symmetric functions defined in (19), and that

\[(24) \quad \prod_{j=1}^l (\text{Sc}_{\lambda_j}(M))^{a_j} (\overline{\text{Sc}_{\lambda_j}(M)})^{b_j} = e_\mu(M) e_{\bar{\mu}}(M),\]

where $\mu$ and $\bar{\mu}$ are partitions $\mu = (1^{a_1} \ldots l^{a_l})$, $\bar{\mu} = (1^{b_1} \ldots l^{b_l})$ and $e_\mu$, $e_{\bar{\mu}}$ are elementary symmetric functions defined in (20). We express the elementary symmetric functions in terms of Schur functions (see p. 335 in [55]):

\[(25) \quad e_\mu = \sum_\lambda K_{\lambda\mu} s_\lambda,\]

where $K_{\lambda\mu}$ is the Kostka number defined preceding (17).

We now integrate over the unitary group and use the fact that the Schur function are irreducible characters expressed in (22), to obtain:

\[(26) \quad \int_{U(N)} e_\mu(M) e_{\bar{\mu}}(M) dM = \sum_{\lambda' \vdash |\mu| = |\bar{\mu}|} K_{\lambda'\mu} K_{\lambda'\bar{\mu}} = N_{\mu\bar{\mu}}\]

where $N_{\mu\bar{\mu}}$ is the number of nonnegative integer matrices $A$ with row$(A) = \mu$ and col$(A) = \bar{\mu}$. The last equality in (26) is the consequence of the Knuth correspondence [43], establishing a bijection between $N$ -matrices $A$ of finite support and ordered pairs of $(P, Q)$ of SSYT of the same shape with type$(P) = \text{col}(A)$ and type$(Q) = \text{row}(A)$. This completes proof of Theorem 2.

2.4. Some consequences. Theorem 2 shows that $\mathbb{E}_U(\text{Sc}_j^a(M)) = 0$ for any fixed $j, a \geq 1$; further for any fixed $j_1, \ldots, j_k$ and $a_1, \ldots, a_k$ we have

\[\mathbb{E}_U(\text{Sc}_{j_1}^{a_1}(M) \ldots \text{Sc}_{j_k}^{a_k}(M)) = 0;\]

it also easily implies that $\text{Sc}_j(M)$ are not independent:

\[\mathbb{E}_U |\text{Sc}_j(M)|^2 |\text{Sc}_k(M)|^2 = j + 1 \neq 1.\]

We further remark, that as a consequence of Theorem 1, Diaconis and Shahshahani have shown that if $M$ is chosen from Haar measure on $U_N$, the
trace of successive powers has limiting Gaussian distributions: as $N \to \infty$, for any fixed $k$ and Borel sets $B_1, \ldots, B_k$

\[(27) \quad P(\text{Tr}M \in B_1, \ldots, \text{Tr}M^k \in B_k) \to \prod_{j=1}^{k} P(\sqrt{j}Z \in B_j),\]

where $Z$ is standard complex normal. This has the following implication for secular coefficients

**Proposition 1.** Let $M$ be chosen uniformly in $U_N$. For fixed $j$ and for any Borel set $B$ we have

\[(28) \quad P\{S_{c_j}(M) \in B\} \to P\{W_j \in B\},\]

where $W_j$ is the polynomial in independent standard complex Gaussian variables $Z_1, \ldots, Z_j$, given by

\[(29) \quad W_j = \frac{1}{j!} \left| \begin{array}{cccc}
Z_1 & 1 & 0 & \cdots & 0 \\
\sqrt{2}Z_2 & Z_1 & 2 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\sqrt{j-1}Z_{j-1} & \sqrt{j-2}Z_{j-2} & \sqrt{j-3}Z_{j-3} & \cdots & j-1 \\
\sqrt{j}Z_j & \sqrt{j-1}Z_{j-1} & \sqrt{j-2}Z_{j-2} & \cdots & Z_1
\end{array} \right|.

For example,

\[S_{c_3}(M) \sim \frac{1}{6}Z_1^3 - \frac{1}{\sqrt{2}}Z_1Z_2 + \frac{1}{\sqrt{3}}Z_3.

This proposition follows easily from (27) and the Newton formula relating elementary and power sum symmetric functions [44, p.28].

Now, since the number of magic squares $H_k(j)$ can be expressed as the $k$-th power of this Gaussian polynomial, this proposition might be useful in computing $H_k(j)$ and its leading coefficient $\text{vol}(B_k)$ — a subject which has received much recent attention (see [6, 11, 19, 20, 24, 46]). The connection with Toeplitz determinants, which is discussed in the next section, might also be of interest in connection with computing $H_k(j)$.

Formula (29) gives the asymptotic distribution of the $j$th secular coefficient for fixed $j$ as $N$ tends to infinity as a polynomial of degree $j$ in independent Gaussian variables. It is natural to ask for limiting distribution as $j$ grows with $N$. For example what is the limiting distribution of the $[N/2]$ secular coefficient? On the one hand, (29) suggests it is a complex sum of independent random variables, so perhaps normal. On the other hand, (5) holds for all $j$ making normality questionable.

Finally, we note that Theorem 2 served as one of the motivations for [17], where integral moments of partial sums of the Riemann zeta function on the critical line were computed and the following result was proved.
Theorem 4. Let \( a_k \) be the arithmetic factor given by

\[
a_k = \prod_p \left( 1 - \frac{1}{p} \right) \left( \sum_{j=0}^{\infty} \frac{d_k(p^j)^2}{p^j} \right) .
\]

Then

\[
\lim_{T \to \infty} \frac{1}{T} \int_0^T \left| \sum_{n=1}^{X} \frac{1}{n^{\frac{1}{2} + it}} \right|^{2k} dt = a_k \gamma_k (\log X)^k + O \left( (\log X)^{k^2 - 1} \right).
\]

Here \( \gamma_k \) is the geometric factor, \( \gamma_k = \text{vol}(P_k) \), where \( P_k \) is the convex polytope of substochastic matrices, defined by the following inequalities (note the similarity with (15)):

\[
P_k = \left\{ (x_{ij}) \in \mathbb{R}^{k \times k} \mid x_{ij} \geq 0; \sum_{i=1}^{k} x_{ij} \leq 1; \sum_{j=1}^{k} x_{ij} \leq 1 \right\}.
\]

3. Connection with the Toeplitz Determinants

For certain functions \( f \) an alternative approach to computing the averages \( \int_{U(N)} f(M) dM \) over the unitary group can be based on the Heine-Szego formula.

Proposition 2. [Heine-Szego formula] For \( f \in L^1(S^1) \) we have:

\[
\frac{1}{(2\pi)^N} \int_0^{2\pi} \cdots \int_0^{2\pi} \prod_{j=1}^{N} f(e^{i\theta_j}) \prod_{1 \leq k \leq l \leq N} |e^{i\theta_k} - e^{i\theta_l}|^2 d\theta_1 \cdots d\theta_N = D_N(f).
\]

Here \( D_N(f) \) is the \( N \times N \) Toeplitz determinant with symbol \( f \):

\[
D_N(f) = \det \left( \hat{f}(j - k) \right)_{0 \leq j, k \leq N},
\]

where \( \hat{f}(\tau) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\tau}) d\theta \). See [9] for a proof and references to early literature.

K. Johansson [38] gave a proof of Diaconis and Shahshahani result (27) using (33) and Szego strong limit theorem for Toeplitz determinants; on the other hand, as explained in [9], the asymptotic normality (27) gives a new proof (and some extensions) of the strong Szego limit theorem.

To apply proposition Proposition 2 in our setting it is convenient to introduce the following polynomial

\[
Q_M(z) = \det(I + Mz) = \sum_{j=0}^{N} S c_j(M) z^j.
\]
The polynomial $Q_M(z)$ is closely related to the characteristic polynomial, in fact
\begin{equation}
Q_M\left(-\frac{1}{z}\right) = \frac{(-1)^N}{z^N} P_M(z).
\end{equation}

With \( Q_M(z) = \sum_{j=0}^{N} \text{Sc}_j(M) z^{-j} \) we then have:
\begin{equation}
\mathbb{E}_{U_N} \left[ Q_M(z_1) \ldots Q_M(z_l) \overline{Q_M(z_{l+1})} \ldots \overline{Q_M(z_m)} \right] = \frac{1}{(z_{l+1} \ldots z_m)^N} D_N(f),
\end{equation}
where
\begin{equation}
f(t) = \frac{1}{t^{m-l}} \prod_{i=1}^{m} (1 + z_i t) = \sum_{r \geq l-m} t^r e_{r+m-l}(z_1, \ldots, z_m).
\end{equation}

Following [9], the Toeplitz determinant with symbol (38) can be computed using the Jacobi-Trudi identity (21) and is found to be equal to \( s_{N-l}(z_1, \ldots, z_m) \). We thus obtain an alternative simple proof of the following result, first established in [16]:

**Theorem 5.** Notation being as above, we have
\begin{equation}
\mathbb{E}_{U_N} \left[ Q_M(z_1) \ldots Q_M(z_l) \overline{Q_M(z_{l+1})} \ldots \overline{Q_M(z_m)} \right] = \frac{s_{N-l}(z_1, \ldots, z_m)}{(z_{l+1} \ldots z_m)^N}.
\end{equation}

We remark that for computing higher moments of secular coefficients the approach presented in section 2.3 seems to be more advantageous. Theorem 5 straightforwardly implies only the following hard-to-unravel result:
\begin{equation}
\mathbb{E}_{U_N} \text{Sc}_{\alpha_1}(M) \ldots \text{Sc}_{\alpha_1}(M) \overline{\text{Sc}_{\alpha_l-1}(M)} \ldots \overline{\text{Sc}_{\alpha_m}(M)} = K_{N-l-m\alpha}.
\end{equation}

The Toeplitz determinant associated with the symbol given by (38) is also closely related to a classical formula of Schmidt and Spitzer; before stating it we briefly review Haake's derivation of (5).

It is implicitly based on the following lemma due to Andréief [3] (see also [58]):

**Lemma 1.** Let \( f(z), g(z) \) be square-integrable functions on \( S^1 \). Then
\begin{equation}
\mathbb{E}_{U_N} \det(f(M)) \det(g(M^1)) = \det \left( \frac{1}{2\pi} \int_{0}^{2\pi} f(e^{i\theta}) g(e^{-i\theta}) e^{i(j-k)\theta} \, d\theta \right)_{0 \leq j, k \leq N}.
\end{equation}

Applying this lemma with \( f(z) = z - \lambda \) and \( g(z) = z - \mu \) with \( z = e^{i\phi} \) and \( \mu = e^{i\chi} \) and letting \( x = e^{i(\phi-\chi)} \), we have that the integral on the right-hand side of equation (41) is given by
\begin{equation}
\frac{1}{2\pi} \int_{0}^{2\pi} (e^{i\theta}-x)(e^{-i\theta}-1) e^{i(j-k)\theta} \, d\theta = (1+x)\delta(j-k) - \delta(j-k+1) - x\delta(j-k-1),
\end{equation}
where \( \delta \) is the Kronecker delta.
where $\delta(k)$ is 0 or 1 as $k$ is nonzero or zero. Denoting the determinant on the right-hand side of (41) by $D_N(x)$ it is easy to see that for this choice of $f$ and $g$ it satisfies the recurrence

$$D_N(x) = D_{N-1}(x)x + 1,$$

whence

$$(43) \quad D_N(x) = \sum_{j=0}^{N} x^j,$$

yielding the proof of (5).

Formula (43) is also an easy consequence of the following result of Schmidt and Spitzer [51] on Toeplitz determinants:

**Theorem 6.** Let $a$ be given by

$$(44) \quad a(t) = t^{-p} \prod_{j=1}^{p+q} (t - \rho_j), \quad (t = e^{i\theta}).$$

If the zeroes $\rho_1, \ldots, \rho_{p+q}$ are pairwise distinct then for every $N \geq 1$,

$$(45) \quad D_N(a) = \sum_{L} C_L w_L^N,$$

where the sum is taken over all $\binom{p+q}{q}$ subsets $L \in \{1, 2, \ldots, p+q\}$ of cardinality $|L| = p$ and with $\bar{L} = \{1, \ldots, p+q\} \setminus L$,

$$(46) \quad w_L = (-1)^q \prod_{j \in L} \rho_j, \quad C_L = \prod_{j \in L} \rho_j \prod_{k \in \bar{L}} (\rho_j - \rho_k)^{-1}.$$

If we let

$$(47) \quad a(t) = a_{-1} t^{-1} + a_0 + t = t^{-1}(t - \rho)(t - \sigma)$$

with $\rho \neq \sigma$, then theorem 6 gives

$$(48) \quad D_N(a) = \frac{\sigma}{\sigma - \rho}(\sigma)^N + \frac{\rho}{\rho - \sigma}(-\rho)^N = (-1)^N \frac{\sigma^{N+1} - \rho^{N+1}}{\sigma - \rho};$$

and setting $\sigma = x, \rho = 1$ this is easily seen to be equivalent to (43).

We now give a simple proof of Theorem 6 using Theorem 5. We have

$$(49) \quad D_N(a(t)) = \prod_{j=1}^{p+q} \left( \frac{1}{-\rho_j} \right)^N D_N(g(t)),$$

where

$$(50) \quad g(t) = t^{-p} \prod_{j=1}^{p+q} (1 - \rho_j t).$$
Consequently
\[(51)\]
\[
D_N(a(t)) = \prod_{j=1}^{p+q} \left( -\frac{1}{\rho_j} \right)^N s_{np}(-\rho_1, \ldots, -\rho_{p+q}) = (-1)^{\rho N} \prod_{j=1}^{p+q} \left( \frac{1}{\rho_j} \right)^N s_{np}(\rho_1, \ldots, \rho_{p+q}).
\]

Now using the classical definition of Schur functions \((18)\), and recalling that
\[(52)\]
\[
\det(x_i^n-j) = \prod_{1 \leq i < j \leq n} (x_i - x_j),
\]
we obtain the Schmidt-Spitzer formula by using Laplace expansion\(^1\) in the first \(p\) rows of the determinant appearing in the numerator of \(S_{np}\) in formula \((51)\):

\[
(53) \begin{vmatrix}
\rho_1^{p+q-1+N} & \rho_1^{p+q-2+N} & \cdots & \rho_1^N \\
\vdots & \vdots & \ddots & \vdots \\
\rho_p^{p+q-1+N} & \rho_p^{p+q-2+N} & \cdots & \rho_p^N \\
\rho_1^{p+q-1} & \rho_1^{p+q-2} & \cdots & \rho_1^0 \\
\rho_p^{p+q-1} & \rho_p^{p+q-2} & \cdots & \rho_p^0 \\
\cdots & \cdots & \ddots & \cdots \\
\rho_1^{p+q} & \rho_1^{p+q} & \cdots & \rho_1^0 \\
\rho_p^{p+q} & \rho_p^{p+q} & \cdots & \rho_p^0 \\
\end{vmatrix}^{-1} \begin{vmatrix}
\rho_1^{p+q-1} & \rho_1^{p+q-2} & \cdots & \rho_1^0 \\
\vdots & \vdots & \ddots & \vdots \\
\rho_p^{p+q-1} & \rho_p^{p+q-2} & \cdots & \rho_p^0 \\
\cdots & \cdots & \ddots & \cdots \\
\end{vmatrix}
\]

4. Secular Coefficients of Random Orthogonal and Symplectic Matrices

In this section we prove the analogues of Theorem 2 for orthogonal group \(O(N)\) and for symplectic group \(Sp(2N)\).

**Theorem 7.** (a) Consider \(a = (a_1, \ldots, a_l)\) with \(a_j\) nonnegative natural numbers. Let \(\mu\) be a partition \(\mu = (1^{a_1} \cdots 1^{a_l})\). Then for \(N \geq \sum_{j=1}^{l} ja_j\) and \(|\mu|\) even we have
\[(54)\]
\[
\mathbb{E}_{O(N)} \prod_{j=1}^{l} (\text{Sc}_j(M))^{a_j} = NSO_{\mu}.
\]

Here \(NSO_{\mu}\) is the number of nonnegative symmetric integer matrices \(A\) with \(\text{row}(A) = \text{col}(A) = \mu\) and with all diagonal entries of \(A\) equal to 0. For \(|\mu|\) odd the expected value in \((54)\) is 0.

(b) In particular, for \(N \geq jk\) and \(jk\) even we have
\[(55)\]
\[
\mathbb{E}_{O(N)} \text{Sc}_j(M)^k = S_k^0(j),
\]

where \(S_k^0(j)\) is the number of \(k \times k\) symmetric nonnegative integer matrices with each row and column summing up to \(j\) and all diagonal entries equal to zero (equivalently \(j\)-regular graphs on \(k\) vertices without loops). For \(jk\) odd the expected value in \((55)\) is 0.

\(^1\)We recall the definition of Laplace expansion. Fix \(p\) rows of matrix \(A\). Then the sum of products of the minors of order \(p\) that belong to these rows by their cofactors is equal to the determinant of \(A\).
Proof: We have $\text{Sc}_j(M) = e_j(M)$ and $\prod_{j=1}^l (\text{Sc}_j(M))^{a_j} = e_{\mu}(M)$; as in the proof of theorem 2, we begin by expressing elementary symmetric functions in terms of Schur functions $e_{\mu} = \sum_{\lambda} K_{\lambda\mu}s_{\lambda}$.

We now integrate over the orthogonal group and use the fact that for integrals of Schur functions we have the following expression (see [48]):

$$E_{O(N)}s_{\lambda}(M) = \begin{cases} 1, & \text{if } \lambda = 2\nu, \ell(\lambda) \leq N; \\ 0, & \text{otherwise}, \end{cases}$$

(56)

where $2\nu$ represents the partition produced by doubling each elements of $\nu$. Note that if $\lambda = 2\nu$, then $\lambda' = \nu^2$, where $\nu^2$ represents the partition produced by writing each element of $\nu$ twice. We thus obtain

$$E_{O(N)}e_{\mu}(M) = \sum_{\lambda = \nu^2} K_{\lambda\mu}.$$  

(57)

Next we observe that the condition $\lambda = \nu^2$ is equivalent to the condition that tableau has all columns of even length. Now we recall that a version of Knuth correspondence [43] establishes the bijection between symmetric matrices of nonnegative integers with column sums given by $\mu$ and tableaux of any shape with content $\mu$, and that furthermore in this correspondence the trace of the matrix is the number of odd length columns of the corresponding tableau. We finally note that symmetric nonnegative matrix whose diagonal elements are all zero corresponds to an adjacency matrix of a graph without loops. This completes proof of Theorem 7.

We remark that for the case of the $2k$-th moment of the first secular coefficient, that is the moments of trace studied in [23], (57) specializes to the following formula:

$$E_{O(N)}\text{Tr}(M)^{2k} = \sum_{\lambda' = 2\nu} K_{\lambda'2\nu} = \sum_{\lambda' = 2\nu} f^\lambda = 1 \times 3 \cdots \times (2k - 1),$$

(58)

where $f^\lambda$ denotes the number of standard tableaux of shape $\lambda$. We refer the reader to [37] for the representation-theoretic significance of this formula and its generalizations.

**Theorem 8.** (a) Consider $a = (a_1, \ldots, a_l)$ with $a_j$ nonnegative natural numbers. Let $\mu$ be a partition $\mu = (1^{a_1} \ldots 1^{a_l})$. Then for $N \geq \sum_{j=1}^l ja_j$ and $|\mu|$ even we have

$$E_{Sp(2N)} \prod_{j=1}^l (\text{Sc}_j(M))^{a_j} = NSP_{\mu}.$$  

(59)

Here $NSP_{\mu}$ is the number of nonnegative symmetric integer matrices $A$ with $\text{row}(A) = \text{col}(A) = \mu$ and with all diagonal entries of $A$ even. For $|\mu|$ odd the expected value in (59) is 0.
(b) In particular, for \( N \geq jk \) and \( jk \) even we have

\[
E_{\text{Sp}(2N)S_c}(M)^k = S_k^{\text{sp}}(j),
\]

where \( S_k^{\text{sp}}(j) \) is the number of \( k \times k \) symmetric nonnegative integer matrices with each row and column summing up to \( j \) and all diagonal entries even (equivalently, the number of \( j \)-regular graphs on \( k \) vertices with loops and multiple edges). For \( jk \) odd the expected value in (60) is 0.

**Proof:** We proceed as in the proof of Theorem 7, this time using the following expression for integrals of the Schur functions over the symplectic group:

\[
E_{\text{Sp}(2N)} s_{\lambda}(M) = \begin{cases} 
1, & \text{if } \lambda = \nu^2, \ l(\lambda) \leq 2N; \\
0, & \text{otherwise},
\end{cases}
\]

(61)

to obtain

\[
E_{\text{Sp}(2N)} e_{\mu}(M) = \sum_{\lambda = 2\nu} K_{\lambda\mu}.
\]

(62)

Next we observe that the condition \( \lambda = 2\nu \) is equivalent to the condition that tableau has all rows of even length. Now we recall that a version of Knuth correspondence introduced by Burge [10] establishes the bijection between symmetric matrices of nonnegative integers with column sums given by \( \mu \) and tableaux of any shape with content \( \mu \); and that furthermore in this correspondence the number of odd diagonal elements of the matrix is equal to the number of odd length rows of the corresponding tableau. We finally note that symmetric nonnegative matrix whose diagonal elements are all even corresponds to an adjacency matrix of a graph with loops and multiple edges. This completes proof of Theorem 8.

**Remark 1.** The limiting distribution of the secular coefficients for both orthogonal and symplectic group can be obtained in exact analogy with the Proposition 1 by invoking Newton's identities to express the secular coefficients in terms of power sums and then using limit theorems for power sums proved in [21, 23]. The analogues of Theorem 4 for \( L \)-functions with orthogonal and symplectic symmetries are proved on [18].

5. **Secular Coefficients of GUE Matrices and Matching Polynomials**

Let \( \mu_N(dM) \) denote the GUE measure on the space \( \mathcal{H}_N \) of hermitian \( N \times N \) matrices; "G" and "U" refer to it being Gaussian and \( U(N) \)-invariant. If we denote the matrix elements by \( m_{jk} = x_{jk} + iy_{jk} \),

\[
(63) \quad \mu_N(dM) = \prod_{1 \leq j < k \leq N} \frac{1}{\pi} e^{-|m_{jk}|^2} dx_{jk} dy_{jk} \prod_{k=1}^N \frac{1}{\sqrt{2\pi}} e^{-\frac{x_{kk}^2}{2}} dx_{kk}.
\]

The eigenvalues of a matrix \( M \) chosen at random with respect to (63) are distributed with the density
\begin{equation}
P_N(\lambda_1, \ldots, \lambda_N) = \prod_{1 \leq j < k \leq N} (\lambda_j - \lambda_k)^2 \prod_{k=1}^{N} \frac{e^{-\frac{\lambda_k^2}{4}}}{k! \sqrt{2\pi}}.
\end{equation}

We will denote the Vandermonde determinant by \( \text{Van}(f_1, \ldots, f_N) \):

\begin{equation}
\text{Van}(f_1, \ldots, f_N) = \prod_{1 \leq j < k \leq N} (f_j - f_k).
\end{equation}

Now consider the characteristic polynomial

\begin{equation}
P_M(x) = \det(Ix - M) = \prod_{j=1}^{N} (x - \lambda_j) = \sum_{j=0}^{N} \text{Sc}_j(M)x^{N-j}(-1)^j,
\end{equation}

where \( \text{Sc}_j(M) \) is the \( j \)-th secular coefficient. We are interested in moments of \( \text{Sc}_j(M) \) with respect to \( \mu_N(M) \).

The combinatorial significance of the higher moments of the first secular coefficient \( \text{Sc}_1(M) \), i.e. the moments of traces, has been thoroughly investigated, starting with the work of Harer and Zagier [33]; see also [32, 29]. Let

\[ C(k, N) = E_{\mu_N} \text{Tr}M^{2k}. \]

Harer and Zagier have shown that this integral is always a positive integer and using the Wick formula obtained the following combinatorial interpretation:

\begin{equation}
C(k, N) = \sum_{0 \leq g \leq k/2} \epsilon_g(k)N^{k+1-2g},
\end{equation}

where \( \epsilon_g(k) \) denotes the number of ways to obtain an orientable surface of genus \( g \) by identifying in pairs the sides of \( 2k \)-gon. They also proved the following formula for \( C(k, N) \) by using rather complicated manipulations with generating functions:

\begin{equation}
C(k, N) = (2k - 1)!! \sum_{j=0}^{k} \binom{N}{j+1} \binom{k}{j} 2^j.
\end{equation}

We refer to [60] for an elegant account of this work and further developments.

In one of his last papers [42] Sergei Kerov gave a simple combinatorial interpretation of the numbers \( C(k, N) \) in terms of rook polynomials or, equivalently, in terms of appropriate involutions. As we show below the combinatorial significance of the moments of higher secular coefficients can also be obtained by considering appropriate involutions, or equivalently, matching polynomials.

We start with the following simple observation (which cannot possibly be new).
Proposition 3. Notation being as above we have

\[ E_{\mu_N}(P_M(x)) = \int P_M(x) d\mu_N(x) = h_N(x), \]

where \( h_N(x) \) is the \( N \)th normalized Hermite polynomial (see the definition in Remark 2 below).

Proof: This follows from Heine's formula [57, p. 27], which can be stated as follows: Let \( \alpha(x) \) be a weight function on the interval \([a, b]\) (where we allow for \( a = -\infty, b = +\infty \)). Then

\[ p_N(x) = \prod_{i=1}^{n} \int_{a}^{b} \cdots \int_{a}^{b} \prod_{i<j}^{N} (x - x_i) \prod_{i=1}^{N} (x_i - x_j)^2 \alpha(x_i) dx_i \tag{69} \]

defines orthogonal polynomials of degree \( N \) with weight function \( \alpha(x) \).

Combining (69) with (64) and (66) yields the proof of the proposition.

Remark 2. Hermite polynomials \( H_n(x) \) are orthogonal with respect to \( e^{-x^2} \) and can be defined by

\[ H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}. \tag{70} \]

Explicitly we have:

\[ H_n(x) = \sum_{k=0}^{[n/2]} \frac{(-1)^k n!}{k! (n-2k)!} (2x)^{n-2k}. \tag{71} \]

The first few Hermite polynomials are as follows:

\[ H_0(x) = 1. \]

\[ H_1(x) = 2x. \]

\[ H_2(x) = 4x^2 - 2. \]

\[ H_3(x) = 8x^3 - 12x. \]

\[ H_4(x) = 16x^4 - 48x^2 + 12. \]

Normalized Hermite polynomials, \( h_N(x) \) are defined by

\[ h_n(x) = 2^{-\frac{n}{2}} H_n\left(\frac{x}{\sqrt{2}}\right), \tag{72} \]

or explicitly:
\begin{equation}
    h_n(x) = \sum_{k=0}^{[n/2]} \frac{(-1)^k n!}{k!(n-2k)!} \frac{x^{n-2k}}{2^k}.
\end{equation}

The first few normalized Hermite polynomials are as follows:

\begin{align*}
    h_0(x) &= 1. \\
    h_1(x) &= x. \\
    h_2(x) &= x^2 - 1. \\
    h_3(x) &= x^3 - 3x. \\
    h_4(x) &= x^4 - 6x^2 + 3.
\end{align*}

Polynomials \( \frac{1}{\sqrt{n!}} h_n(x) \) are orthonormal with respect to the weight \( \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \) and satisfy the following orthogonality relations:

\begin{equation}
    \int_{-\infty}^{\infty} h_n(x) h_m(x) e^{-\frac{x^2}{2}} = \sqrt{2\pi n!} \delta_{nm}
\end{equation}

To bring out the combinatorial significance of Proposition 3, let us recall the following definition.

\textbf{Definition 1.} Given a graph \( G \) with \( n \) vertices and \( m \) edges, let \( p(G, k) \) denote the number of ways in which one can select \( k \) independent edges in \( G \). Let further \( p(G, 0) = 1 \) for all \( G \). Then the matching polynomial \( \alpha(G) \) of the graph \( G \) is given by

\begin{equation}
    \alpha(G) = \alpha(G, x) = \sum_{0}^{m} (-1)^k p(G, k) x^{n-2k}.
\end{equation}

Heilmann and Lieb [34] (see also Godsil and Gutman [28]) have proved that for a complete graph \( K_n \) the matching polynomial \( \alpha(K_n, x) = h_n(x) \); this immediately implies the following corollary.

\textbf{Corollary 1.} Notation being as above, \( \mathbb{E}_{\mu_N} |\text{Sc}_j(M)| \) is equal to the number of \( j \)-matchings in the complete graph \( K_N \):

\begin{equation}
    \mathbb{E}_{\mu_N} |\text{Sc}_j(M)| = \frac{N!}{2^j j!(n-2j)!}.
\end{equation}

We can apply the known results about matching polynomials to \( \mathbb{E}_{\mu_N} |\text{Sc}_j(M)| \). For example, the asymptotic normality of \( \mathbb{E}_{\mu_N} |\text{Sc}_j(M)| \), in the sense made precise below, is implied by the following result of Godsil:
Theorem (Godsil [27]). Let $G$ be a graph and let $p(G)$ be the total number of matchings in $G$. Let $X$ be the random variable whose value is the number of edges in a randomly chosen matching; denote by $m(G)$ its mean and by $\sigma(G)$ its standard deviation. There exist numbers $K$ and $L$ such that for any graph $G$ where $\sigma(G) > K$,

$$
|\frac{\sigma(G)p(G,k)}{p(G)} - \frac{1}{\sqrt{2\pi}} e^{-(k-m(G))^2/2\sigma^2(G)}| < \frac{L}{\sqrt{\sigma(G)}}.
$$

(77)

We also easily obtain the following result which can be viewed as a very simple instance of the universality phenomenon in random matrix theory.

Proposition 4. Let $M$ be a random symmetric matrix of size $N$ with zeroes on the main diagonal and off-diagonal entries taking values $\{+1, -1\}$ with probability $\frac{1}{2}$. Then the expected value of the characteristic polynomial of $M$ is $h_N(x)$.

Proof: Let $G$ be a graph with $n$ vertices and $m$ edges. Let $u = (u_1, \ldots, u_m) \in \{-1, 1\}^m$. Let $G_u$ be the weighted graph obtained from $G$ by associating the weight $u_i$ with the edge $e_i$ for $i = 1, \ldots, m$. Let further $P(G_u)$ be the characteristic polynomial of $G_u$. It was proved by Godsil and Gutman [28] (Corollary 2.2) that

$$
\alpha(G) = 2^{-m} \sum_u P(G_u),
$$

(78)

where the summation is ranging over all $2^m$ distinct $m$-tuples $u$. The proposition follows at once by applying this result to the complete graph $G = K_N$ and recalling that $\alpha(K_N) = h_N$.

Remark 3. In fact it is easy to see that the proof of Corollary 2.2. in [28] applies to any symmetric distribution of the entries.

Now we prove the following generalization of Proposition 3:

Theorem 9. Let $2k$ be a nonnegative integer. Notation being as above, we have

$$
\mathbb{E}_{\mu_N}(P_M^{2k}(x)) = h_N^{(k)}(x),
$$

where $h_N^{(k)}(x)$ is the $N$th monic generalized Hermite polynomial. These are orthogonal polynomials with respect to the weight $|x|^{2k} e^{-x^2}$.

We will give explicit formulas for generalized Hermite polynomials and will discuss their combinatorial significance following the proof of the Theorem.

Proof: Consider the product of characteristic polynomials at $k$ values of $x$ and integrate it with respect to $\mu_N$: 
\[ f_k(x_1, \ldots, x_k) = \mathbb{E}_{\mu_N}[P_M(x_1) \ldots P_M(x_k)] \]

\[
\sum_{j_1, \ldots, j_k} \mathbb{E}_{\mu_N}[S_{c_{j_1}}(M) \ldots S_{c_{j_k}}(M)] x_1^{N-j_1} \cdots x_k^{N-j_k} (-1)^{j_1 + \cdots + j_k} 
\]

\[
= \frac{1}{Z_N} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{N} d\mu(\lambda_i) \text{Van}^2(\lambda_1, \ldots, \lambda_N) \prod_{i=1}^{k} \prod_{j=1}^{N} (x_i - \lambda_j),
\]

where \( Z_N \) is the normalization in the definition of \( \mu_N \):

\[ \mu_N = \frac{1}{Z_N} e^{-\frac{1}{2} \sum_{i=1}^{N} \lambda_i^2} \text{Van}^2(\lambda_1, \ldots, \lambda_N) \]

and

\[ d\mu(\lambda) = e^{-\frac{\lambda^2}{2}} d\lambda. \]

As is well known [60],

\[ Z_N = \prod_{k=1}^{N} k! \]

If we set all of the \( x_1, \ldots, x_k \) to be equal, we obtain a generating function \( f_k(x) \):

\[ f_k(x) = \mathbb{E}_{\mu_N}(P_M^k(x)) = \sum_{j_1, \ldots, j_k} \mathbb{E}_{\mu_N}[S_{c_{j_1}}(M) \ldots S_{c_{j_k}}(M)] x^{N_k - \sum_{i=1}^{k} j_i} (-1)^{\sum_{i=1}^{k} j_i}. \]

Our approach up to (86) follows the method in [8]. In formula (79) we can rewrite the integrand as follows:

\[ \text{Van}(\lambda_1, \ldots, \lambda_N) \prod_{i=1}^{k} \prod_{j=1}^{N} (x_i - \lambda_j) = \frac{\text{Van}(\lambda_1, \ldots, \lambda_N; x_1, \ldots, x_k)}{\text{Van}(x_1, \ldots, x_k)}. \]

Furthermore, we can express \( \text{Van}(\lambda_1, \ldots, \lambda_N) \) as the determinant

\[ \text{Van}(\lambda_1, \ldots, \lambda_N) = \det[p_n(\lambda_j)]_{0 \leq n \leq N-1; 1 \leq j \leq N}, \]

where \( p_n(x) \) are arbitrary monic polynomials of degree \( n \).

Combining (84) with (83) we can rewrite the latter as follows

\[ \text{Van}(\lambda_1, \ldots, \lambda_N; x_1, \ldots, x_k) = \det[p_\alpha(v_\beta)], \]

where

\[
\begin{cases}
0 \leq \alpha \leq N + k - 1 \\
1 \leq \beta \leq N + k
\end{cases}
\]

and
\[ v_\beta = \begin{cases} 
\lambda_\beta, & \text{if } \beta \leq N; \\
2^{\frac{\beta}{2}} \cdot \beta^{\frac{1}{2} - \frac{\beta}{2}}, & \text{if } N < \beta \leq N + k; 
\end{cases} \]

If we now choose the polynomials \( p_n \) to be orthogonal with respect to the measure \( d\mu(\lambda) = e^{-\lambda^2} \), i.e. choose them to be monic Hermite polynomials \( h_n \) we obtain after integrating over the \( N \) eigenvalues:

\[ f_k(x_1, \ldots, x_k) = \frac{1}{Z_N} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{N} d\mu(\lambda_i) \text{Van}(\lambda_1, \ldots, \lambda_N; x_1, \ldots, x_k) \text{Van}(\lambda_1, \ldots, \lambda_N) \]

\[ = \frac{N! \prod_{n=0}^{N-1} n!}{Z_N} \det[h_\alpha(x_\beta)]_{1 \leq \alpha \leq N+k-1, 1 \leq \beta \leq k}. \]

Now (81) and (74) imply that the factor \( \frac{N! \prod_{n=0}^{N-1} n!}{Z_N} \) is equal to 1, so we obtain (86)

\[ f_k(x_1, \ldots, x_k) = \frac{1}{\text{Van}(x_1, \ldots, x_k)} \det \begin{vmatrix} h_N(x_1) & h_{N+1}(x_1) & \cdots & h_{N+k-1}(x_1) \\
h_N(x_2) & h_{N+1}(x_2) & \cdots & h_{N+k-1}(x_2) \\
\vdots & \vdots & \ddots & \vdots \\
h_N(x_k) & h_{N+1}(x_k) & \cdots & h_{N+k-1}(x_k) \end{vmatrix}. \]

If we now set all \( x_1 = \cdots = x_k = x \) in (86) we get:

\[ f_k(x) = \left( -1 \right)^{\frac{k(k-1)}{2}} \pi \prod_{j=0}^{k-1} j! \det \begin{vmatrix} h_N(x) & h_{N+1}(x) & \cdots & h_{N+k-1}(x) \\
h_N'(x) & h_{N+1}'(x) & \cdots & h_{N+k-1}'(x) \\
\vdots & \vdots & \ddots & \vdots \end{vmatrix} \]

\[ = C_k W(h_N(x), h_{N+1}(x), \ldots, h_{N+k-1}(x)). \]

Here \( W(h_N(x), h_{N+1}(x), \ldots, h_{N+k-1}(x)) \) is the Wronskian of polynomials \( h_N(x), h_{N+1}(x), \ldots, h_{N+k-1}(x) \). Now by Christoffel formula [57, p.30],

\[ W(h_N(x), h_{N+1}(x), \ldots, h_{N+k-1}(x)) = h_N^{(k/2)}(x), \]

where \( h_N^{(k/2)}(x) \) are monic generalized Hermite polynomials [13, p.156-157], orthogonal with respect to the weight \( |x|^k e^{-x^2} \). This concludes the proof of Theorem 9.

Monic generalized Hermite polynomials \( h_N^{(k)}(x) \) are related to generalized Hermite polynomials \( H_n^{(k)}(x) \) by the following formula (cf. (72)):

\[ h_n^{(k)}(x) = 2^{-\frac{3}{2}} H_n^{(k)}\left( \frac{x}{\sqrt{2}} \right). \]

Hermite polynomials can be expressed in terms of Laguerre polynomials as follows:

\[ H_{2n}^{(k)}(x) = (-1)^k 2^{2n} n! L_n^{k-\frac{1}{2}}(x^2), \]

\[ H_{2n+1}^{(k)}(x) = (-1)^k 2^{2n+1} n! L_n^{k+\frac{1}{2}}(x^2). \]

Recall that Laguerre polynomials [57, p. 100], \( L_n^\alpha(x) \) with \( \alpha > -1 \) are orthogonal on \([0, \infty)\) with respect to the weight \( e^{-x} x^\alpha \); they are explicitly given by

\[ L_n^\alpha(x) = \sum_{j=0}^{n} \binom{n + \alpha}{n - j} \frac{(-x)^j}{j!}. \]  \(89\)

The first few generalized Hermite polynomials are as follows:

\[ H_0^{(k)}(x) = 1. \]

\[ H_1^{(k)}(x) = 2x. \]

\[ H_2^{(k)}(x) = 4x^2 - 2(1 + 2k). \]

\[ H_3^{(k)}(x) = 8x^3 - 4(3 + 2k)x. \]

\[ H_4^{(k)}(x) = 16x^4 - 16(3 + 2k)x^2 + 4(1 + 2k)(3 + 2k). \]

The connection between Laguerre polynomials and rook placements goes back to Kaplansky and Riordan [39, 49]. Combinatorial properties of generalized Hermite polynomials \( H_n^{(k)}(x) \) were studied by Strehl [56]; we summarize his results in (91) and (92) below. First of all, we note that the number of \( j \)-matchings in a complete graph \( K_N \), figuring in Corollary 1, is equal to the number of involutions of \( S_N \) having \( j \) transpositions. Thus, denoting the set of involutions of the set \( \{1, \ldots, N\} \) by \( \text{Inv}_{[N]} \), and for any \( \sigma \in \text{Inv}_{[N]} \) letting \( \text{fix}(\sigma) \) stand for the number of fixed points and \( \text{trans}(\sigma) \) stand for the number of transpositions, we have the following alternative description of normalized Hermite polynomials \( h_N(x) \):

\[ h_N(x) = \sum_{\sigma \in \text{Inv}_{[N]}} x^{\text{fix}(\sigma)} (-1)^{\text{trans}(\sigma)}. \]  \(90\)

We now give a parallel combinatorial description of \( h_N^{(k)} \). Let \([-N, N] = \{-N, \ldots, -1, 1, \ldots, N\} \) and \([-N, N]_0 = \{-N, \ldots, -1, 0, 1, \ldots, N\} \). Denote by \( \text{Inv}_{[-N, N]} \) the set of involutions of \([-N, N] \) (and by \( \text{Inv}_{[-N, N]}_0 \) the set of involutions of \([-N, N]_0 \) ) and for \( \sigma \in \text{Inv}_{[-N, N]} \) let \( \text{cyc}(\sigma) \) stand for the
number of cycles in the action of $\sigma$ combined with the action of the "mirror" involution sending each $i$ to $-i$ for all $i = 1, \ldots, N$. Then

$$h_{2N}^{(k)}(x) = \sum_{\sigma \in \text{inv}^{[N]}_{N}} (2k + 1)^{\text{cyc}(\sigma)} x^{\text{fix}(\sigma)} (-1)^{\text{trans}(\sigma)},$$

and

$$h_{2N+1}^{(k)}(x) = \sum_{\sigma \in \text{inv}^{[N]}_{N} \cup \text{inv}^{[N]}_{N} \cup \text{inv}^{[N]}_{0}} (2k + 1)^{\text{cyc}(\sigma)} x^{\text{fix}(\sigma)} (-1)^{\text{trans}(\sigma)}.$$

6. Concluding Remarks

(1) It should be possible to prove part (a) of Theorem 3, purely probabilistically from the expression for $H_{k}(j)$ as moments of the polynomial in the i.i.d. Gaussians given in Proposition 1. On the other hand, it would be interesting to understand probabilistic implications of part(b) of Theorem 3.

(2) We also note that $E_{U(N)}|S_{c_j}(M)|^{2\nu} = H_{\nu}(j)$ makes sense for any real values of $\nu$ and that the resulting expression can be thought of as a generalization of $H_{k}(j)$ for integral values of $k$. The methods based on using symmetric functions theory do not easily extend to the non-integral situation; however, the methods based on using the Toeplitz determinants, as discussed in Section 3, do apply.

(3) The type of determinant given in (86) appears in the work of Karlin and McGregor [40] in connection with coincidence probabilities so there might be connections with queues, etc.

(4) It would be very interesting to generalize the result of Proposition 4 to the higher moments, in particular to $h_{N}^{k}$ and to the expression for $C(k, N) = E_{\mu, \nu} \text{Tr}(M)^{2k}$.

(5) In the limit (suitably interpreted) as $N \to \infty$, the coefficients of CUE and GUE matrices should have the same universal behavior. The GUE case seems to be more amenable to Riemann-Hilbert asymptotic analysis.

(6) Finally we remark that a unitary matrix $M$ is conjugate on the one hand to the diagonal matrix with eigenvalues on the diagonal, and on the other hand to the Frobenius, or companion matrix, with first row consisting of the secular coefficients, ones below the main diagonal, and remaining entries zero. This strongly suggests that secular coefficients are (as Gian-Carlo Rota might have put it) "nearly equi-primal" with the eigenvalues and indicates that computing their moments is not the only, and perhaps not the most natural, question to ask.
References


DEPARTMENTS OF MATHEMATICS AND STATISTICS, STANFORD UNIVERSITY, STANFORD, CA 94305

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CA 94305

E-mail address: agamburd@math.stanford.edu