A COMBINATORIAL APPROACH TO MIXED RATIOS OF CHARACTERISTIC POLYNOMIALS

HELEN RIEDTMANN

Abstract. We provide a combinatorial derivation of an asymptotic formula for averages of mixed ratios of characteristic polynomials over the unitary group, where mixed ratios are products of ratios and/or logarithmic derivatives. Our proof of this formula is a generalization of Bump and Gamburd's elegant combinatorial proof of Conrey, Forrester and Snaith's formula for averages of ratios of characteristic polynomials over the unitary group. One application of this formula is an asymptotic expression for sums over zeros of a random characteristic polynomial from the unitary group, which we call an explicit formula for eigenvalues in analogy to what is called an explicit formula in the context of $L$-functions.

1. Introduction

We present a combinatorial method to derive formulas for averages of products of ratios and/or logarithmic derivatives of characteristic polynomials over the unitary group. More concretely, we give combinatorial expressions for integrals of the following type: let $A, B, \ldots, F$ be finite sets of complex variables, then

$$\text{mixed ratio}(A, B, C, D, E, F) = \int_{U(N)} \prod_{\alpha \in A} \chi_g(\alpha) \prod_{\beta \in B} \chi_g^{-1}(\beta) \prod_{\gamma \in C} \chi_g(\gamma) \prod_{\varepsilon \in E} \chi_g(\varepsilon) \prod_{\varphi \in F} \chi_g^{-1}(\varphi) \, dg$$

where $\chi_g$ denotes the characteristic polynomial of the unitary matrix $g \in U(N)$, and the integral over $U(N)$ is taken with respect to Haar measure (normalized so that the total volume is 1). The study of averages of characteristic polynomials of random matrices has attracted considerable attention in recent years. Our interest in averages of the type mixed ratio$(A, B, C, D, E, F)$ is motivated by conjectured connections with number theory discovered by Keating and Snaith [KS00].

Our method produces results of the form

$$\text{mixed ratio} = \text{combinatorial main term} + \text{error term}.$$  

In fact, we are only able to provide a neat asymptotic bound for the error term of mixed ratio$(A, B, C, D, E, F)$ as $N \to \infty$ under the assumption that at least one of the sets of variables is empty. Setting some sets of variables equal to the empty set results in the four theorems discussed in Section 4.3:

- Upon specializing our formula for mixed ratios to $E = \emptyset = F$, we essentially recover a formula for ratios due to Conrey, Forrester and Snaith [CFS03].
- If we prescribe $F = \emptyset$ or $A = \emptyset = D$, we obtain new expressions for the corresponding mixed ratios.

2010 Mathematics Subject Classification. Primary 05E05, 11M50, 15B52.

This research was partially supported by the Forschungskredit of the University of Zurich, grant no. FK-15-089.
• Glossing over a few technical details, mixed ratio(∅, ∅, ∅, ∅, E, F) provides a compact combinatorial expression for the main term of the average of products of logarithmic derivatives. In [CS08], Conrey and Snaith derive a different formula for averages of products of logarithmic derivatives, without using any combinatorial tools. By definition the two expressions for the main term are equal; however, ours is a sum of products of monomial symmetric polynomials, while theirs is a rather complicated ad-hoc expression.

The following expression for products of logarithmic derivatives of completed characteristic polynomials constitutes the principal application of our formula for products of logarithmic derivatives of classic characteristic polynomials (stated in Theorem 4.11). Completed characteristic polynomials Λ₉ are introduced on page 32. Let E and F be sets of non-zero variables having absolute value strictly less than 1, then

\[
\int_{U(N)} \prod_{\varepsilon \in E} \Lambda'_\varepsilon(\varepsilon) \prod_{\varphi \in F} \frac{\Lambda'_{\varphi^{-1}}(\varphi)}{\Lambda_{\varphi^{-1}}(\varphi)} \, dg = \sum_{\lambda} \left( -\frac{N}{2} \right)^{l(E) + l(F) - 2l(\lambda)} z_{\lambda} m_{\lambda}(E) m_{\lambda}(F) + \text{error}.
\]

An asymptotic bound on the error term as \( N \to \infty \) is given in Theorem 5.3. This equality allows us to derive what we call an explicit formula for the eigenvalues of unitary matrices. More precisely, let the function \( z \mapsto f(z, z^2, \ldots, z_n) \) and \( h(z) \) "behave well" in a neighborhood of the unit circle and let \( f(z_1, \ldots, z_n) \) be symmetric, then Theorem 5.4 provides a combinatorial formula for the following expression:

\[
\int_{U(N)} \sum_{1 \leq j_1, \ldots, j_n \leq N} h(\rho_{j_1}) \cdots h(\rho_{j_n}) f(\rho_{j_1}, \ldots, \rho_{j_n}) \, dg
\]

where for every matrix \( g \in U(N) \), \( \mathcal{R}(g) = \{\rho_1, \ldots, \rho_N\} \) stands for the multiset of its eigenvalues.

Our main focus lies on the analogy to explicit formulae for the zeros of so-called L-functions, which generalize the celebrated Riemann \( \zeta \)-function. Following [CFK+05], we view characteristic polynomials as a model for L-functions. In consequence, we regard eigenvalues, which by definition are the zeros of characteristic polynomials, as a model for zeros of L-functions – as conjectured in Montgomery’s pioneering work [Mon73]. Not only is the question motivated by this analogy between characteristic polynomials and L-functions, but our proof of Theorem 5.4 also mirrors the proof of the explicit formula for L-functions given in [RS96], thus making the conjectured connections between eigenvalues of random matrices and zeros of L-functions deeper. The principal difficulty is that the derivation of the explicit formula for L-functions is based on the fact that sufficiently far to the right of the critical line L-functions can be written as Euler products; however, there is no natural analogue for the Euler product in the characteristic polynomials for unitary matrices. In order to circumvent this obstacle, we need an alternative way to describe characteristic polynomials inside the unit circle. This is exactly what the equality in (1.1) provides.

In addition to providing new formulas for averages of mixed ratios of characteristic polynomials over the unitary group (stated in Theorems 4.9 and 4.10), our framework covers averages of both ratios and products of logarithmic derivatives of characteristic polynomials. Having a unified approach might be of relevance to number theoretic interpretations, given that products (i.e. a particular type of ratio) have been used to make predictions about the values taken by L-functions, while products of logarithmic derivatives are connected to predictions about the
zeros of \( L \)-functions. For an overview on random matrix theory predictions for \( L \)-function, we refer the interested reader to the introduction of the author’s thesis [Rie18a].

Our method for computing mixed ratio(\( A, B, C, D, E, F \)) is a generalization of Bump and Gamburd’s combinatorial approach to a formula for averages of ratios of the form mixed ratio(\( A, B, C, D, \emptyset, \emptyset \)) [BG06]. As such it is based on the observation that the integrand is symmetric in both \( R(g) \) and \( K(g) \), which implies that it can be written as a linear combination of the form

\[
\sum_{\mu, \nu} s_{\mu}(R(g)) s_{\nu}(R(g))
\]

where \( s_{\lambda} \) is the Schur function associated to the partition \( \lambda \) (which we define on page 6). We emphasize three new ingredients that make this generalization possible, namely the so-called first overlap identity, a new variant of the Murnaghan-Nakayama rule and an equality that is inspired by the vertex operator formalism.

- We only need the simplest case of the first overlap identity, which is derived in [Rie18b].
- Our variant of the Murnaghan-Nakayama rule (stated in Proposition 3.5) provides an explicit expression for the following signed sum, under quite restrictive assumptions:

\[
\sum_{\mu \cap \lambda \text{ is a } k \text{-ribbon}} (-1)^{ht(\mu \cap \lambda)} s_{\lambda}(X).
\]

Ribbons are defined at the very end of Section 2.1.
- The equality that is related to the vertex operator formalism (stated in Lemma 2.9) describes the interaction between two “power sum” operators on the ring of symmetric functions.

1.1. Structure of this paper. In Section 2 we collect the combinatorial definitions and formulas that our results are based on. Section 3 contains some extensions and variations of the Murnaghan-Nakayama rule, which are used to prove formulas for the average of mixed ratios of characteristic polynomials over the unitary matrices in Section 4. In Section 5 we first introduce the notion of the completed characteristic polynomial of a unitary matrix, and then present an expression for the average of its logarithmic derivatives, which will allow us to derive an explicit formula for the eigenvalues of unitary matrices. We conclude by a brief explanation why an explicit formula of this type is of interest from a number theoretic perspective.

2. Background and notation

Before presenting the required combinatorial background, let us fix some general notation: we use the symbol \( \triangleq \) to denote an equality between the quantities on its left-hand side and its right-hand side which defines the quantity on its left-hand side. Furthermore, LHS/RHS always denotes the left-hand/right-hand side of the equality under consideration.

2.1. Sequences and partitions. For us a sequence is a finite enumeration of elements, such as \( \mathcal{X} = (X_1, \ldots, X_n) \). Its length is the number of its elements, denoted by \( l(\mathcal{X}) = n \). A subsequence \( \mathcal{Y} \) of \( \mathcal{X} \) is a sequence given by \( Y_k = X_{n_k} \) \( 1 \leq n_1 < n_2 < \cdots < n = n \) is an increasing sequence of indices. If \( K \) is a subsequence of \( [n] = (1, \ldots, n) \), then \( \mathcal{X}_K \) is shorthand for the subsequence of \( \mathcal{X} \) that corresponds to the indices in \( K \). In consequence, any sequence of length \( n \)
contains exactly $2^n$ subsequences regardless of the number of repeated elements. If two sequences $X$ and $Y$ of the same length are equal up to reordering their elements, we write $X^{\equiv^\text{sort}} Y$. The union of two sequences $X \cup Y$ is obtained by appending $Y$ to $X$; we sometimes add subscripts to indicate the lengths of the two sequences in question. The complement of a subsequence $Y \subset X$ is the subsequence $X \setminus Y$ of $X$ that satisfies $Y \cup (X \setminus Y) \equiv^\text{sort} X$. All operations on sequences that have not been mentioned are understood to be element wise. For instance, $\text{abs}(X)$ is the sequence of absolute values of the elements of $X$. Moreover, we will write $\text{abs}(\lambda) < 1$ to indicate that all elements of the sequence $\lambda$ are strictly less than 1 in absolute value. We do not denote the sequence of absolute values by $|X|$ (which would be more consistent with our usage of applying operations on sequences element by element) because vertical bars traditionally denote the size of a sequence or partition.

For sequences whose elements can be subtracted and multiplied, we define the following two functions:

$$\Delta(X) = \prod_{1 \leq i < j \leq n} (X_i - X_j) \quad \text{and} \quad \Delta(X; Y) = \prod_{x \in X, y \in Y} (x - y).$$

We implicitly view all sets of variables as sequences but for simplicity of notation we will not fix the order of the variables explicitly. It is important, however, to stick to one order throughout a computation or within a formula.

A partition is a non-increasing sequence $\lambda = (\lambda_1, \ldots, \lambda_n)$ of non-negative integers, called parts. If two partitions only differ by a sequence of zeros, we regard them as equal. By an abuse of notation, we say that the length of a partition is the length of the subsequence that consists of its positive parts. The size of a partition $\lambda$ is the sum of its parts, denoted $|\lambda|$. For any positive integer $i$, $m_i(\lambda)$ is the number of parts of $\lambda$ that are equal to $i$. It is sometimes convenient to use a notation for partitions that makes multiplicities explicit:

$$\lambda = \left\{1^{m_1(\lambda)}2^{m_2(\lambda)}\ldots i^{m_i(\lambda)}\ldots \right\}.$$  

The following statistic on the multiplicities will appear in some of our results:

$$z_\lambda = \prod_{i \geq 1} i^{m_i(\lambda)} m_i(\lambda)!$$  

We will frequently view partitions as Ferrers diagrams. The Ferrers diagram associated to a partition $\lambda$ is defined as the set of points $(i, j) \in \mathbb{Z} \times \mathbb{Z}$ such that $1 \leq i \leq \lambda_j$; it is often convenient to visualize the points as square boxes. For instance, the Ferrers diagrams associated to partitions of the type $\langle m^n \rangle$ are just rectangles. The conjugate partition $\lambda'$ of $\lambda$ is given by the condition that the Ferrers diagram of $\lambda'$ is the transpose of the Ferrers diagram of $\lambda$. We note for later reference that if the union of two partitions $\mu$ and $\nu$ happens to be a partition, then $(\mu \cup \nu)' = \mu' + \nu'$. Given two partitions $\kappa$ and $\lambda$, we say that $\kappa$ is a subset of $\lambda$ if their Ferrers diagrams satisfy that containment relation. Note that $\kappa \subset \lambda$ is our shorthand for both subset and subsequence. It will be clear from the context whether we view $\kappa$ and $\lambda$ as sequences or diagrams. For a partition $\lambda$ that is contained in the rectangle $\langle m^n \rangle$, we call the partition

$$\tilde{\lambda} = (m - \lambda_n, \ldots, m - \lambda_1) \subset \langle m^n \rangle$$

its $(m, n)$-complement.

If $\mu$ is a subset of $\lambda$, then the corresponding skew diagram is the set of boxes $\lambda \setminus \mu$ that are contained in $\lambda$ but not in $\mu$. A ribbon is a skew diagram that is connected and contains no $2 \times 2$ subset of boxes. What we call ribbon is also known as skew or rim hook [Sag01, p. 180], and as border strip [Mac98, p. 5]. Let us illustrate this...
definition by some examples. The left-most diagram is a ribbon, while the other two both violate one of the conditions.

![Diagrams](image)

The diagram in the middle illustrates that we only consider edgewise connected skew diagrams connected.

The size of a ribbon is the number of its boxes. We sometimes call a ribbon of size $k$ a $k$-ribbon. The height (ht) of a ribbon is one less than the number of its rows. We use the following shorthand for the property that $\lambda \downarrow \mu$ is a $k$-ribbon:

$$\mu \xrightarrow{k} \lambda.$$ 

We note for later reference that $\lambda \downarrow \mu$ is a $k$-ribbon if and only if $\lambda' \downarrow \mu'$ is. In that case,

$$\text{ht} (\lambda' \downarrow \mu') = k - 1 - \text{ht} (\lambda \downarrow \mu). \quad (2.1)$$

For sequences $\lambda^{(0)} \xrightarrow{k_1} \lambda^{(1)} \xrightarrow{k_2} \ldots \xrightarrow{k_n} \lambda^{(n)}$, the symbol $\text{ht} (\lambda^{(n)} \downarrow \lambda^{(0)})$ denotes the sum of the heights of the intermediate ribbons.

### 2.2. The ring of symmetric functions

In this section we introduce the most commonly used symmetric polynomials. In addition, we will briefly discuss the more abstract concept of symmetric functions, which is necessary in order to define specializations and operators.

**Definition 2.1** (monomial symmetric polynomials) Let $\mathcal{X} = (x_1, \ldots, x_n)$ be a set of variables and let $\lambda$ be a partition. If $l(\lambda) > n$, then the monomial symmetric polynomial $m_\lambda (\mathcal{X})$ is identically zero; otherwise,

$$m_\lambda (\mathcal{X}) = \sum_{(\alpha_1, \ldots, \alpha_n) \vdash \lambda} x_1^{\alpha_1} \cdots x_n^{\alpha_n}.$$ 

We remark that this definition makes use of the convention that any partition of length less than $n$ may be viewed as a sequence of length exactly $n$ by appending zeros.

These polynomials are called symmetric because they are invariant under permutations of the elements of $\mathcal{X}$. The following definition lists three other commonly used families of symmetric polynomials.

**Definition 2.2.** (power sums, elementary and complete symmetric polynomials) Let $k$ be a positive integer and let $\mathcal{X}$ be a set of variables.

1. The $k$-th elementary symmetric polynomial $e_k (\mathcal{X})$ is given by $m_{(k)} (\mathcal{X})$, which is equal to the sum of all products of $k$ variables with distinct indices.
   
   We use the convention that $e_0 (\mathcal{X}) = 1$.

2. The $k$-th complete symmetric polynomial $h_k (\mathcal{X})$ is equal to $\sum_{\lambda : |\lambda| = k} m_\lambda (\mathcal{X})$.
   
   We use the convention that $h_0 (\mathcal{X}) = 1$.

3. The $k$-th power sum $p_k (\mathcal{X})$ is defined by $m_{(k)} (\mathcal{X}) = \sum_{x \in \mathcal{X}} x^k$.
   
   We remark that for any set of variables $\mathcal{X}$, the $l(\mathcal{X})$-th elementary polynomial $e_{l(\mathcal{X})} (\mathcal{X})$ is simply the product of all variables. This observation motivates the following non-standard notation:

$$e (\mathcal{X}) = \prod_{x \in \mathcal{X}} x.$$
For theoretical considerations, it is often more convenient to work with symmetric functions instead of symmetric polynomials as they are not dependent on a set of variables. The monomial symmetric function corresponding to $\lambda$, which we denote by $m_{\lambda}$, is determined by the condition that for any set of variables $X$, $m_{\lambda}(X)$ is the monomial symmetric polynomial defined above. We will freely use this trick to get rid of the set of variables for all symmetric polynomials.

**Definition 2.3 (ring of symmetric functions).** The ring of symmetric functions $(\text{Sym})$ is the complex vector space spanned by the monomial symmetric functions $m_{\lambda}$ where $\lambda$ runs over all partitions.

Owing to the fact that the product of two symmetric polynomials is again symmetric, $\text{Sym}$ is endowed with a natural ring structure. For a rigorous definition of the ring of symmetric functions consult [Mac95, p. 17-19]. It turns out that the monomial symmetric functions are not the only natural basis for $\text{Sym}$. If we use the convention that for any partition $\lambda$,

$$p_{\lambda} = \prod_{i \geq 1} p_{m_{i}(\lambda)}^i,$$

then the $p_{\lambda}$ also form a basis of the ring of symmetric functions [Mac95, p. 24]. In fact, the same holds for the elementary and complete symmetric functions [Mac95, p. 20 and 22].

2.2.1. **Schur functions.** Arguably the most natural basis for $\text{Sym}$ is given by the Schur functions. We will see that they are orthonormal with respect to the Hall inner product. Moreover, they are the main link between the theory of symmetric functions and representation theory. We follow [Mac95] in our presentation of Schur functions.

**Definition 2.4 (Schur functions).** Let $X$ be a set of $n$ pairwise distinct variables and $\lambda$ a partition. If $l(\lambda) > n$, then $s_{\lambda}(X) = 0$; otherwise,

$$s_{\lambda}(X) = \frac{\det (x^{\lambda_j+n-j})_{x \in X, 1 \leq j \leq n}}{\Delta(X)}$$

where $\Delta(X)$ denotes the product of all pairwise differences of elements in $X$. The fact that the polynomial $\Delta(X)$ is a divisor of the determinant in the numerator implies that $s_{\lambda}(X)$ is a homogeneous polynomial of degree $|\lambda|$, which allows us to extend this definition to all sets of variables of length $n$.

Technically, this defines a symmetric polynomial - not a symmetric function. For historical reasons, we call both $s_{\lambda}(X)$ and $s_{\lambda}$ the Schur function indexed by the partition $\lambda$. There are various definitions for Schur functions, each emphasizing a different aspect. In fact, their combinatorial definition, for which we refer the interested reader to [Sag01], will also play a minor role.

The Hall inner product on $\text{Sym}$ is given by the condition that $\langle h_{\lambda}, m_{\mu} \rangle = \delta_{\lambda\mu}$ for all partitions $\lambda, \mu$ where $\delta_{\lambda\mu}$ is the Kronecker delta. In order to state the main property of this inner product, we need to introduce the vector space $\text{Sym}^k$, which is spanned by $m_{\lambda}$ where $\lambda$ runs over all partitions of size $k$. For each $k \geq 0$, let $u_{\lambda}, v_{\lambda}$ be bases of $\text{Sym}^k$, indexed by partitions of size $k$. Then the following conditions are equivalent:

1. For all partitions $\lambda, \mu$, $\langle u_{\lambda}, v_{\mu} \rangle = \delta_{\lambda\mu}$.
2. For all sets of complex variables $X, Y$ so that $|xy| < 1$ for all $x \in X, y \in Y$,

$$\sum_{\lambda} u_{\lambda}(X)v_{\lambda}(Y) = \prod_{x \in X, y \in Y} (1 - xy)^{-1}.$$
Lemma 2.5 (Cauchy identities). Let $\mathcal{X}$ and $\mathcal{Y}$ be two sets of variables with elements in $\mathbb{C}$ so that the product of any element in $\mathcal{X}$ with any element in $\mathcal{Y}$ is strictly less than 1 in absolute value. The Cauchy identity states that

$$
\sum_{\lambda} s_\lambda(\mathcal{X}) s_\lambda(\mathcal{Y}) = \prod_{x \in \mathcal{X}, y \in \mathcal{Y}} (1 - xy)^{-1}.
$$

Furthermore, what we call the power sum version of the Cauchy identity states that

$$
\sum_{\lambda} s_\lambda(\mathcal{X}) s_\lambda(\mathcal{Y}) = \sum_{\mu} z_{\mu}^{-1} p_\mu(\mathcal{X}) p_\mu(\mathcal{Y})
$$

where the three sums range over all partitions.

In consequence, the Schur functions form an orthonormal basis for the ring of symmetric function, while the power sums satisfy $\langle p_\lambda, p_\mu \rangle = z_{\lambda\lambda}^0$ for all partitions $\lambda, \mu$. The following Lemma gives another point of view on the orthonormality of Schur functions.

Lemma 2.6 (Schur orthogonality, [Bum13]). Let $U(N)$ denote the unitary group of degree $N$. As $U(N)$ is compact it possesses a unique Haar measure normalized so that the volume of the entire group is 1. Whenever we integrate over $U(N)$, we integrate with respect to this measure. If for each matrix $g \in U(N)$ we write $R(g)$ for the multiset of its eigenvalues, then

$$
\int_{U(N)} s_\mu(R(g)) s_\nu(R(g)) dg = \begin{cases} 1 & \text{if } \mu = \nu \text{ and } l(\mu) \leq N, \\ 0 & \text{otherwise}. \end{cases}
$$

2.2.2. Specializations of the ring of symmetric functions. The definitions given in this paragraph are taken from [BC14, p. 259]. A specialization $\rho$ of the ring of symmetric functions is an algebra homomorphism from Sym to $\mathbb{C}$. We denote the application of $\rho$ to a symmetric function $f$ as $f(\rho)$. For two specializations $\rho_1$ and $\rho_2$ we define their union $\rho = \rho_1 \cup \rho_2$ as the specialization defined on power sum symmetric functions via

$$
\rho_k(\rho_1 \cup \rho_2) = \rho_k(\rho_1) + \rho_k(\rho_2)
$$

for all $k \geq 1$, and extended to all symmetric functions by the fact that the power sum symmetric functions form an algebraic basis of Sym. We note for later reference that

$$
z_{\lambda}^{-1} p_\lambda(\rho_1 \cup \rho_2) = \sum_{\mu, \nu} \rho_\mu(\rho_1) z_{\nu}^{-1} p_\nu(\rho_2)
$$

and

$$
s_\lambda(\rho_1 \cup \rho_2) = \sum_{\mu, \nu} c^\lambda_{\mu\nu} \rho_\mu(\rho_1) s_\nu(\rho_2).
$$

where $c^\lambda_{\mu\nu}$ are Littlewood-Richardson coefficients; their definition can be found in [Mac95, p. 142].

Definition 2.7. Let $\omega$ be the involution on the ring of symmetric functions given by $\omega(\xi_r) = h_r$. Recall that $\omega(p_n) = (-1)^{n-1} p_n$ and $\omega(s_\lambda) = s_\lambda$. We define the following two specializations:

$$
\rho^0_\lambda : \text{Sym} \to \mathbb{C}; f \mapsto f(\mathcal{X}) \quad \text{and} \quad \rho^1_\lambda : \text{Sym} \to \mathbb{C}; f \mapsto \omega(f)(\mathcal{X}).
$$

Borodin and Corwin call specializations of type $\rho^0$ and $\rho^1$ finite length specializations and finite length dual specializations, respectively.
2.2.3. Power sum operators. We define two types of power sum operators on the vector space Sym. For \( k \geq 1 \), the \( k \)-th product operator, which we denote \( p_k \) by a slight abuse of notation, maps the symmetric function \( f \) to the product \( p_k f \). In order to define the second type of operators, recall that any symmetric function \( f \) can be uniquely written as a polynomial in the power sums \( p_1, p_2, \ldots \). For \( k \geq 1 \), the \( k \)-th derivation operator maps \( f \in \text{Sym} \) to the formal derivative of this polynomial with respect to \( p_k \); we denote it by \( \frac{\partial}{\partial p_k} \). In analogy to power sums, we use the convention that the \( \lambda \)-th product/derivation operator is given by the corresponding compositions of the respective operators: for a partition \( \lambda \) of length \( n \), \( p_\lambda = p_{\lambda_1} \cdots p_{\lambda_n} \).

A definition of the two power sum operators as well as most of the properties described in the following two lemmas are given in [Mac95, p. 76].

Lemma 2.8. Let \( f, g \) be symmetric functions and \( k, l \) strictly positive integers.

1. The two power sum operators satisfy the following commutation relations:

\[
\frac{\partial}{\partial p_k} p_l f = \begin{cases} 
 p_l \frac{\partial}{\partial p_k} f & \text{if } l \neq k, \\
 p_k \frac{\partial}{\partial p_k} f + f & \text{if } l = k.
\end{cases}
\]

2. The product and derivation operators are almost adjoint with respect to the Hall inner product; more precisely, \( \langle k \frac{\partial}{\partial p_k} f, g \rangle = \langle f, p_k g \rangle \).

Proof. The commutation relations are a direct consequence of the product rule for the derivative. To show the second property, it is enough to consider \( f = p_\mu \) and \( g = p_\nu \) since the \( p_\lambda \), where \( \lambda \) ranges over all partitions, form a (linear) basis of Sym. The fact that this basis is orthogonal implies that both sides of the equation vanish unless \( \mu = \nu \cup (k) \). In this case,

\[
\langle k \frac{\partial}{\partial p_k} p_\mu, p_\nu \rangle = \langle km_k(\mu)p_\nu, p_\nu \rangle = km_k(\mu)z_\nu = z_\mu = \langle p_\mu, p_k p_\nu \rangle.
\]

Lemma 2.9. Let \( \mu \) and \( \nu \) be partitions. Then

\[
\frac{\partial}{\partial p_\mu} p_\nu = \begin{cases} 
 \prod_{i \geq 1} \frac{m_i(\nu)!}{m_i(\nu \setminus \mu)!} p_{\nu \setminus \mu} & \text{if } \mu \subset \nu \text{ as sequences} \\
0 & \text{otherwise}
\end{cases}
\]

as elements of the ring of symmetric functions.

Proof. Given that the \( l \)-th product and the \( k \)-th derivation operator commute whenever \( l \neq k \), we may write the left-hand side in (2.6) for any \( \mu \) as

\[
\frac{\partial}{\partial p_\mu} p_\nu = \prod_{i \geq 1} \frac{\partial}{\partial p_i} \left( \frac{m_i(\nu)!}{m_i(\nu \setminus \mu)!} p_{\nu \setminus \mu} \right).
\]

Handling each factor separately gives the right-hand side in (2.6).

Lemma 2.9 can be interpreted as moving the derivation operator to the right by means of the commutation relations in order to obtain a more concrete expression. We have taken this idea from [BCC17], in which it is used to simplify expressions involving another pair of operators that satisfy similar commutation relations.
2.3. **Littlewood-Schur functions.** Littlewood-Schur functions are a generalization of Schur functions, whose combinatorial definition appeared for the first time in the work of Littlewood [Lit36]. These functions were studied under a variety of different names: they are called hook Schur functions by Berele and Regev [BR87], supersymmetric polynomials by Nicoletti, Metropolis and Rota [NMR81], super-Schur functions by Brenti [Bre93], and Macdonald denotes them $s_{\lambda}(x/y)$ [Mac95, p. 58ff]. We follow Bump and Gamburd in calling them Littlewood-Schur functions and denoting them $LS_{\lambda}(X, Y)$ [BG06].

**Definition 2.10** (Littlewood-Schur functions). Let $X$ and $Y$ be two sets of variables. For any partition $\lambda$, define

$$LS_{\lambda}(X; Y) = \sum_{\mu, \nu} c_{\mu \nu}^\lambda s_{\mu}(X)s_{\nu}(Y)$$

where $c_{\mu \nu}^\lambda$ are Littlewood-Richardson coefficients; their definition can be found in [Mac95, p. 142].

The Littlewood-Schur function $LS_{\lambda}(X; Y)$ is a homogeneous polynomial of degree $|\lambda|$ in the variables $X \cup Y$. In contrast to the polynomials defined in the preceding section, Littlewood-Schur functions are not symmetric. However, this definition makes it apparent that $LS_{\lambda}(X; Y)$ is symmetric in both sets of variables separately. This combinatorial approach can also be used to prove the following formula that generalizes the Cauchy identity (i.e. the equality in (2.2)) as well as the dual Cauchy identity (which will not be relevant for our purposes).

**Proposition 2.11** (generalized Cauchy identity, [BRS5]). Let $S$, $T$, $U$ and $V$ be sets of variables with elements in $\mathbb{C}$. Suppose that all numbers of the form $uv$ or $st$ with $s \in S$, $t \in T$, $u \in U$ and $v \in V$ are strictly less than 1 in absolute value. If the same holds for all numbers of the form $ut$ or for all numbers of the form $sv$, then

$$\sum_{\lambda} LS_{\lambda}(S; U)LS_{\lambda}(T; V) = \prod_{s \in S} (1 + sv) \prod_{s \in S} (1 - st)^{-1} \prod_{u \in U} (1 - uv)^{-1} \prod_{t \in T} (1 + ut).$$

In particular, $\sum_{\lambda} |LS_{\lambda}(S; U)LS_{\lambda}(T; V)|$ possesses an upper bound that only depends on the absolute values of the elements in the four sets of variables in question.

The last sentence is a consequence of the fact that Littlewood-Richardson coefficients are non-negative, which entails that

$$|LS_{\lambda}(X; Y)| \leq \sum_{\mu, \nu} c_{\mu \nu}^\lambda |s_{\mu}(X)||s_{\nu}(Y)| \leq \sum_{\mu, \nu} c_{\mu \nu}^\lambda s_{\mu}(abs(X))s_{\nu}(abs(Y)) = LS_{\lambda}(abs(X); abs(Y)).$$

We remark that the second inequality follows immediately from the combinatorial definition for Schur functions.

**Remark 2.12.** The theory of specializations provides an alternative expression for $LS_{\lambda}(X; Y)$. Indeed,

$$LS_{\lambda}(X; Y) = \sum_{\mu, \nu} c_{\mu \nu}^\lambda s_{\mu}(X)s_{\nu}(Y) = \sum_{\mu, \nu} c_{\mu \nu}^\lambda s_{\mu}(\rho_X^\lambda) s_{\nu}(\rho_Y^\lambda) = s_{\lambda}(\rho_X^\lambda \cup \rho_Y^\lambda).$$

The last equality is due to (2.5). This perspective allows us to consider Littlewood-Schur functions a special type of Schur functions, which renders the following specialization intuitive: $LS_{\lambda}(X; \emptyset) = s_{\lambda}(X)$. 

---

MIXED RATIOS OF CHARACTERISTIC POLYNOMIALS
Yet another way to view Littlewood-Schur functions is given by Moens and Van der Jeugt’s determinantal formula. Their expression for $LS_{\lambda}(-X;Y)$ depends on the index of the partition $\lambda$.

**Definition 2.13 (index of a partition).** The $(m,n)$-index of a partition $\lambda$ is the largest (possibly negative) integer $k$ with the properties that $(m + 1 - k, n + 1 - k) \not\in \lambda$ and $k \leq \min\{m,n\}$.

If $(m,n) \not\in \lambda$, then $k$ is the side of the largest square with bottom-right corner $(m,n)$ that fits next to the diagram of the partition $\lambda$. If $(m,n) \in \lambda$, then $-k$ is the side of the largest square with top-left corner $(m,n)$ that fits inside the diagram of $\lambda$. Let us illustrate this by a sketch: the hatched area is the diagram of some partition $\lambda$.

![Diagram](image)

We remark that the definition given above is not equivalent to the definition of index used in [MdJ03]. Our notion has the advantage of being invariant under conjugation.

**Theorem 2.14 (determinantal formula for Littlewood-Schur functions, adapted from [MdJ03]).** Let $X$ and $Y$ be sets of variables of length $n$ and $m$, respectively, so that the elements of $X \cup Y$ are pairwise distinct. Let $\lambda$ be a partition with $(m,n)$-index $k$. If $k$ is negative, then $LS_{\lambda}(-X;Y) = 0$; otherwise,

$$LS_{\lambda}(-X;Y) = \varepsilon(\lambda) \frac{\Delta(Y;X)}{\Delta(X)\Delta(Y)} \times \det \begin{pmatrix} (x - y)^{-1} & x^{|\lambda_j + n - m - j|} & 0 \\ (y^{|\lambda_i| + m - n - i}) & 1 \leq j \leq m - k & 0 \\ 1 \leq i \leq n - k & y \in Y & 0 \end{pmatrix}$$

where $\varepsilon(\lambda) = (-1)^{|\lambda|-k}|(-1)^{mk}(-1)^{k(k-1)/2}$.

This theorem makes it easy to see that Littlewood-Schur functions behave well under transposition of their indexing partition; more concretely, for any partition $\lambda$, $LS_{\lambda}(X;Y) = LS_{\lambda}(Y;X)$. The following is another immediate consequence that will prove useful for the computations in Section 4. Corollary 2.15 is a special case of Berele and Regev’s factorization formula [BR77], which was originally derived without the help of the (more recent) determinantal formula.

**Corollary 2.15 ([BR77]).** Let $X$ and $Y$ be sets of variables with $n$ and $m$ elements, respectively, and let $\lambda$ be a partition with $(m,n)$-index $0$. If $l(\lambda) \leq n$, then

$$LS_{\lambda}(-X;Y) = \Delta(Y;X) \delta_{\lambda - (-m^\circ)}(-X).$$

**Proof.** First suppose that the elements in of $X \cup Y$ are pairwise distinct. Then

$$LS_{\lambda}(-X;Y) = \varepsilon(\lambda) \frac{\Delta(Y;X)}{\Delta(X)\Delta(Y)} \det \begin{pmatrix} (x - y)^{-1} & x^{|\lambda_j + n - m - j|} & 0 \\ (y^{|\lambda_i| + m - n - i}) & 1 \leq j \leq n & 0 \\ 1 \leq i \leq m & y \in Y & 0 \end{pmatrix}.$$
The Schur blocks in the bottom-left and the top-right corner of the matrix are squares. Hence,

$$LS_{\lambda}(-\mathcal{X}; \mathcal{Y}) = \varepsilon(\lambda) \frac{\Delta(\mathcal{Y}; \mathcal{X})}{\Delta(\mathcal{X}) \Delta(\mathcal{Y})} \times (-1)^{mn} \det \left( x^{(\lambda_i + n - m - j)} \right)_{1 \leq i, j \leq n} \det \left( y^{(\lambda_i + m - n - i)} \right)_{1 \leq i \leq m}. $$

On the one hand, the assumption that the length of $\lambda$ is less than $n$ implies that $\lambda'_i \leq n$ for $i \geq 1$. On the other hand, the assumption that the $(m, n)$-index of $\lambda$ is 0 implies that $\lambda'_i \geq n$ for $1 \leq i \leq m$. Therefore, the second determinant is actually a Vandermonde determinant, which cancels with $\Delta(\mathcal{Y})$. This allows us to conclude that

$$LS_{\lambda}(-\mathcal{X}; \mathcal{Y}) = \varepsilon(\lambda) \Delta(\mathcal{Y}; \mathcal{X}) (-1)^{mn} s_{\lambda' - (m n)}(\mathcal{X}) = \Delta(\mathcal{Y}; \mathcal{X}) s_{\lambda - (m n)}(-\mathcal{X})$$

since $\varepsilon(\lambda) = (-1)^{|\lambda|}$ by the assumptions on $\lambda$. If the elements of $\mathcal{X} \cup \mathcal{Y}$ are not pairwise distinct, the equality in (2.7) follows from the observation that both sides are polynomials in $\mathcal{X} \cup \mathcal{Y}$, which agree on infinitely many points. \hfill \Box

A less immediate but equally useful consequence of the determinantal formula is the simplest case of the first overlap identity:

**Lemma 2.16.** [Rie18b] Let $\mathcal{X}$ and $\mathcal{Y}$ be sets of variables with $n$ and $m$ elements, respectively, so that the elements of $\mathcal{X}$ are pairwise distinct. Let $\lambda$ be a partition with $(m, n)$-index $k$. If $0 \leq l \leq \min\{n - k, n\}$, then

$$LS_{\lambda}(-\mathcal{X}; \mathcal{Y}) = \sum_{S, T \subset \mathcal{X}, S \cup l_{n - l} T \supseteq \mathcal{X}} \frac{LS_{\lambda_{l(\mathcal{S})}} + (l - 1) \cdot LS_{\lambda_{l(\mathcal{S}) + 1, l + 2, \ldots}}(-\mathcal{S}; \mathcal{Y}) LS_{\mathcal{X} \setminus l_{n - l} \mathcal{T}}(-\mathcal{T}; \mathcal{Y})}{\Delta(\mathcal{T})}. $$

Recall that the subscripts in $S \cup l_{n - l} T$ indicate that $l(S) = l$ and $l(T) = n - l$, respectively.

3. On the Murnaghan-Nakayama rule

This section is dedicated to the Murnaghan-Nakayama rule. After stating the rule in its original form we present a few generalizations and variations, some of which are new and some are already known.

**Theorem 3.1** (Murnaghan-Nakayama rule, [Mur37, Nak40]). Let $\mu$ be a partition. For any strictly positive integer $k$,

$$p_{\lambda} s_{\mu} = \sum_{\lambda, \mu \subseteq \lambda} (-1)^{ht(\lambda \setminus \mu)} s_{\lambda}. $$

The following Corollary demonstrates what a powerful tool the theory of specializations of the symmetric group can be. It delivers the generalization of the Murnaghan-Nakayama rule to Littlewood-Schur almost for free. Neither the statement nor its proof is new but we did not manage to find an exact reference in the literature.

**Corollary 3.2** (Murnaghan-Nakayama rule for Littlewood-Schur functions). Let $\mu$ be a partition. For any strictly positive integer $k$,

$$(3.1) \quad LS_{\mu}(\mathcal{X}; \mathcal{Y}) \left[ p_{\lambda}(\mathcal{X}) + (-1)^{k - 1} p_{\lambda}(\mathcal{Y}) \right] = \sum_{\lambda, \mu \subseteq \lambda} (-1)^{ht(\lambda \setminus \mu)} LS_{\lambda}(\mathcal{X}; \mathcal{Y}).$$
Proof. View the Littlewood-Schur functions $LS\lambda(X; Y)$ on the right-hand side in (3.1) as specializations of $s_\lambda$, following Remark 2.12. Then use the Murnaghan-Nakayama rule to write the resulting sum as a specialization of $p_k s_\lambda$, which equals the left-hand side in (3.1) when written out. □

On the one hand, the left-hand side of the Murnaghan-Nakayama rule can be viewed as a product of a power sum and a Schur function. On the other hand, it can be seen as applying the product operator to a Schur function. The second point of view immediately leads to the question whether there is a similar "rule" for the derivation operator.

**Corollary 3.3** (dual Murnaghan-Nakayama rule). Let $\lambda$ be a partition. For any strictly positive integer $k$,

\[
(3.2) \quad k \frac{\partial}{\partial p_k} s_\lambda = \sum_{\mu : \mu \rightarrow k \lambda} (-1)^{ht(\lambda \setminus \mu)} s_\mu.
\]

The statement we have chosen to call the dual Murnaghan-Nakayama rule is standard but very elusive in the literature.

Proof. Exploit that Schur functions form an orthonormal basis of Sym to write the left-hand side in (3.2) as a linear combination of Schur functions:

\[
k \frac{\partial}{\partial p_k} s_\lambda = \sum_{\mu} \left( k \frac{\partial}{\partial p_k} s_\lambda, s_\mu \right) s_\mu.
\]

Using that derivation and product are almost adjoint (i.e. Lemma 2.8), switch operators, and then apply the Murnaghan-Nakayama rule:

\[
k \frac{\partial}{\partial p_k} s_\lambda = \sum_{\mu} \left( s_\lambda, p_k s_\mu \right) s_\mu = \sum_{\mu} \sum_{\nu : \nu \rightarrow k \mu} (-1)^{ht(\nu \setminus \mu)} \left( s_\lambda, s_\nu \right) s_\mu.
\]

The equality in (3.2) now follows from the orthonormality of Schur functions. □

The derivation operator thus allows us to give a neat expression for the signed sum of Schur functions associated to $\mu$ where $\mu$ ranges over all partitions so that $\lambda \setminus \mu$ is a $k$-ribbon. However, this expression can be difficult to work with because for a general symmetric polynomial $f(X)$ it is hard to give an explicit expression for $\frac{\partial}{\partial p_k} f(\rho_{\lambda}^n)$. The following proposition solves this problem under very specific assumptions, which will turn out to be sufficient for our purposes.

**Definition 3.4.** Let $X$ be a set of non-zero variables. For any partition $\lambda$, we define the $-\lambda$-th power sum of $X$ by

\[
p_{-\lambda}(X) = p_\lambda (X^{-1}).
\]

We remark that $p_{-\lambda}(X)$ is not a symmetric polynomial in $X$, but a symmetric Laurent polynomial.

**Proposition 3.5.** Let $X$ consist of $n$ non-zero variables and let $\mu$ be a partition of length $n$. For any integer $k$ with $1 \leq k \leq \mu_n$,

\[
(3.3) \quad s_\mu(X) p_{-k}(X) = \sum_{\lambda : \lambda \rightarrow k \mu} (-1)^{ht(\mu \setminus \lambda)} s_\lambda(X).
\]

Proof. Choose an integer $m$ such that $\mu$ is contained in the rectangle $(m^n)$. Let $\tilde{\mu}$ denote the $(m, n)$-complement of $\mu$. We reformulate the left-hand side of the equation in (3.3) as a function in the variables $X^{-1}$:

\[
s_\mu(X) p_{-k}(X) = c(X)^m s_{\tilde{\mu}} (X^{-1}) p_k (X^{-1}).
\]
This trick is an immediate consequence of the determinantal definition for Schur functions. Applying the Murnaghan-Nakayama rule yields
\[ s_\mu(\mathcal{X})p_{-k}(\mathcal{X}) = e(\mathcal{X})^m \sum_{\nu, \mu \vdash (m)} (-1)^{ht(\nu \setminus \mu)} s_\nu(\mathcal{X}^{-1}). \]

Notice that \( \nu_i \leq \mu_1 + k \leq \mu_1 + m = m. \) Hence, all partitions \( \nu \) that contribute to the sum are contained in \( (m^n) \). In consequence, the \((m,n)\)-complement of \( \nu \) is well defined. Replace \( \lambda \) by \( \tilde{\nu} \) in the summation index and reuse the trick to obtain
\[ s_\mu(\mathcal{X})p_{-k}(\mathcal{X}) = \sum_{\lambda, \mu \vdash (m,n)} (-1)^{ht(\tilde{\lambda} \setminus \lambda)} s_\lambda(\mathcal{X}). \]

It is easy to see that \( \tilde{\lambda} \setminus \tilde{\mu} \) is a \( k \)-ribbon if and only if \( \mu \setminus \lambda \) is. Together with the fact that the height remains unaltered this proves the claim. \( \square \)

**Corollary 3.6.** Let \( \mathcal{X} \) consist of \( n \) non-zero variables and let \( \mu \) be a partition of length at most \( n \). For any partition \( \lambda \) with \( |\lambda| \leq \mu_n \),
\[ \left( \prod_{l \geq 1} t^{m_l(\lambda)} \frac{\partial}{\partial t^{\lambda}} s_\mu \right)(\rho_X^\mu) = s_\mu(\mathcal{X})p_{-\lambda}(\mathcal{X}). \]  

**Proof.** Set \( l(\lambda) = l \). Repeated application of Corollary 3.3 to the left-hand side in (3.4) allows us to reformulate it as
\[ \sum_{\mu(1) \leq \mu \vdash (m,n)} (-1)^{ht(\mu \setminus \lambda)} s_\mu(\mathcal{X})p_{-\lambda}(\mathcal{X}). \]

By definition, \( s_\mu(\rho_X^\mu) = s_{\mu(\lambda)}(\mathcal{X}) \). Given that
\[ \lambda_i \leq |\lambda| - \lambda_{i-1} - \cdots - \lambda_2 \leq \mu_n - \lambda_{i-1} - \cdots - \lambda_1 \leq \mu_n \]
for all \( 1 \leq i \leq l \), repeatedly applying Proposition 3.5 results in the right-hand side of the equation in (3.4). \( \square \)

4. Averages of mixed ratios of characteristic polynomials

In this section we present a unified way to derive formulas for averages of products of ratios and/or logarithmic derivatives of characteristic polynomials over the group of unitary matrices \( U(N) \). Most of these formulas are not exact but contain an error that decreases exponentially as \( N \) goes to infinity.

4.1. **Tricks for bounding the error term.** As the heading suggests, this section is a collection of observations that will allow us to give asymptotic bounds for the various error terms. They are not particularly hard to prove or interesting in their own right.

**Lemma 4.1.** Fix a positive integer \( k \) and a partition \( \lambda \subset \langle m^n \rangle \). Then there are at most \( \min\{m,n\} \) partitions \( \mu \) such that
\[ \begin{cases} \lambda \setminus \mu \text{ is a } k\text{-ribbon.} \\ \mu \setminus \lambda \text{ is a } k\text{-ribbon and } \mu \subset \langle m^n \rangle. \end{cases} \]

**Proof.** Let \( \tilde{\lambda} \) denote the \((m,n)\)-complement of the partition \( \lambda \). For every partition \( \mu \subset \langle m^n \rangle \), \( \mu \setminus \lambda \) is a \( k \)-ribbon if and only if \( \tilde{\lambda} \setminus \tilde{\mu} \) is a \( k \)-ribbon, where \( \tilde{\mu} \) denotes the \((m,n)\)-complement of \( \mu \). Hence, it is sufficient to bound the number of partitions \( \mu \) such that \( \lambda \setminus \mu \) is a \( k \)-ribbon.

The condition that \( \mu \) be a partition implies that the top-right box of any ribbon \( \lambda \setminus \mu \) must not have any box to its right that is contained in \( \lambda \). This gives at most \( n \) possible positions for the top-right box, which entails that there are at most \( n \)
partitions $\mu$ such that $\lambda \setminus \mu$ is a $k$-ribbon. An analogous argument based on the bottom-left box of the ribbons bounds their number by $m$, thus concluding the proof.

Before going on to the next trick we recall the big-$O$ notation – primarily to fix notation. Given two functions $f$ and $g$ with domain $\mathcal{X}$, we write $f = O(g)$ if there exists a real constant $c(\mathcal{P})$ that may depend on the set of parameters $\mathcal{P}$ such that $|f(x)| \leq c(\mathcal{P})|g(x)|$ for all $x \in \mathcal{X}$. In this setting, we call $c(\mathcal{P})$ the implicit constant.

The following notation will also appear in the bounds for the error terms: the positive part of a real number $x$ is denoted by $x^+ = \max\{x, 0\}$.

**Lemma 4.2.** Fix a natural number $n$. For all square matrices $A$ whose size is less than $n$,

\[
\begin{align*}
(1) \quad & \det A = O_n \left( \prod_{j=1}^{m} \max_{1 \leq i \leq m} |a_{ij}| \right) \\
(2) \quad & \det A = O_n \left( \prod_{j=1}^{m} \max_{1 \leq j \leq m} |a_{ij}| \right)
\end{align*}
\]

where $m$ denotes the size of $A$.

**Proof.** Both statements follow directly from the Leibniz formula for determinants.

We only give a justification for the first statement, as they are exact analogues. We have that

\[
|\det A| = \left| \sum_{\sigma \in S_m} \varepsilon(\sigma) \prod_{j=1}^{m} a_{\sigma(j)j} \right| \leq \sum_{\sigma \in S_m} \prod_{j=1}^{m} \max_{1 \leq i \leq m} |a_{ij}| \leq n! \prod_{j=1}^{m} \max_{1 \leq i \leq m} |a_{ij}|. \quad \square
\]

We will use this lemma to infer asymptotic bounds for Schur and Littlewood-Schur functions based on their determinantal definitions.

**Lemma 4.3.** Fix a positive number $r$ and a set $\mathcal{X}$ of pairwise distinct variables.

1. If $\text{abs}(\mathcal{X}) \leq r$, then $s_{\lambda}(\mathcal{X}) = O_{\mathcal{X}} \left( r^{\lambda} \right)$ as a function of $\lambda$.
2. If $\mathcal{Y}$ is the subsequence of $\mathcal{X}$ that consists of the elements of absolute value greater than $1$, $s_{\lambda}(\mathcal{X}) = O_{\mathcal{X}} \left( |\mathcal{Y}|^{\lambda} \right)$ as a function of $\lambda$.

**Proof.** Set $n = l(\mathcal{X})$. To show the first bound, suppose that $\text{abs}(\mathcal{X}) \leq r$. By the determinantal definition for Schur functions,

\[
s_{\lambda}(\mathcal{X}) = O_{\mathcal{X}} \left( \det \left( x^{\lambda_j + n-j} \right)_{x \in \mathcal{X}, 1 \leq j \leq n} \right)
\]

as the denominator only depends on $\mathcal{X}$. Applying the first statement of Lemma 4.2 yields

\[
s_{\lambda}(\mathcal{X}) = O_{\mathcal{X}} \left( \prod_{j=1}^{n} \max_{x \in \mathcal{X}} |x^{\lambda_j + n-j}| \right)
\]

\[
= O_{\mathcal{X}} \left( r^{\lambda} \right).
\]

The second bound in this lemma is a consequence of the second statement of Lemma 4.2.

\[
s_{\lambda}(\mathcal{X}) = O_{\mathcal{X}} \left( \det \left( x^{\lambda_j + n-j} \right)_{x \in \mathcal{X}, 1 \leq j \leq n} \right)
\]

\[
= O_{\mathcal{X}} \left( \prod_{x \in \mathcal{X}, 1 \leq j \leq n} \max_{1 \leq j \leq n} |x^{\lambda_j + n-j}| \right).
\]
By assumption all variables \( x \in \mathcal{X} \) that are not elements of \( \mathcal{Y} \) are less than 1 in absolute value. Hence,

\[
\mathfrak{s}_\lambda(\mathcal{X}) = O_\mathcal{X} \left( \prod_{y \in \mathcal{Y}} \max_{1 \leq j \leq n} |y^{\lambda_j + n-j}| \right)
\]

\[
= O_\mathcal{X} \left( \prod_{y \in \mathcal{Y}} y^{\lambda_1} \right)
\]

This concludes the proof since \( \varepsilon(\mathcal{Y}) \) without index is our notation for the product. \( \square \)

The first bound in Lemma 4.3 can be viewed as a special case of Lemma 4.4 which gives an analogous statement for Littlewood-Schur functions.

**Lemma 4.4.** Fix a natural number \( l \), a positive number \( r \) and two sets of variables \( \mathcal{X} \) and \( \mathcal{Y} \) such that \( \text{abs}(\mathcal{X}) \leq r \) and the elements of \( \mathcal{X} \cup \mathcal{Y} \) are pairwise distinct. As a function of partitions \( \lambda \) with \( l(\lambda) \leq l \),

\[
LS_\lambda(-\mathcal{X} ; \mathcal{Y}) = O_\mathcal{P} \left( r^{|\lambda|} \right)
\]

where the implicit constant depends on \( \mathcal{P} = \{ \mathcal{X}, \mathcal{Y}, l \} \).

**Proof.** Set \( n = l(\mathcal{X}), m = l(\mathcal{Y}) \) and denote the \((m,n)\)-index of \( \lambda \) by \( k \). We remark that \( k \) depends on \( \lambda \), while \( m \) and \( n \) are constants. The determinantal formula for Littlewood-Schur functions (i.e. Theorem 2.14) entails that if \( k \) is non-negative

\[
LS_\lambda(-\mathcal{X} ; \mathcal{Y}) = O_{\mathcal{X}, \mathcal{Y}} \left( \det \left( \begin{pmatrix} \frac{(y - x)^{-1}}{y \in \mathcal{Y}}_{x \in \mathcal{X}} \frac{x^{\lambda_j + n - m - j}}{1 \leq j \leq n - k} \left( \begin{pmatrix} 0 \end{pmatrix}_{y \in \mathcal{Y}} \right) \right) \right)
\]

otherwise, the Littlewood-Schur function \( LS_\lambda(-\mathcal{X} ; \mathcal{Y}) \) vanishes, allowing us to ignore the case \( k < 0 \). Let us call this matrix \( A \). As the size of \( A \) is \( m + n - k \leq m + n \) for all partitions \( \lambda \), Lemma 4.3 states that

\[
LS_\lambda(-\mathcal{X} ; \mathcal{Y}) = O_{\mathcal{X}, \mathcal{Y}} \prod_{j=1}^{\max_{1 \leq i \leq m + n - k} |a_{ij}|}
\]

The condition that \( l(\lambda) \leq l \) is equivalent to \( \lambda^i \leq l \), and thus implies that \( \lambda^i \leq l \) for all \( i \). Hence, the \( m \) first columns of \( A \) make no asymptotically relevant contribution to the bound. Therefore,

\[
LS_\lambda(-\mathcal{X} ; \mathcal{Y}) = O_{\mathcal{X}, \mathcal{Y}, l} \left( \prod_{j=1}^{n-k} \max_{x \in \mathcal{X}} |x^{\lambda_j + n - m - j}| \right) = O_{\mathcal{X}, \mathcal{Y}, l} \left( r^{|\lambda|-m-k} \right)
\]

By the definition of index, \( \lambda_i \leq m-k \) for all indices \( i > n-k \). Combined with the condition that \( l(\lambda) \leq l \), we infer that

\[
|\lambda_{n-k}| \leq |\lambda| \leq |\lambda_{n-k}| + (l - (m-k))(m-k) \leq |\lambda_{n-k}| + lm.
\]

Therefore, if \( \begin{cases} r \geq 1, & \text{then } r^{|\lambda_{n-k}|} \leq O_{l,m,r} \left( r^{|\lambda|} \right) \end{cases} \)

If we drop the condition that the variables in \( \mathcal{X} \cup \mathcal{Y} \) be pairwise distinct, we can no longer use Lemma 4.3 to obtain an asymptotic bound on \( LS_\lambda(\mathcal{X} ; \mathcal{Y}) \), given that the implicit constant might grow arbitrarily large whenever elements of \( \mathcal{X} \cup \mathcal{Y} \) converge towards each other. The following lemmas, which are based on the combinatorial definitions for Schur and Littlewood-Schur functions, provide bounds that do not
depend on $X$ and $Y$. In particular, the variables need not to be pairwise distinct. However, the bounds based on the combinatorial definitions are not as good as the bounds based on the determinantal definitions.

**Lemma 4.5.** Fix a positive number $r$ and a set of variables $X$ such that $\text{abs}(X) \leq r$. As a function of $\lambda$,

$$s_\lambda(X) = O \left( |\lambda|^{t(X)^2} r^{|\lambda|} \right).$$

**Proof.** Owing to the combinatorial definition for Schur functions (which can be found in [Sag01]),

$$|s_\lambda(X)| \leq \sum_T r^{|\lambda|}$$

where the sum runs over all semistandard $\lambda$-tableaux $T$ whose entries do not exceed $l(X)$. Hence, it suffices to bound the number of tableaux that contribute to the sum. Given that the rows/columns of $T$ are weakly/strongly increasing, there are at most $\lambda_1 \cdot \cdots \cdot \lambda_l \leq |\lambda|^l$ possible choices for the boxes of $T$ that contain the positive integer $i$. Multiplying over all $1 \leq i \leq l(X)$ gives the desired bound.

**Lemma 4.6.** Fix a natural number $l$, a positive number $r$ and two sets of variables $X$ and $Y$ such that $\text{abs}(X) \leq r$. As a function of partitions $\lambda$ with $l(\lambda) \leq l$,

$$LS_\lambda(X; Y) = O_{\mathcal{P}} \left( |\lambda|^{t(X)^2+r^2} r^{|\lambda|} \right)$$

where the implicit constant depends on $\mathcal{P} = \{l, r, l(Y), \max(\text{abs}(Y))\}$.

**Proof.** Applying Lemma 4.5 to the combinatorial definition for Littlewood-Schur functions (i.e. Definition 2.10) gives us

$$LS_\lambda(X; Y) = \sum_{\mu,\nu, \nu \leq l(Y)} c^\lambda_{\mu\nu} O \left( |\mu|^{t(X)^2} r^{|\mu|} |\nu|^{t(Y)^2} R^{|\nu|} \right)$$

where $R = \max(\text{abs}(Y))$. The next step in this proof relies on some basic properties of Littlewood-Richardson coefficients, a justification of which can be found in [Rie18a, p. 54-55]. Since $c^\lambda_{\mu\nu}$ vanishes unless $\nu$ is a subset of $\lambda$, only partitions $\nu$ contained in the rectangle $\langle l(Y)^l \rangle$ appear in the sum. Hence, the fact that $|\nu| + |\mu| = |\lambda|$ for all partitions that contribute to the sum entails that $|\mu| \leq |\lambda| \leq |\mu| + l(Y)l$, which allows us to replace $|\mu|$ by $|\lambda|$. Keeping track of the fact that $\mu \subset \lambda$ whenever $c^\lambda_{\mu\nu} \neq 0$, we thus have that

$$LS_\lambda(X; Y) = O_{l, r, l(Y), R} \left( |\lambda|^{t(X)^2} r^{|\lambda|} \right) \sum_{\mu,\nu, \nu \leq l(Y)} \sum_{\mu \subset \lambda, |\mu|+|\nu|=|\lambda|} c^\lambda_{\mu\nu}.$$

According to the Littlewood-Richardson rule (stated and proved in [Sag01]), $c^\lambda_{\mu\nu}$ can be bounded by the number of skew semistandard $\lambda \setminus \nu$-tableaux $T$ with weight $\mu$. For each positive integer $i$, there are at most $\lambda_1 \cdot \cdots \cdot \lambda_i \leq |\lambda|^i$ ways to choose the boxes of $T$ that contain $i$ (given that $l(\lambda) \leq l$). The condition that $l(\mu) \leq l$ thus implies that $c^\lambda_{\mu\nu} \leq |\lambda|^l$.

The bound stated above now follows from the observation that the number of pairs $\mu, \nu$ to sum over is less than $l(Y)^l \times (l(Y)l)^l$. Indeed, there are less than $l(Y)^l$ partitions $\nu$ that are contained in the rectangle $\langle l(Y)^l \rangle$. Fixing a partition $\nu$, the conditions that $\mu \subset \lambda$ and $|\mu| = |\lambda| - |\nu|$ allow us to infer that there are at most $|\mu| \leq l(Y)l$ ways to choose a part $\mu_i$ for $1 \leq i \leq l$. 

$\square$
4.2. The recipe. Before stating our recipe for computing averages of mixed ratios of characteristic polynomials over the group of unitary matrices, we quickly recall the notion of a characteristic polynomial from linear algebra. The fact that we only consider unitary — and thus invertible — matrices allows us to work with a variant of the standard definition, which possesses close conjectural ties to the theory of \( L \)-functions.

Definition 4.7 (characteristic polynomial). The characteristic polynomial of a unitary matrix \( g \in U(N) \) is given by \( \chi_g(z) = \det (I - zg^{-1}) \) where \( I \) is the identity matrix.

Recipe (ratios and logarithmic derivatives). Let \( A, B, C, D, E \) and \( F \) be sets of non-zero variables so that the four latter only contain elements that are strictly less than 1 in absolute value. If \( |l(D)| \leq |l(A)| \) and the elements of \( A \cup B^{-1} \) are pairwise distinct, then

\[
\int_{U(N)} \frac{\prod_{\alpha \in A} \chi_{g}(\alpha) \prod_{\beta \in B} \chi_{g^{-1}}(\beta) \prod_{\gamma \in C} \chi_{g^{-1}}(\gamma)}{\prod_{\varepsilon \in E} \chi_{g}(\varepsilon) \prod_{\varphi \in F} \chi_{g^{-1}}(\varphi)} dg = (-1)^{|E|+|F|} \epsilon(-B)^N \sum_{S,T \subset A \cup B^{-1}} \sum_{S \cup \{g\}, I \subset T \setminus \{g\}} \Delta(D;S) \Delta(T;S)
\]

\[\times \sum_{E', s \geq l(E') \in E} \sum_{q+n \geq l(E') \in E} \left( \sum_{\chi_{\psi^{-1}(E')} \subset I(E')} \sum_{l(\chi_{\psi^{-1}(E')}) = q+n} m_{\chi_{\psi^{-1}(E')}} \right)\]

\[\times \sum_{\psi^{-1}(E')} \sum_{l(\psi^{-1}(E'))} \left( E' \right) \sum_{\omega^{-1}(\psi^{-1}(E'))} \left( F \right) m_{\omega^{-1}(\psi^{-1}(E'))}
\]

\[\times \sum_{\lambda \in \omega_{\xi^{-1}(\psi^{-1}(E'))}} \sum_{\omega_{\xi^{-1}(\psi^{-1}(E'))} \subset D} \sum_{\omega_{\xi^{-1}(\psi^{-1}(E'))} \subset D} \left( \rho_{E' \cup \psi^{-1}(E')} \prod_{i \geq 1} \frac{\prod_{m_i(\omega) m_i(\psi^{-1}(E'))}!}{m_i(\xi)!!} \right) p_{\zeta}(\psi^{-1}(E'))
\]

\[+ \text{error}.
\]

An asymptotic bound for the error is given in (4.9) on page 17.

We call this statement a recipe rather than a theorem because we are not able to give a neat bound for the error term. In particular, the error term might be larger than the main term. However, when some of the sets of variables are empty the error term becomes more tractable, which will allow us to prove the results presented in Section 4.3. In this sense the recipe provides a unified way of showing formulas for products of ratios and/or logarithmic derivatives.

On a more technical note, observe that for \( g \in U(N) \) and \( z \in \mathbb{C} \setminus \{0\} \),

\[
\chi_g(z) = \det (I - zg^{-1}) = \det (-zg^{-1}) \det (I - z^{-1}g) = (-z)^N \epsilon(\mathcal{R}(g)) \chi_{g^{-1}} (z^{-1})
\]

where \( \mathcal{R}(g) \) is the multiset of eigenvalues of \( g \). Considering the integrand on the left-hand side of (4.1), we see that this observation allows us to replace \( \chi_g(\delta) \) by \( \chi_{g^{-1}}(\gamma) \) with \( \gamma = \delta^{-1} \) at the cost of a factor which is easy to handle. Hence, for any \( r \in \mathbb{R} \setminus \{0\} \), the condition that \( \text{abs}(C), \text{abs}(D) \leq r \) is essentially equivalent to the condition that all elements of \( C \cup D \) are less than \( r \) or greater than \( r^{-1} \) in absolute value. Moreover, the same holds for the sets of variables \( A \) and \( B \). In particular, prerequisites of the type \( \text{abs}(A), \text{abs}(B) \leq 1 \) are essentially empty conditions. Of
Proof. This proof is based on the observation that the integrand on the left-hand side is symmetric in the eigenvalues of the unitary matrix \( g \), say \( \mathcal{R}(g) \), as well as in their complex conjugates \( \overline{\mathcal{R}(g)} \). It is thus (at least theoretically) possible to express the integrand as an infinite linear combination of products of Schur functions of the form
\[
\overline{s}_\lambda(\mathcal{R}(g))s_\kappa(\mathcal{R}(g)).
\]
Once the coefficients of this linear combination are known, Schur orthogonality immediately gives an expression for the integral on the left-hand side. At the cost of an error (which is ultimately due to the fact that Schur functions are only essentially orthonormal), we then simplify this expression by applying results presented in the preceding sections.

Given that \( \text{abs}(\mathcal{C}) \), \( \text{abs}(\mathcal{D}) < 1 \) elementary linear algebra manipulations together with the generalized Cauchy identity (i.e. Proposition 2.11) give the following expression for the ratios on the left-hand side in [111]:
\[
\prod_{\alpha \in A} \chi_\alpha(\lambda) \prod_{\beta \in B} \chi_\beta^{-1}(\beta) \prod_{\gamma \in C} \chi_\gamma^{-1}(\gamma)
\]
\[
= \prod_{\alpha \in A} \det(1 - \alpha g^{-1}) \prod_{\beta \in B} \det(g) \det(-\beta I) \det(-\beta^{-1} g^{-1} + I)
\]
\[
\times \prod_{\gamma \in C} \det(1 - \gamma g^{-1})^{-1}
\]
\[
= e(-B)^N \prod_{x \in A \cup B^{-1}} (1 - x \overline{p}) \prod_{\rho \in \mathcal{R}(g)} (1 - p)^{-1} \prod_{\gamma \in C} (1 - \gamma p)^{-1}
\]
\[
= e(-B)^N \prod_{\rho \in \mathcal{R}(g)} \left[ \sum_\lambda L \chi_\lambda \left( - (A \cup B^{-1}) \cap \mathcal{D} \right) \overline{s}_\lambda(\mathcal{R}(g)) \right] \sum_\kappa s_\kappa(\mathcal{C}) s_\kappa(\mathcal{R}(g)).
\]

In their combinatorial proof of a formula for averages of ratios of characteristic polynomials over the unitary group, Bump and Gamburd use the same algebraic manipulations and similar Cauchy identities to write ratios of characteristic polynomials in terms of Schur functions [BG06, p. 245-246]. Furthermore, Dehaye remarks that for \( \varepsilon \in \mathbb{C} \) with \( |\varepsilon| < 1 \) [Deit08],
\[
\chi_\rho(\varepsilon) = \sum_{\rho \in \mathcal{R}(g)} \varepsilon^p \overline{p} = - \sum_{m=1}^{\varepsilon m-1} p_m(\mathcal{R}(g)).
\]
Setting \( e = l(\mathcal{E}) \), \( f = l(\mathcal{F}) \) and \( \mathcal{E} = (\varepsilon_1, \ldots, \varepsilon_e) \), \( \mathcal{F} = (\varphi_1, \ldots, \varphi_f) \), we may thus reformulate the integral on the left-hand side in (4.1) as
\[
\text{LHS} = e(-B)^N \sum_\lambda L \chi_\lambda \left( - (A \cup B^{-1}) \cap \mathcal{D} \right) \sum_\kappa s_\kappa(\mathcal{C})
\]
\[
\times \left( -1 \right)^e \varepsilon^{e+f} \sum_{m_1, \ldots, m_e \geq 1} \left( \prod_{i=1}^{e} \varepsilon_{m_i}^{m_i-1} \right) \sum_{n_1, \ldots, n_f \geq 1} \left( \prod_{j=1}^{f} \varphi_{n_j}^{n_j-1} \right)
\]
\[
\times \int_{U(1)} e(\mathcal{R}(g))^{f(\mathcal{B})} \overline{s}_\lambda(\mathcal{R}(g)) s_\kappa(\mathcal{R}(g)) \prod_{i=1}^{e} p_{m_i}(\mathcal{R}(g)) \prod_{j=1}^{f} p_{n_j}(\mathcal{R}(g)) dg.
\]
In order to write the integrand as a linear combination of products of Schur functions, we repeatedly apply the Murnaghan-Nakayama rule (i.e. Theorem 3.1):

\[
\text{LHS} = e(-B)^N \sum_{\lambda} LS_{\lambda'} \left( - (A \cup B^{-1}) ; D \right) \sum_{\kappa} s_\kappa(C)
\]

\[
\times (-1)^{e+f} \sum_{m_1, \ldots, m_e \geq 1} \left( \prod_{i=1}^e \varphi_j^{-1} \right) \sum_{n_1, \ldots, n_f \geq 1} \left( \prod_{j=1}^f \varphi_j^{-1} \right) \sum_{\lambda \geq \kappa} (-1)^{ht(\lambda(\nu) \setminus \lambda)} \sum_{\nu = \lambda(\mu) \setminus \lambda(\nu)} (-1)^{ht(\lambda(\nu) \setminus \lambda(\mu))} \sum_{l(\nu) \leq N} \left( \prod_{i=1}^e \varphi_j^{-1} \right) \sum_{l(\nu) \leq N} \left( \prod_{j=1}^f \varphi_j^{-1} \right)
\]

\[
\times \int_{U(N)} e(\mathcal{R}(g))^{(l(B))^N} s_{\lambda(\nu)}(\mathcal{R}(g)) s_{\kappa(\lambda)}(\mathcal{R}(g)) dg.
\]

It is a straightforward linear algebra exercise to show that for sequences $\mathcal{X}$ of length $N$, $e(\mathcal{X})^{l(B)} s_\kappa(\mathcal{X}) = s_{\kappa + (A)^N}(\mathcal{X})$. Hence, Schur orthogonality (i.e. Lemma 2.6) allows us to compute the integral. In practice, we just introduce the dummy variable $\pi$ to ensure that $\kappa(\nu) + (l(B))^N = \lambda(\mu)$, and that the length of the partition does not exceed $N$:

\[
\text{LHS} = e(-B)^N \sum_{\lambda} LS_{\lambda'} \left( - (A \cup B^{-1}) ; D \right) \sum_{\kappa} s_\kappa(C)
\]

\[
\times (-1)^{e+f} \sum_{m_1, \ldots, m_e \geq 1} \left( \prod_{i=1}^e \varphi_j^{-1} \right) \sum_{n_1, \ldots, n_f \geq 1} \left( \prod_{j=1}^f \varphi_j^{-1} \right) \sum_{\lambda \geq \kappa} (-1)^{ht(\lambda(\nu) \setminus \lambda)} \sum_{\nu = \lambda(\mu) \setminus \lambda(\nu)} (-1)^{ht(\lambda(\nu) \setminus \lambda(\mu))} \sum_{l(\nu) \leq N} \left( \prod_{i=1}^e \varphi_j^{-1} \right) \sum_{l(\nu) \leq N} \left( \prod_{j=1}^f \varphi_j^{-1} \right)
\]

\[
\times \left[ \sum_{\lambda(\mu) = \lambda(\nu) \setminus \lambda(\mu)} (-1)^{ht(\lambda(\nu) \setminus \lambda)} \right] \left[ \sum_{\lambda(\nu) = \lambda(\mu) \setminus \lambda(\nu)} (-1)^{ht(\lambda(\nu) \setminus \lambda(\nu))} \right].
\]

The remainder of the proof is dedicated to simplifying the expression above, which seems to come at the cost of introducing an error term. We will replace $\lambda(i)$ by the following sum of partitions: $\lambda(i) = \nu(i) + \mu(i)$ where $\nu(i)$ is the intersection of $\langle l(B)^N \rangle$ and $\lambda(i)$. Notice that every $m_i$-ribbon $\lambda(i) \setminus \lambda(i-1)$ that appears in the expression above can be cut into two ribbons: a $q_i$-ribbon $\nu(i) \setminus \nu(i-1)$ that is a subset of the rectangle $\langle l(B)^N \rangle$, and a $p_i$-ribbon $\mu(i) \setminus \mu(i-1)$ whose boxes lie strictly to the right of the vertical line given by $x = l(B)$.

For the main term, we restrict ourselves to ribbon sizes that satisfy

\[
q_1 + \cdots + q_e + n_1 + \cdots + n_f \leq N - l(C).
\]

This restriction leads to a number of simplifications: Given that only partitions $\kappa$ of length less than $l(C)$ contribute to the sum (since otherwise $s_\kappa(C)$ vanishes), the fact that $n_1 + \cdots + n_f + l(C) \leq N$ entails that $l(\pi) \leq N$. Moreover, the restriction implies that for every $m_i$-ribbon that appears in the main term, $p_i = 0$ or $q_i = 0$. This last simplification is probably best explained by means of a sketch. The following drawing depicts possible Ferrers diagrams of the partition $\langle l(B)^N \rangle + \pi$ (white) and its subset $\lambda$ (hatched).
By definition of the \( q_i \), \( |\lambda \cap (l(B))^N| = Nl(B) - q_1 - \cdots - q_c \), which implies that \( N - \lambda_B \leq q_1 + \cdots + q_c \), as indicated on the sketch. In addition, we have already seen that \( l(\pi) \leq l(B) + n_1 + \cdots + n_f \). Therefore, the condition given in [4.3] implies that the box with coordinates \((l(B), l(\pi))\) must be contained in \( \lambda \). The box in question is marked by a slightly darker pattern. We thus conclude that every \( m_i \)-ribbon \( \lambda \setminus \lambda^{(i-1)} \) lies either to the left or strictly to the right of this box, which is the graphical way of saying that either \( m_i = q_i \) or \( m_i = p_i \).

In sum, the main term is equal to

\[
\text{main} = \epsilon(-B)^N \sum_{\mu, \nu} \sum_{\nu' \cup \mu'} \mathcal{L}S_{\nu' \cup \mu'}(- (A \cup B^{-1}) ; D) \sum_{\kappa} g_{\kappa}(C)
\]

\[
\times (-1)^{s+f} \sum_{g, h \geq 0} \sum_{G, H \subseteq [e]} \sum_{p_1, \ldots, p_g \geq 1} \left( \prod_{i=1}^{g} p_i^{-1} \right)
\]

\[
\times \sum_{q, n \geq 0} \sum_{q_1 + \cdots + q_c \geq 1} \left( \prod_{i=1}^{h} q_i^{-1} \right) \sum_{n_1 + \cdots + n_f \geq n} \left( \prod_{j=1}^{f} n_j^{-1} \right) \sum_{\pi}
\]

\[
\sum_{\mu^{(1)}, \ldots, \mu^{(h)}} \left( -1 \right)^{ht(\nu^{(h)} \setminus \nu)} \sum_{\mu^{(1)} \mu^{(2)} \cdots \mu^{(h)}} \left( -1 \right)^{ht(\mu^{(1)} \setminus \mu)}
\]

\[
\sum_{\kappa^{(1)}, \ldots, \kappa^{(f)}} \left( -1 \right)^{ht(\kappa^{(f)} \setminus \kappa)}
\]

First notice that under the assumption that the condition given in [4.3] is satisfied, the restriction to pairs of partitions \( \mu, \nu \) so that \( \nu' \cup \mu' \) is a partition is actually superfluous. Indeed,

\[
\nu_{(B)} \geq N - q_1 - \cdots - q_h \geq l(C) + n_1 + \cdots + n_f \geq l(\pi) \geq l(\mu) = \mu'.
\]
We use Lemma 2.13 to write \( LS_{\nu' \cup \mu'} \left( - \left( A \cup B^{-1} \right); D \right) \) as a sum of products of Littlewood-Schur functions that depend on \( \nu \) or \( \mu \) but not on both. This is permissible given that the elements of \( A \cup B^{-1} \) are pairwise distinct and that \( l(D) \leq l(A) \), which implies that the \( (l(D), l(A) + l(B)) \)-index of any partition is less than \( l(A) \). More concretely, we obtain
\[
LS_{\nu' \cup \mu'} \left( - \left( A \cup B^{-1} \right); D \right) = \sum_{S,T \subseteq A \cup B^{-1}} \frac{LS_{\nu' + \{l(A); \{S; D\} \} \left( -S; D \right) \cdot LS_{\mu'} \left( -T; D \right)}}{\Delta(T; S)}.
\]
Again due to the fact that \( l(D) \leq l(A) \), Corollary 2.15 states that
\[
LS_{\nu' + \{l(A); \{S; D\} \} \left( -S; D \right) = \Delta(D; S) \cdot \varepsilon(-S)^{l(A) - l(D)} \frac{\Delta(D; S)}{\Delta(T; S)} \cdot \varepsilon(-S)^{l(A) - l(D)} \cdot \varepsilon(-S).
\]
Hence, rearranging the various sums in the main term yields
\[
\text{main} = \varepsilon(-B)^N \sum_{S,T \subseteq A \cup B^{-1}} \varepsilon(-S)^{l(A) - l(D)} \frac{\Delta(D; S)}{\Delta(T; S)} \cdot \varepsilon(-S)^{l(A) - l(D)} \cdot \varepsilon(-S).
\]
\[
\times \left( -1 \right)^{\rho + f} \sum_{g,h \geq 0; \, G,H \subseteq [e]} \sum_{p_1, \ldots, p_g \geq 1} \left( \prod_{i=1}^{g} \frac{1}{G_i} \right) \sum_{g + h = e} \sum_{f \geq 1} \left( \prod_{j=1}^{f} \frac{\varphi_j^{n_j-1}}{n_j} \right)
\]
\[
\times \sum_{\nu' \in \{1, \ldots, \nu(h); \nu(h) \nu(h)^{2}, \ldots, \nu(h)^{k(h)}\}} \varepsilon(-S) \cdot \left( -1 \right)^{\delta(S \setminus \nu(h))} \cdot \left( -1 \right)^{\delta(\mu(h) \setminus \mu)}
\]
\[
\times \sum_{\mu} LS_{\mu'} \left( -T; D \right) \cdot \sum_{\pi \in \{1, \ldots, \mu(h)\} \setminus \mu(1) \ldots \mu(h)} \left( -1 \right)^{\delta(\kappa(f) \setminus \kappa)}.
\]
Let us now focus on the three sums over ribbons:
\[
\text{ribbon}(q_1, \ldots, q_h) = \sum_{\nu} \varepsilon(-S) \cdot \sum_{\nu \subseteq \nu(1) \ldots \nu(h); \nu(h) = \{i \in (B)\}^{N}} \left( -1 \right)^{\delta(\nu(h) \setminus \nu)} \cdot \varepsilon(-S) \cdot \sum_{\mu} LS_{\mu'} \left( -T; D \right) \cdot \sum_{\pi \in \{1, \ldots, \mu(h)\} \setminus \mu(1) \ldots \mu(h)} \left( -1 \right)^{\delta(\kappa(f) \setminus \kappa)}.
\]
The equality in (2.1) allows us to get rid of the conjugation in the index of the Schur function. Since \( (l(B))^N = \langle N_l(B) \rangle \),

\[
\text{ribbon}(q_1, \ldots, q_h) = (-1)^{q-h} \sum_{\mu^{(0)} \mu^{(1)} \ldots \mu^{(h-1)}} (-1)^{ht}(\langle N^{(m)} \rangle^{(0)})_{\mu^{(0)}}(\rho_+^{-})
\]

Repeatedly applying Corollary 3.6 results in

\[
\text{ribbon}(q_1, \ldots, q_h) = (-1)^{q-h} \left[ \frac{\partial}{\partial p_{q_1}} \ldots \frac{\partial}{\partial p_{q_h}} \rho \langle N^{(m)} \rangle \right](\rho_+^{-}).
\]

The theory of operators makes it apparent that ribbon\((q_1, \ldots, q_h)\) is independent of the order of the \(q_i\). Without loss of generality, we may thus assume that \((q_1, \ldots, q_h)\) is a partition of length \(h\), say \(\chi\). In this notation the preceding equality reads

\[
\text{ribbon}(\chi) = (-1)^{|\chi|-h} \left[ \prod_{i \geq 1} c_{\mu_i}(\chi) \frac{\partial}{\partial p_{\chi}} \rho \langle N^{(m)} \rangle \right](\rho_+^{-})
\]

Given that \(|\chi| = q \leq N\) and \(l(S) = l(B)\), Corollary 3.6 states that

\[
\text{ribbon}(\chi) = (-1)^{|\chi|-h} \rho \langle N^{(m)} \rangle(-S)\rho - \chi(-S) = (-1)^{|\chi|-h} e(-S)^N \rho - \chi(-S).
\]

The remaining two sums over ribbons that appear in (4.4) can in fact be viewed as one sum:

\[
\text{ribbon}(p_1, \ldots, p_g; n_1, \ldots, n_f) \equiv \sum_{\mu} L_S \mu'(-T; D) \sum_{\pi} \sum_{\mu^{(1)} \ldots \mu^{(g)}} (-1)^{ht}(\mu^{(g)} \backslash \mu)
\]

\[
\times \sum_{\kappa} \delta_\kappa(C) \sum_{\kappa^{(1)} \ldots \kappa^{(f)}} (-1)^{ht}(\kappa^{(f)} \backslash \kappa)
\]

\[
\sum_{\mu} L_S \mu'(-T; D)
\]

\[
\times \sum_{\mu^{(1)} \ldots \mu^{(g)}; \kappa^{(1)} \ldots \kappa^{(f)}} (-1)^{ht}(\mu^{(g)} \backslash \mu) (-1)^{ht}(\kappa^{(f)} \backslash \kappa) \delta_\kappa(\rho_+^C).
\]

Repeatedly applying the Murnaghan-Nakayama rule and its dual (i.e. Theorem 3.1 and Corollary 3.3) gives

\[
\text{ribbon}(p_1, \ldots, p_g; n_1, \ldots, n_f) = \sum_{\mu} \delta_\mu \left( \rho_+^B \cup \rho_+^C \right)
\]

\[
\times \left[ n_1 \frac{\partial}{\partial p_{n_1}} \ldots n_f \frac{\partial}{\partial p_{n_f}} p_{p_g} \ldots p_{p_1} \delta_\mu \right](\rho_+^C)
\]
where we view the Littlewood-Schur function as a specialization of a Schur function, following Remark 2.12. As above the theory of operators makes it obvious that ribbon\((p_1, \ldots, p_g; n_1, \ldots, n_f)\) is symmetric in both \((p_1, \ldots, p_g)\) and \((n_1, \ldots, n_f)\), which we thus replace by the partitions \(\psi\) and \(\omega\) of lengths \(g\) and \(f\), respectively. This substitution yields

\[
\text{ribbon}(\psi; \omega) = \left[ \prod_{i \geq 1} i^{m_i(\omega)} \frac{\partial}{\partial p_{\omega}} \sum_{\mu} g_{\mu} \left( \rho^\beta_{\mu} \cup \rho^\alpha_{\mu} \right) \right] (\rho^\omega). 
\]

Due to the power sum version of the Cauchy identity given in (2.3), this is equal to

\[
\text{ribbon}(\psi; \omega) = \prod_{i \geq 1} i^{m_i(\omega)} \sum_{\lambda} \sum_{\psi \cup \lambda} \frac{1}{\lambda!} \left( \rho^\beta_{\lambda} \cup \rho^\alpha_{\lambda} \right) \left[ \frac{\partial}{\partial p_{\psi \cup \lambda}} \right] (\rho^\omega). 
\]

According to Lemma 2.9, this is equal to

\[
\text{ribbon}(\psi; \omega) = \prod_{i \geq 1} i^{m_i(\omega)} \sum_{\lambda, \xi} \sum_{\psi \cup \lambda} \frac{1}{\lambda!} \left( \rho^\beta_{\lambda} \cup \rho^\alpha_{\lambda} \right) \left[ \frac{\partial}{\partial p_{\psi \cup \lambda}} \right] (\rho^\omega). 
\]

This substitution yields

\[
\text{ribbon}(\psi; \omega) = \sum_{\lambda} \sum_{\xi} \sum_{\psi \cup \lambda} \frac{1}{\lambda!} \left( \rho^\beta_{\lambda} \cup \rho^\alpha_{\lambda} \right) \left[ \frac{\partial}{\partial p_{\psi \cup \lambda}} \right] (\rho^\omega). 
\]

Incorporating these simplifications into the expression for the main term given in (4.3) on page 21 results in

\[
\text{main} = \epsilon (-B)^N \sum_{S, T \subset A \cup B^{-1}, \mu, \nu} \epsilon (-S)^{N+1(l(A)-1l(D)} \frac{\Delta(D; S)}{\Delta(T; S)} \times \left[ \prod_{i \geq 1} i^{m_i(\omega)} \right] \left( \rho^\beta_{\omega} \cup \rho^\alpha_{\omega} \right) \left[ \frac{\partial}{\partial p_{\omega}} \right] (\rho^\omega). 
\]

This is the main term stated in the Recipe up to elementary algebraic manipulations. Going back to our expression in (1.2) for the integral on the left-hand side, we obtain the error term by considering all ribbons that do not satisfy the condition given in (1.3). Taking absolute values inside the sums results in

\[
|\text{error}| \leq |\epsilon (B)^N | \sum_{\lambda} |L_{\lambda}^{\epsilon} \left( - (A \cup B^{-1}) ; D \right) | \sum_{\kappa} |s_{\kappa}(\lambda)| \times \left[ \prod_{i \geq 1} i^{m_i(\omega)} \right] \left( \rho^\beta_{\omega} \cup \rho^\alpha_{\omega} \right) \left[ \frac{\partial}{\partial p_{\omega}} \right] (\rho^\omega). 
\]
we eliminate the dummy variable 
and
µ
error
n
and
(4.5)
| error | \leq \|\varepsilon(B)\|_N \sum_{\mu' \cup \mu'} |\mathcal{L}S_{\nu' \cup \mu'} (- (A \cup B^{-1}) ; D)| \sum_{\kappa} |s_\kappa(C)|
\times \sum_{q + n > N - \|\varepsilon\|} \sum_{p_1, \ldots, p_k \geq 0} \left( \prod_{i = 1}^e |q_i + p_i| \right) \sum_{n_1, \ldots, n_f \geq 1; n_1 + \cdots + n_f = n} \left( \prod_{j = 1}^f \varphi_j^{n_j - 1} \right)
\times \sum_{l(\pi) \leq N} \left( \sum_{\mu' = (\lambda)^{(1)} \cdots (\lambda)^{(s)}} \left( \prod_{\mu'} |\mu'|^{q_1 \cdots q_s} \right) \right) \left( \sum_{\mu = (\mu)^{(1)} \cdots (\mu)^{(s)}} \left( \prod_{\mu} |\mu|^{p_1 \cdots p_s} \right) \right) \left( \sum_{\kappa = (\kappa)^{(1)} \cdots (\kappa)^{(s)} \mu = (\mu)^{(1)} \cdots (\mu)^{(s)}} \left( \prod_{\kappa} |\kappa|^{n_1 \cdots n_f} \right) \right).

The next step mirrors our derivation of the main term: We separate the partitions \mu and \nu in \mathcal{L}S_{\nu' \cup \mu'} (- (A \cup B^{-1}) ; D) by an application of Lemma 2.16 and then forget the condition that the union \nu' \cup \mu' must still be a partition. In addition, we eliminate the dummy variable \pi to combine the sequences of sums over the p_i- and n_j-ribbons:

| error | \leq \|\varepsilon(B)\|_N \sum_{\kappa, \mu, \nu} |\varepsilon(S)^{\|\lambda\| - \|\nu\|} | \Delta(D; S) | \frac{\Delta(D; S)}{\Delta(T; S)} | S, T \subset A \cup B^{-1}; S \cup \{A\} \cap \{A\} \leq (A \cup B^{-1}) \sum_{\kappa, \mu, \nu} |s_\kappa(S)| |\mathcal{L}S_{\nu' \cup \mu'} (- T; D)| |s_\kappa(C)|
\times \sum_{q + n > N - \|\varepsilon\|} \sum_{p_1, \ldots, p_k \geq 0} \left( \prod_{i = 1}^e |q_i + p_i| \right) \sum_{n_1, \ldots, n_f \geq 1; n_1 + \cdots + n_f = n} \left( \prod_{j = 1}^f \varphi_j^{n_j - 1} \right)
\times \left( \sum_{\nu' = (\lambda)^{(1)} \cdots (\lambda)^{(s)}} \left( \prod_{\nu'} |\nu'|^{q_1 \cdots q_s} \right) \right) \left( \sum_{\mu = (\mu)^{(1)} \cdots (\mu)^{(s)}} \left( \prod_{\mu} |\mu|^{p_1 \cdots p_s} \right) \right) \left( \sum_{\kappa = (\kappa)^{(1)} \cdots (\kappa)^{(s)} \mu = (\mu)^{(1)} \cdots (\mu)^{(s)}} \left( \prod_{\kappa} |\kappa|^{n_1 \cdots n_f} \right) \right).
Lemma 4.1 provides an upper bound for the two ribbon-counting sequences of sums. Indeed, all partitions \( \lambda^{(s)} \) are contained in a rectangle of width \( l(B) \), while all \( \mu^{(s)} \) and \( \kappa^{(s)} \) are contained in a rectangle of height \( N \). We conclude that

\[
\text{error} = O(l(A), l(B), l(D), l(E), A \cup B^{-1}) \left( N^{l(E) + l(F) - 1} \right) |e(B)^N| \\
\times \sum_{q + n > N - l(C)} \left( \sum_{q_1, \ldots, q_l(E) \geq 0; q_1 + \cdots + q_l(E) = q} \left( \prod_{i=1}^{l(E)} |E_{q_i + p_i - 1}| \right) \sum_{p_1, \ldots, p_l(E) \geq 0; p_1 + \cdots + p_l(E) = n} \prod_{j=1}^{l(F)} |F_{j}||^{n_j - 1} \right) \\
\times \sum_{S, T \subset A \cup B^{-1}; S \cup l(B), l(A) T \nsubseteq A \cup B^{-1}} \left( \sum_{\nu \in \{ l(B)^N \}} \sum_{\lambda \in l(C)^N} |s_{\nu}^{(s)}(S)| \right) \left( \sum_{\kappa, \mu} |LS_{\mu}(D; -T)| |s_{\kappa}^{(s)}(C)| \right)
\]

where the last sum is over pairs of partitions \( \kappa, \mu \) so that there exists a partition which can be obtained by adding \( n \) boxes to \( \kappa \) or by adding \( p = p_1 + \cdots + p_e \) boxes to \( \mu \).

\[ \Box \]

4.3. The results. In this section we present four theorems that can be viewed as special cases of the Recipe. In these instances we are able to give reasonable bounds for the error terms, unlike in the full generality of the Recipe. While the two formulas for averages of products of ratios and logarithmic derivatives seem to be new, formulas for pure ratios and pure products of logarithmic derivatives can be found in the literature. In fact, our expression for the ratios is just a reformulation of Bump and Gamburd’s ratio theorem \[BG06\]. The logarithmic derivative theorem presented here gives a neater and more combinatorial expression for the leading term of Conrey and Snaith’s expression for averages of logarithmic derivatives \[CS08\].

**Theorem 4.8** (ratios). Let \( A, B, C \) and \( D \) be sets of non-zero variables so that the latter two only contain elements that are strictly less than 1 in absolute value. Let the elements of \( A \cup B^{-1} \) be pairwise distinct. If \( l(D) \leq N + l(A) \) and \( l(C) \leq N \), then

\[
\int_{U(N)} \frac{\prod_{\alpha \in A} \chi_{\alpha}(\alpha) \prod_{\beta \in B} \chi_{\beta}^{-1}(\beta) \prod_{\gamma \in C} \chi_{\gamma}^{-1}(\gamma)}{\prod_{\delta \in D} \chi_{\delta}(\delta)} d\theta_N \\
= \epsilon(-B)^N \sum_{S, T \subset A \cup B^{-1}; S \cup l(B), l(A) T \nsubseteq A \cup B^{-1}} \epsilon(-S)^{N + l(A) - l(D)} \Delta(D; S) \Delta(T; S) \prod_{\gamma \in C} (1 - \gamma^{-1}) \prod_{\delta \in D} (1 - t_{\delta}).
\]

Remark. Theorem 4.8 is basically the ratio theorem presented in \[BG06\], except for the assumptions on the lengths of the sets of variables. This does not come as a surprise given that the proof of the Recipe is based on Bump and Gamburd’s approach. Their theorem holds under the assumption that \( l(C) + l(D) \leq N \). In fact, they only state \( l(C) \), \( l(D) \leq N \) as a requirement but their proof implicitly makes us of the stronger assumption: on page 246 of \[BG06\] they apply Proposition 8 (a weaker version of Lemma 2.10), which is only permissible if \( l(C) + l(D) \leq N \).
Proof. In a first step, let us suppose that \( l(\mathcal{D}) \leq l(\mathcal{A}) \), in which case the equality in (4.7) follows from the Recipe. We set \( \mathcal{E} = \emptyset = \mathcal{F} \) in (4.12). Under the assumption that \( l(\mathcal{C}) \leq N \), the error term vanishes. The main term simplifies to

\[
\sum_{\lambda} \varepsilon(-\mathcal{B})^{N-\varepsilon(\mathcal{A})-l(\mathcal{D})} \frac{\Delta(\mathcal{D} ; \mathcal{S})}{\Delta(\mathcal{T} ; \mathcal{S})} \sum_{\lambda} z_{\lambda}^{-1} p_{\lambda} \left( \rho_{\mathcal{A}}^{\mathcal{B}} \cup \rho_{\mathcal{D}}^{\mathcal{B}} \right) p_{\lambda}(\mathcal{C}).
\]

We write the sum over \( \lambda \) as a product: According to Lemma 2.5 and Remark 2.12

\[
\sum_{\lambda} z_{\lambda}^{-1} p_{\lambda} \left( \rho_{\mathcal{A}}^{\mathcal{B}} \cup \rho_{\mathcal{D}}^{\mathcal{B}} \right) p_{\lambda}(\mathcal{C}) = \sum_{\mu} g_{\mu} \left( \rho_{\mathcal{A}}^\mathcal{B} \cup \rho_{\mathcal{D}}^\mathcal{B} \right) g_{\mu}(\mathcal{C})
\]

\[
= \sum_{\mu} LS_{\mu}(\mathcal{D} ; -\mathcal{T}) g_{\mu}(\mathcal{C}) = \prod_{\gamma \in \mathcal{C}} \left( 1 - \gamma \delta \right)^{-1} \prod_{t \in T} \left( 1 - t \gamma \right)
\]

where the last equality is a consequence of the generalized Cauchy identity (i.e. Proposition 2.11).

In order to justify the equality in (4.7) in case \( l(\mathcal{D}) \leq N + l(\mathcal{A}) \), it suffices to note that in the proof of the Recipe the role of the assumption that \( l(\mathcal{D}) \leq l(\mathcal{A}) \) is to ensure that the \( (l(\mathcal{D}) , l(\mathcal{A}) + l(\mathcal{B})) \)-index of of \( \nu' \cup \mu' \) is less than \( l(\mathcal{A}) \), which makes it permissible to apply Lemma 2.10 to \( LS_{\nu' \cup \mu'}(-(-\mathcal{A} \cup \mathcal{B}^{-1}) ; \mathcal{D}) \). Given that \( \mathcal{E} = \emptyset \), the partition \( \nu' \) is equal to the rectangle \( \langle \mathcal{N}(\mathcal{B}) \rangle \). Therefore, the \( (l(\mathcal{D}) , l(\mathcal{A}) + l(\mathcal{B})) \)-index of \( \nu' \cup \mu' \) is less than \( l(\mathcal{A}) \) whenever \( l(\mathcal{D}) \leq N + l(\mathcal{A}) \). \( \square \)

**Theorem 4.9.** Let \( \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} \) and \( \mathcal{E} \) be sets of non-zero variables so that the elements of \( \mathcal{A} \cup \mathcal{B}^{-1} \) are pairwise distinct and \( l(\mathcal{D}) \leq l(\mathcal{A}) \). If \( \text{abs}(\mathcal{A}) , \text{abs}(\mathcal{B}) \leq 1 , \text{abs}(\mathcal{C}) , \text{abs}(\mathcal{D}) < 1 \) and there exists \( r \in \mathbb{R} \) so that \( \text{abs}(\mathcal{E}) \leq r < 1 \), then

\[
\int \frac{\prod_{\alpha \in \mathcal{A}} x_{\alpha}(\alpha) \prod_{\beta \in \mathcal{B}} x_{\beta}^{-1}(\beta) \prod_{\gamma \in \mathcal{C}} x_{\gamma}(\gamma) \prod_{t \in \mathcal{T}} x_{t}(t)}{\prod_{\mathcal{D}} x_{\mathcal{D}}(\mathcal{D})} \frac{\Delta(\mathcal{D} ; \mathcal{S})}{\Delta(\mathcal{T} ; \mathcal{S})} \prod_{\gamma \in \mathcal{C}} \left( 1 - \gamma \delta \right)^{-1} \prod_{t \in \mathcal{T}} \left( 1 - t \gamma \right)
\]

\[
\times (-1)^{l(\mathcal{E})} \sum_{\mathcal{E}' \subseteq \mathcal{E} : \mathcal{E}' \cup \mathcal{E}'^{-1} \subseteq \mathcal{E}} \left( \sum_{\psi \leq \mathcal{C} : l(\mathcal{E}') = l(\mathcal{E})} \left( \sum_{\psi_{\mathcal{C}}} \left( m_{\mathcal{C}}(\mathcal{C}') \right) p_{\mathcal{C}}(\mathcal{C}) \right) + O_{\mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} , \text{max} \{ \text{abs}(\mathcal{C}) , \text{abs}(\mathcal{D}) \} \right) \left( r^{N(\mathcal{B}^{-1}) + 2 \mathcal{E}^{-1}} \right)
\]

If, in addition to the conditions stated above, \( \text{abs}(\mathcal{B}) = r_1 \leq r \) for some \( r_1 \in \mathbb{R} \), then the bound on the error term can be improved by a factor of \( r^N \).
Proof. We set \( F = \emptyset \) in the statement of the Recipe. The main term of the expression on the right-hand side simplifies to

\[
\epsilon(-B)^N \sum_{S, T \in A: \cup B^{-1}, \cup A_{l(B), l(A)} T \in A: B^{-1}} \epsilon(-S)^{N + l(A) - l(D)} \frac{\Delta(D; S)}{\Delta(T; S)}
\]

\[
\times (-1)^{|E|} \sum_{E, E' \subseteq E: E \cup E' = E} \left( \sum_{\lambda \in (E')} \left( \sum_{\rho \in \lambda} \frac{c_{\mu}(-1)^{\nu}}{\rho_{\lambda^{(\rho)}}} (\rho_{E, E'}) \frac{c_{\mu}(-1)^{\nu}}{\rho_{\lambda^{(\rho)}}} \frac{c_{\mu}(-1)^{\nu}}{\rho_{\lambda^{(\rho)}}} \right) \right)
\]

\[
\times \left( \sum_{\lambda \in (E')} \left( \sum_{\rho \in \lambda} \frac{c_{\mu}(-1)^{\nu}}{\rho_{\lambda^{(\rho)}}} (\rho_{E, E'}) \frac{c_{\mu}(-1)^{\nu}}{\rho_{\lambda^{(\rho)}}} \frac{c_{\mu}(-1)^{\nu}}{\rho_{\lambda^{(\rho)}}} \right) \right)
\]

As in the proof of Theorem 4.8, the generalized Cauchy identity allows us to replace the sum over \( \lambda \) by a product. This yields the main term of this theorem.

It remains to bound the error given in equation (4.6) on page 25. We exploit \( |\epsilon| \leq r \) and \( n = 0 \) (since \( F = \emptyset \)) to infer the following bound:

\[
\text{error} = O_{r, l, l(A), l, l(D), l, l(E), A: B^{-1}} \left( N^{(l(E) - 1)^+} \right) |\epsilon(B)|^N
\]

\[
\times \sum_{q > N^2 l(C)} r^q \left( \sum_{\nu \geq 0; \nu \geq 0; \nu = q} 1 \right) \sum_{p \geq 0} r^p \left( \sum_{p_1 + \ldots + p_{l(E)} = p} 1 \right)
\]

\[
\times \sum_{S, T \in A: \cup B^{-1}, \cup A_{l(B), l(A)} T \in A: B^{-1}} \left| s_{\nu, \nu} (S) \right|
\]

\[
\times \left( \sum_{S, T \in A: \cup B^{-1}, \cup A_{l(B), l(A)} T \in A: B^{-1}} \left| L_{S, T} \right| \left| s_{\nu, \nu} (C) \right| \right)
\]

As \( s_{\nu, \nu} (C) \) vanishes if \( l(\nu) > l(C) \), we have that \( l(\mu) \leq l(\nu) \leq l(C) \) for all partitions that appear in the sum over \( \nu \) and \( \mu \). Setting \( R = \max\{|\nu|, |\epsilon(D)|\} < 1 \), Lemmas 4.5 and 4.6 thus entail that

\[
\sum_{\nu, \nu} \left| L_{S, T} \right| \left| s_{\nu, \nu} (C) \right| =
\]

\[
\sum_{\nu, \nu} O_{l(C), l(D), l(A), B^{-1}} \left( |\nu|^{l(D) - 1 + l(C)} R^{|\nu|} \right)
\]

\[
O_{l(C), l(D), l(A), B^{-1}} \left( R^{l(C) - 1} R^{|\nu|} \right)
\]
where we have crudely bounded the number of partitions of length $n$ and size $m$ by $(m + 1)^{(n-1)+}$. Bounding the number of integers $p_1, \ldots, p_l$ whose sum equals $p$ by $(p + 1)^{(l - 1)+}$, another argument based on geometric series thus allows us to conclude that the sum over $p$ is $O_{l(C), l(D), l(E), \max\{\abs{C}, \abs{D}\}, A, \nu} (1)$. Two applications of Lemma 4.3 will allow us to bound

\[
S(N) = \sum_{S, T \subseteq A \cup B^{-1} : S \cup l(B), l(A), T \subseteq A \cup B^{-1}}^{|e(B)^N|} \sum_{\nu \subset l(B)^N, |\nu| = l(B)^N - q} |g_{\nu}(S)|.
\]

Suppose that $\abs{B} = r_1$ for some $r_1 \leq r$. As $\abs{A} \leq 1 < r_1^{-1}$, the first statement in Lemma 4.3 implies that

\[
S(N) = O_{A \cup B^{-1}} \left( r_1^{N l(B)} \sum_{\nu \subset l(B)^N, |\nu| = l(B)^N - q} r_1^{-|\nu|} \right) = O_{A \cup B^{-1}} \left( r^q N^{l(B)^N - 1} \right).
\]

This last bound is due to the fact that for $l(B) \geq 1$, the number of partitions $\nu$ of some fixed size $Q$ that are contained in the rectangle $l(B)^N$ is at most $(N + 1)^{(l(B)^N - 1)}$. Indeed, $\nu_1'$ to $\nu_{l(B)}'$ are some integers between 0 and $N$, while $\nu_{l(B)}'$ is determined by the condition that $\nu_1' + \cdots + \nu_{l(B)}' = Q$. If we only assume that $\abs{A}$, $\abs{B} \leq 1$, the second bound in Lemma 4.3 allows us to infer that

\[
S(N) = O_{A \cup B^{-1}} \left( \sum_{\nu \subset l(B)^N, |\nu| = l(B)^N - q} e(B)^N \nu_1' \right) = O_{A \cup B^{-1}} \left( N^{l(B)^N - 1} \right),
\]

since $\nu_1' \leq N$ for all partitions $\nu$ that appear in the sum. In conclusion,

\[
\text{error} = O_p \left( N^{l(E)^N - 1 + (l(E)^N - 1)^+} \sum_{q > N-l(C)} r^q (1 + \delta_{\abs{B} = r_1}) (q + 1)^{(l(E)^N - 1)^+} \right)
\]

where $\delta_{\abs{B} = r_1}$ indicates whether the additional condition on $B$ is satisfied. Here, the implicit constant depends on

\[
P = \{r, l(C), l(D), l(E), \max\{\abs{C}, \abs{D}\}, A, B\}.
\]

The bound stated in the theorem follows from yet another argument based on geometric series. □
Theorem 4.10. Let \( r \in \mathbb{R} \) with \( r < 1 \). Let \( \mathcal{B}, \mathcal{C}, \mathcal{E} \) and \( \mathcal{F} \) be sets of non-zero variables so that \( \abs{\mathcal{B}} \leq 1 \), \( \abs{\mathcal{C}} < 1 \) and \( \abs{\mathcal{E}}, \abs{\mathcal{F}} \leq r \). Then

\[
\int_{U(N)} \frac{\prod_{\beta \in \mathcal{B}} \chi_{g^{-1}(\beta)}}{\prod_{\gamma \in \mathcal{C}} \chi_{g^{-1}(\gamma)}} \prod_{e \in \mathcal{E}} \chi_{g(e)} \prod_{\varphi \in \mathcal{F}} \chi_{g^{-1}(\varphi)} \, dg
\]

\[
= (-1)^{l(\mathcal{E})+l(\mathcal{F})} \sum_{\mathcal{E}^{'}, \mathcal{E}'' \subseteq \mathcal{E}, \mathcal{E} \cup \mathcal{E}'' \subseteq \mathcal{E}} \left[ \sum_{l(\chi) = l(\mathcal{E}''')} m_{\chi} \langle 1(\mathcal{E}''), \mathcal{E}'' \rangle \mathcal{P}_\chi(-\mathcal{B}) \right] 
\]

\[
\times \sum_{l(\psi) = l(\mathcal{E}')} m_{\psi} \langle 1(\mathcal{E}'), \mathcal{E}'' \rangle \sum_{l(\omega) = l(\mathcal{F})} m_{\omega} \langle 1(l(\mathcal{F})), \mathcal{F} \rangle \prod_{i \geq 1} \frac{\epsilon^{m_i}(\omega)m_i(\psi)!}{m_i(\psi \setminus \omega)!} \mathcal{P}_\omega(\mathcal{C}) 
\]

+ error.

In particular, the main term vanishes unless \( l(\mathcal{F}) \leq l(\mathcal{E}) \). To provide a bound for the error term we require one of the following additional conditions on the set of variables \( \mathcal{B} \). If there exists a real number \( r_1 \leq r \) with \( \abs{\mathcal{B}} = r_1 \), then

\[
\text{error} = O_P \left( r^{3N} A^{l(\mathcal{B})^2 + l(\mathcal{B}) - 1} + (l(\mathcal{E}) - 1)^+ + (l(\mathcal{F}) - 1)^+ + (l(\mathcal{F}) + l(\mathcal{E}) + l(\mathcal{F}) - 1)^+ + 2 \right),
\]

where the implicit constant depends on \( \mathcal{P} = \{r, l(\mathcal{B}), l(\mathcal{C}), l(\mathcal{E}), l(\mathcal{F}), \max(\abs{\mathcal{C}})\} \).

If the elements of \( \mathcal{B} \) are pairwise distinct, then

\[
\text{error} = O_P \left( r^{N} N^{l(\mathcal{B}) - 1} + l(\mathcal{E}) - 1)^+ + (l(\mathcal{F}) - 1)^+ + (l(\mathcal{E}) + l(\mathcal{F}) - 1)^+ + 2 \right),
\]

where the implicit constant depends on \( \mathcal{P} = \{r, \mathcal{B}, l(\mathcal{C}), l(\mathcal{E}), l(\mathcal{F}), \max(\abs{\mathcal{C}})\} \).

Proof. The idea of the proof is to view this statement as a special case of the Recipe by setting \( \mathcal{A} = \emptyset = \mathcal{D} \). Technically, this is not permissible since the elements of \( \mathcal{B} \) are not assumed to be pairwise distinct. However, for the main term it is enough to slightly perturb the elements of \( \mathcal{B} \) before applying the Recipe, and then make the perturbations vanish. For the error term, one quickly checks the proof of the Recipe to see that the implicit constant does in fact not depend on \( \mathcal{B} \) itself - but only on its length - in case \( l(A) = 0 = l(D) \): the terms on the right-hand side in (1.5) that depend on \( \mathcal{B} \) are

\[
\text{terms}(\mathcal{B}) = \left| \epsilon(\mathcal{B})^N \right| \sum_{\substack{\mu, \nu \in \text{partition} \\ \nu' \cup \mu' \subseteq \nu \subseteq \langle l(\mathcal{B}) \rangle}} \left| L_{S_{\nu' \cup \mu'}} \left( - (A \cup B^{-1}) ; D \right) \right|.
\]

Under the assumption that \( \mathcal{A} = \emptyset = \mathcal{D} \),

\[
\text{terms}(\mathcal{B}) = \left| \epsilon(\mathcal{B})^N \right| \sum_{\substack{\mu, \nu \in \text{partition} \\ \nu' \cup \mu' \subseteq \nu \subseteq \langle N(\mathcal{B}) \rangle}} \left| g_{\nu' \cup \mu'} \left( - B^{-1} \right) \right|.
\]

Moreover, the fact that any Schur function \( g_{\lambda}(X) \) vanishes whenever \( l(\lambda) > l(X) \) entails that only terms with \( \mu = \emptyset \) contribute to the sum. Hence, we are left with

\[
\text{terms}(\mathcal{B}) = \left| \epsilon(\mathcal{B})^N \right| \sum_{\nu' \subseteq \langle N(\mathcal{B}) \rangle} \left| g_{\nu'} \left( - B^{-1} \right) \right|,
\]

which also appears in (1.6), restricted to \( \mathcal{A} = \emptyset = \mathcal{D} \).
Having resolved this technicality, we now set $A = \emptyset = D$ in the Recipe. It easily follows from the explicit expression for $\varepsilon_{\lambda}^{-1} p_{\lambda} \left( \rho_{\phi}^{\beta} \right)$ given in (2.4) that this power sum vanishes unless $\lambda = \emptyset$. Thus, the main term on the right-hand side of the equality in (4.1) simplifies to

$$
\text{main} = (-1)^n l(\mathcal{E}) + l(\mathcal{F}) \sum_{\mathcal{E}' \mathcal{E}'' \subset \mathcal{E}} \sum_{q,n \geq 0; \mathcal{E}' \cup \mathcal{E}'' \mathcal{E}} \sum_{q+n \leq N - l(\mathcal{C})} \left( \sum_{\chi; \mathcal{E}'} m_{\chi} \overline{\langle i(\mathcal{E}'') \rangle} \langle -\mathcal{E}'' \rangle \right) \left( \sum_{\psi; \mathcal{F}} m_{\psi} \langle i(\mathcal{E}'') \rangle \langle -\mathcal{B} \rangle \right) \\
\times \sum_{\psi; \mathcal{F}} m_{\psi} \langle i(\mathcal{E}'') \rangle \langle \mathcal{F} \rangle \\
\times \sum_{\omega; \mathcal{C}} m_{\omega} \langle i(\mathcal{F}) \rangle \langle \mathcal{F} \rangle \prod_{i \geq 1} \frac{i^{m_{i}(\omega)} m_{i}(\psi)!}{m_{i}(\psi - \omega)!} p_{\psi \omega}(\mathcal{C}).
$$

Owing to the condition that $\omega$ be a subsequence of $\psi$, this expression vanishes unless $l(\mathcal{F}) \leq l(\mathcal{E})$. In addition, this condition allows us to eliminate the dependence on $N$ at the cost of incurring an error that is

$$
O_{r, l(\mathcal{B}), l(\mathcal{C}), l(\mathcal{E}), l(\mathcal{F})} \left( r N^{l(\mathcal{E}) + 1} + (l(\mathcal{F}) - 1)^{+} + (l(\mathcal{F}) + 1) \right)
$$

Under the assumption that $|\text{abs}(\mathcal{B})| \leq r$, the bound may even be multiplied by $r^N$. Indeed, the sum over $q, n \geq 0$ so that $q + n > N - l(\mathcal{C})$ is

$$
O \left( \sum_{q,n \geq 0; \mathcal{C}} r^{q+n} \right) = \sum_{n+p+q \geq 0; \mathcal{C}} r^{q+n} \left( N^{l(\mathcal{E}) + l(\mathcal{F}) - 1} \right) \left( l(\mathcal{F}) \right) \prod_{i \geq 1} \frac{i^{m_{i}(\omega)} m_{i}(\psi)!}{m_{i}(\psi - \omega)!} p_{\psi \omega}(\mathcal{C})
$$

where $\delta(\text{abs}(\mathcal{B}) \leq r)$ indicates whether $|\text{abs}(\mathcal{B})| \leq r$. Handling the sums counting partitions as in the preceding proof, and then employing an argument based on geometric series gives the desired bound.

It remains to show the bound on the error inherited from the Recipe. Given that $A = \emptyset = D$, the formula in (4.0) simplifies to

$$
\text{error} = O_{r, l(\mathcal{B}), l(\mathcal{E}), l(\mathcal{F})} \left( N^{l(\mathcal{E}) + l(\mathcal{F}) - 1} \right) \left( l(\mathcal{B}) \right) N \sum_{q,n \geq 0, \mathcal{C}} r^{q+n} \\
\times \left( \sum_{p \geq 0} r^p \right) \left( \sum_{p \geq 0} r^p \right) \left( \sum_{p \geq 0} r^p \right) \left( \sum_{p \geq 0} r^p \right)
$$

$$
\times \left( \sum_{\nu \subset \{ l(\mathcal{B}) \}^N \mathcal{C}} |\varphi_{\nu}(\mathcal{B}^{-1})| \right) \left( \sum_{\nu \subset \mathbb{N}} |\varphi_{\nu}(\mathcal{C})| \right)
$$
where we have also used that $|\text{abs}(E)|, |\text{abs}(F)| \leq r$. First, consider the following function of $N$, which also depends on $B$:

$$S_q(N) \triangleq |\varepsilon(\mathcal{B})|^N \sum_{\nu \subseteq \{l(\mathcal{B})\}^N, \nu \neq \emptyset} |s_{\nu'}(B^{-1})| .$$

Hence, error that is independent of $B$.

$$\left| |s_{\nu'}(B^{-1})|\right| \text{ where we have used that the condition on } \nu \text{ implies that } n \geq p. \text{ An argument based on geometric series gives}$$

$$\text{error} = O_{r, l(\mathcal{B}), l(\mathcal{C}), l(E), l(F), \max(\text{abs}(\mathcal{C}))} \left( N(l(\mathcal{E}) + l(\mathcal{F}) - l(\mathcal{C})) \right)$$

$$\times \sum_{q, n \geq 0; q + n > N - l(\mathcal{C})} S_q(N) r^{q + n + l(\mathcal{E}) - l(\mathcal{F})} \sum_{p \geq n} r^p l(\mathcal{E}) l(\mathcal{F}) \cdots l(\mathcal{F})$$

Replacing $S_q(N)$ by the appropriate bound concludes the proof. □
Theorem 4.11 (logarithmic derivatives). Let \( r \in \mathbb{R} \), let \( \mathcal{E} \) and \( \mathcal{F} \) be sets of variables so that \( 0 < \text{abs}(\mathcal{E}) \), \( \text{abs}(\mathcal{F}) \leq r < 1 \). Then

\[
\int_{U(N)} \prod_{\varepsilon \in \mathcal{E}} \frac{\chi'_{\varepsilon}(\varphi)}{\chi_{\varepsilon}(\varphi)} \prod_{\varphi \in \mathcal{F}} \frac{\chi'_{\varphi}(\varphi)}{\chi_{\varphi}(\varphi)} d\varphi
\]

\[
= \begin{cases} 
\sum_{\lambda} \chi_{\lambda}(\mathcal{E}) \chi_{\lambda}(\mathcal{F}) & \text{if } l(\mathcal{E}) = l(\mathcal{F}) \\
0 & \text{otherwise}
\end{cases}
\]

\[
+ O_{r,l(\mathcal{E}),l(\mathcal{F})} \left( r^{2N^l(\mathcal{E})-1} + l(\mathcal{F})^1 + l(\mathcal{E}) + l(\mathcal{F})^{-1} + 2 \right).
\]

We have made no effort to optimize the exponent of \( N \) in the bound for the error term.

Proof. We set \( \mathcal{B} = \emptyset = \mathcal{C} \) in Theorem 4.10. As \( p_\lambda(\emptyset) = 0 \) unless \( \lambda = \emptyset \), the right-hand side of the equality in (4.13) simplifies to

\[
(-1)^{l(\mathcal{E})+l(\mathcal{F})} \sum_{\psi} \chi_{\psi}(\mathcal{E}) \chi_{\psi}(\mathcal{F}) \prod_{\omega \in \mathcal{B}} \sum_{l(\omega) = l(\mathcal{F})} \sum_{i \geq 1} \chi_{\omega}^{i}(\psi) \chi_{\omega}^{-i}(\mathcal{F}) + O_{r,l(\mathcal{E}),l(\mathcal{F})} \left( r^{2N^l(\mathcal{E})-1} + l(\mathcal{F})^1 + l(\mathcal{E}) + l(\mathcal{F})^{-1} + 2 \right),
\]

which entails that the main term vanishes unless \( l(\mathcal{E}) = l(\mathcal{F}) \). The expression stated in the theorem is obtained by substituting \( \lambda \) for both \( \psi \) and \( \omega \). \( \square \)

Remark. In [CS08, p. 486], Conrey and Snaith derive a formula for

\[
\int_{U(N)} \prod_{\alpha \in \mathcal{A}} \left( -e^{-\alpha} \right) \frac{\chi'_{\alpha}(e^{-\alpha})}{\chi_{\alpha}(e^{-\alpha})} \prod_{\beta \in \mathcal{B}} \left( -e^{-\beta} \right) \frac{\chi'_{\beta}(e^{-\beta})}{\chi_{\beta}(e^{-\beta})} d\varphi
\]

without employing any combinatorial methods. Compared to the logarithmic derivative theorem presented in this paper, their formula has the distinct advantage of providing an exact expression for the integral. Its principal disadvantage is that this expression is rather complicated, which makes it cumbersome to use. In Conrey and Snaith’s theorem, it is not immediately obvious, for instance, that the leading term vanishes unless \( l(\mathcal{A}) = l(\mathcal{B}) \). Hence, Theorem 4.11 is an improvement because it provides a simple expression in terms of one of the standard bases for the ring of symmetric functions.

5. From logarithmic derivatives to an explicit formula

This section is dedicated to an application of the logarithmic derivative theorem, which is motivated by the analogy between \( L \)-functions and characteristic polynomials alluded to in the introduction. We present an explicit formula for eigenvalues whose derivation mirrors the proof of the explicit formula for zeros of \( L \)-functions given in [RS96]. As Rudnick and Sarnak’s proof is based on completed \( L \)-functions, which are more natural to work with than classic \( L \)-functions, we introduce the analogous notion of completed characteristic polynomials. In addition, we give a formula for products of logarithmic derivatives of completed characteristic polynomials.

Definition 5.1 (completed characteristic polynomial). For unitary matrices \( g \) that satisfy \( \det(-g) \neq -1 \), we define the completed characteristic polynomial as

\[
\Lambda_g(z) = \det(-g)^{1/2} z^{-N/2} \chi_g(z).
\]
Notice that while the characteristic polynomial $\chi_g$ is an entire function, $\Lambda_g$ might only be defined on $\mathbb{C} \setminus \mathbb{R}^-$.

Our reason for considering the completed characteristic polynomial is the following symmetry with respect to the transformation given by $z \mapsto z^{-1}$, which is a basic linear algebra exercise.

**Lemma 5.2** (functional equation). For $g \in U(N)$ with $\det(-g) \neq -1$ the following equalities hold.

1. For all $z \in \mathbb{C} \setminus \mathbb{R}^-$, $\Lambda_g(z) = \Lambda_{g^{-1}}(z^{-1})$.
2. For all $z \in \mathbb{C}$ that are not eigenvalues of $g$, $z \frac{\Lambda'_g(z)}{\Lambda_g(z)} = -z^{-1} \frac{\Lambda'_{g^{-1}}(z^{-1})}{\Lambda_{g^{-1}}(z^{-1})}$.

A formula for products of logarithmic derivatives of completed characteristic polynomials is easily deduced from the logarithmic derivative theorem for classic characteristic polynomials.

**Theorem 5.3** (completed logarithmic derivatives). Let $r \in \mathbb{R}$, let $\mathcal{E}$ and $\mathcal{F}$ be sets of non-zero variables so that $\text{abs}(\mathcal{E}), \text{abs}(\mathcal{F}) \leq r < 1$. Then

\[
\int_{U(N)} \prod_{\varepsilon \in \mathcal{E}} \frac{\Lambda'_g(\varepsilon)}{\Lambda_g(\varepsilon)} \prod_{\varphi \in \mathcal{F}} \frac{\Lambda'_{g^{-1}}(\varphi)}{\Lambda_{g^{-1}}(\varphi)} \, dg
\]

\[
= \sum_{\lambda} \left( -\frac{N}{2} \right)^{l(\mathcal{E})+l(\mathcal{F})-2l(\lambda)} z_\lambda m_\lambda(\mathcal{E}) m_\lambda(\mathcal{F})
+ O_{r,l(\mathcal{E}),l(\mathcal{F})} \left( 2^N N^{l(\mathcal{E})+l(\mathcal{F})+l(l(\mathcal{E})-1)^+ + l(l(\mathcal{F})-1)^+ + l(l(\mathcal{E})+l(l(\mathcal{F})-1)^+ \right).
\]

The two logarithmic derivative theorems presented in this chapter are part of the reasons why we consider it more natural to work with completed characteristic polynomials in the context of viewing random matrix theory as a model for number theory: the main term in Theorem 4.11 is a sum that ranges over all partitions, while the main term in Theorem 4.11 is a sum that ranges over all partitions of a fixed length, which we consider an "unnatural" restriction.

**Proof.** Notice that $\{ g \in U(N) : \det(-g) = -1 \}$ is a null set with respect to Haar measure on $U(N)$. Hence, the fact that $\Lambda_g$ is not defined on this set is of no concern.

We reformulate the left-hand side in (5.9) such that we can apply Theorem 4.11.
We remark that this equality holds thanks to the convention fixed in Section 2.1, which ensures that every sequence of length \( n \) has exactly \( 2^n \) subsequences. Theorem 5.4 allows us to compute the integral:

\[
\text{LHS} = \sum_{\substack{E' \subset E \\ E \subset E}} \left( -\frac{N}{2} \right)^{l(E) + l(F) - 2l(E')} \left( \prod_{\varepsilon \in E'} \varepsilon \right) \left( \prod_{\varphi \in F'} \varphi \right) \\
\times \sum_{\lambda} \frac{z_\lambda m_\lambda \left( \varepsilon \right) m_\lambda \left( F' \right)}{l(\lambda) = l(E')} + O_{r,l(E),l(F)} \left( 2^N N^{l(E) + l(F) + (l(E) - 1)^+ + (l(F) - 1)^+ + l(E) + l(F) - 1)^+ + 2 \right) \\
= \sum_{\lambda} \left( -\frac{N}{2} \right)^{l(E) + l(F) - 2l(\lambda)} \sum_{\substack{E' \subset E \\ E \subset E \subset E}} \frac{z_\lambda m_\lambda \left( E' \right) m_\lambda \left( F' \right)}{l(\lambda) = l(E')} + O_{r,l(E),l(F)} \left( 2^N N^{l(E) + l(F) + (l(E) - 1)^+ + (l(F) - 1)^+ + l(E) + l(F) - 1)^+ + 2 \right).
\]

By the definition of the monomial symmetric polynomials, the main term simplifies to the desired expression.

A formula for the average of products of logarithmic derivatives of completed characteristic polynomials over the unitary group which holds inside the unit circle is the only tool we need to derive an explicit formula for eigenvalues of unitary matrices.

**Theorem 5.4 (explicit formula).** Fix \( r \in \mathbb{R} \) with \( 0 < r < 1 \). Let \( A(r) \) denote the closed annulus (about the origin) with inner radius \( r \) and outer radius \( r^{-1} \), and \( D \) (\( r^{-1} \)) the closed disc (about the origin) of radius \( r^{-1} \). Let \( h \) be a meromorphic function on \( D \) (\( r^{-1} \)) which is holomorphic on \( A(r) \). Let \( f \) be a symmetric function in \( n \) variables such that \( z \mapsto f(z, z_2, \ldots, z_n) \) is meromorphic on \( D \) (\( r^{-1} \)) and holomorphic on \( A(r) \). If \( \{p_1, \ldots, p_N\} \) are the eigenvalues of \( g \in U(N) \), then

\[
\int_{U(N)} \sum_{1 \leq j_1 \leq \ldots \leq j_n \leq N} h(p_{j_1}) \ldots h(p_{j_n}) f(p_{j_1}, \ldots, p_{j_n}) dg \\
= \sum_{\lambda} \left( -\frac{N}{2} \right)^{n - 2l(\lambda)} \frac{z_\lambda}{(2\pi)^n} \sum_{k=0}^{n} \binom{n}{k} \\
\times \int_{[0,2\pi]^n} \left( \prod_{j=1}^{k} h \left( r e^{-it_j} \right) \right) \left( \prod_{j=k+1}^{n} h \left( \frac{e^{it_j}}{r} \right) \right) \\
\times f \left( r e^{-it_1}, \ldots, r e^{-it_k}, \frac{e^{it_{k+1}}}{r}, \ldots, \frac{e^{it_n}}{r} \right) \\
\times m_\lambda \left( r e^{-it_1}, \ldots, r e^{-it_k} \right) m_\lambda \left( r e^{-it_{k+1}}, \ldots, r e^{-it_n} \right) dt_1 \ldots dt_n \\
+ O_{r,n,h,f} \left( 2^N N^{3n+2} \right).
\]

*In the context of this theorem we call a function \( f \) symmetric if it is invariant under the permutation of its variables, which means that \( f \) need not be an element of the ring of symmetric functions.*
Proof. The function \( \Lambda_\delta'(z) / \Lambda_\delta(z) \) is meromorphic on the entire complex plane with simple poles at \( \{0, \rho_1, \ldots, \rho_N\} \); its residue at \( \rho_i \) is the multiplicity of \( \rho_i \). We consider the following path integral along the border of \( A(r) \), i.e. along \( \delta = \delta(r) + \delta(r^{-1}) \)

where \( \delta(r) : [0, 2\pi] \to \mathbb{C}; t \mapsto re^{-i\theta} \) and \( \delta(r^{-1}) : [0, 2\pi] \to \mathbb{C}; t \mapsto r^{-1}e^{i\theta} \):

\[
Eig(g) \triangleq \frac{1}{(2\pi i)^n} \int_\delta \cdots \int_\delta \prod_{i=1}^n \frac{\Lambda'_{\delta}(z_i)}{\Lambda_{\delta}(z_i)} h(z_i) f(z_1, \ldots, z_n) dz_1 \ldots dz_n.
\]

Given that the interior of \( \delta \) does not contain the origin, repeated application of the residue theorem allows us to infer that the above expression is equal to the integral on the left-hand side in (5.2).

In a next step, we show that the integral of Eig\((g)\) over the unitary group is also equal to the right-hand side in (5.2). Recalling that each integral along the path \( \delta \) is the sum of the integrals along \( \delta(r) \) and \( \delta(r^{-1}) \), we multiply out (exploiting the fact that \( f \) is symmetric), and then apply the functional equation for the completed characteristic polynomial \( i.e. \) Lemma 5.2 to the logarithmic derivatives that are integrated along \( \delta(r^{-1}) \):

\[
Eig(g) = \sum_{\mathcal{E}, \mathcal{F} \subset [n]: \mathcal{E} \cup \mathcal{F} = \mathcal{N}} \frac{(-1)^{l(\mathcal{E})}}{(2\pi i)^n} \int_{\delta(r)} \cdots \int_{\delta(r^{-1})} \prod_{\phi \in \mathcal{F}} h(z_\phi) \prod_{\phi \in \mathcal{E}} h(z_\phi) f(z_\mathcal{E} \cup z_\mathcal{F}) \prod_{\phi \in \mathcal{E}} \frac{\Lambda_{\delta(r^{-1})}(z_\phi)}{z_\phi}
\]

Here the superscripts of the integrals indicate which variables are integrated along \( \delta(r) \), and which along \( \delta(r^{-1}) \). Using Theorem 5.3 to integrate this expression over \( U(N) \) gives

\[
\int_{U(N)} Eig(g) dg = \sum_{\mathcal{E}, \mathcal{F} \subset [n]: \mathcal{E} \cup \mathcal{F} = \mathcal{N}} \frac{(-1)^{l(\mathcal{F})}}{(2\pi i)^n} \int_{\delta(r)} \cdots \int_{\delta(r^{-1})} \prod_{\phi \in \mathcal{E}} h(z_\phi) \prod_{\phi \in \mathcal{F}} h(z_\phi) f(z_\mathcal{E} \cup z_\mathcal{F}) \prod_{\phi \in \mathcal{E}} \frac{\Lambda_{\delta(r^{-1})}(z_\phi)}{z_\phi}
\]

\[
= \sum_{\lambda} \left( \frac{N}{2} \right)^{l(\mathcal{E})+l(\mathcal{F})-2l(\lambda)} \lambda \mathcal{M}_\lambda(z_\mathcal{E}) \mathcal{M}_\lambda(z_\mathcal{F}^{-1}) \int_{z_\mathcal{E} \cup z_\mathcal{F}} f(z_\mathcal{E} \cup z_\mathcal{F}) + O_{r,n,h,f}(r^{2N}N^{3n+2}).
\]

Notice that we have exchanged the order of integration, which is permissible since we are only integrating continuous functions over compact spaces with respect to finite measures. Further notice that the terms only depend on \( l(\mathcal{E}) \), and not on the subsequence itself. Hence,

\[
\int_{U(N)} Eig(g) dg = \sum_{\lambda} \left( \frac{N}{2} \right)^{-2l(\lambda)} \lambda \mathcal{M}_\lambda(z_\mathcal{E}) \mathcal{M}_\lambda(z_\mathcal{F}^{-1}) \int_{z_\mathcal{E} \cup z_\mathcal{F}} f(z_\mathcal{E} \cup z_\mathcal{F}) + O_{r,n,h,f}(r^{2N}N^{3n+2}).
\]

Writing out the path integrals gives the desired formula. \( \square \)
In our opinion, the main interest of our explicit formula for eigenvalues of a random unitary matrix lies in the fact that its derivation has the same basic structure as the derivation of the explicit formula for zeros of $L$-functions in [RS96]. This similarity in structure might give a deeper insight into the conjectured connection between $L$-functions and characteristic polynomials from the unitary group. Rudnick and Sarnak’s explicit formula for zeros of $L$-functions is an application of the functional equation and the Euler product. Hence, our proof of the explicit formula for eigenvalues is based on two analogous properties of characteristic polynomials.

- The functional equation for $L$-functions used in [RS96] encodes a symmetry between the value of the completed $L$-function attached to some irreducible cuspidal automorphic representation $\pi$ of $GL_m$ over $\mathbb{Q}$ at the point $s$ and the value of the completed $L$-function associated to the contragredient of $\pi$ at the point $1 - s$. According to [CFK+05], the transformation $s \mapsto 1 - s$ corresponds to the transformation $z \mapsto z^{-1}$. Hence, it is reasonable that the equality
  \[ \Lambda_g(z) = \Lambda_{g^{-1}}(z^{-1}) \]
  plays the role of the functional equation in our derivation of the explicit formula for eigenvalues, where $g \in U(N)$ and the inverse $g^{-1}$ is analogous to the contragredient $\tilde{\pi}$.

- If we view the Euler product as a connector between $L$-functions and prime numbers, there is no hope of finding a random matrix theory analogue. However, if we view the Euler product as an explicit expression for the logarithmic derivative $\Lambda'(s,\pi)/\Lambda(s,\pi)$ that holds sufficiently far to the right of the critical line, then Theorem 5.3 is a possible analogue. Indeed, it provides an explicit expression for (the main term of) the average of logarithmic derivatives of completed characteristic polynomials that holds inside the unit circle.

  As the unit circle is the “critical line” for the completed characteristic polynomial $\Lambda_g(z)$ [CFK+05, p. 39], the unit disc (i.e. the inside of the unit circle) should correspond to either the half-plane to the left or the half-plane to the right of the critical line for completed $L$-functions. The substitute for the Euler product proposed above suggests that the unit disc is associated to the half-plane on the right-hand side. Another argument in support of this correspondence (which is also mentioned in [CFK+05]) is that under the assumption of the Riemann hypothesis, the zeros of $\zeta'(s)$ all lie to the right of the critical line (according to [LM74]), while the zeros of the derivative of any characteristic polynomial $\chi_g(z)$ with $g \in U(N)$ lie inside the unit circle (according to the Gauss-Lucas Theorem).

The proofs of the explicit formulæ (for zeros of $L$-functions and for eigenvalues) are both structured as follows: Consider the sum on the left-hand side of the equality to be proved:

\[
\sum_{\rho_\pi} h(\rho_\pi) - \delta(\pi) [h(0) + h(1)]
\]

where $\rho_\pi$ is over the nontrivial zeros of $L(s,\pi)$ and the second term vanishes unless $\pi$ corresponds to the $\zeta$-function, or

\[
\int_{U(N)} \sum_{1 \leq j_1, \ldots, j_n \leq N} h(\rho_{j_1}) \cdots h(\rho_{j_n}) f(\rho_{j_1}, \ldots, \rho_{j_n}) dg
\]

where $\{\rho_1, \ldots, \rho_N\}$ is the multiset of eigenvalues of $g \in U(N)$. Use Cauchy’s argument principle to express this sum over zeros as a contour integral. This
MIXED RATIOS OF CHARACTERISTIC POLYNOMIALS

results in an integral of the following abstract form:

\[ \frac{1}{2\pi i} \int_{\gamma_1} \frac{\Lambda'(s)}{\Lambda(s)} h(s) ds - \frac{1}{2\pi i} \int_{\gamma_2} \frac{\Lambda'(s)}{\Lambda(s)} h(s) ds \]

where \( \Lambda \) stands for a completed \( L \)-function or a random completed characteristic polynomial, and the contours \( \gamma_1 \) and \( \gamma_2 \) are vertical lines that are located to the right and to the left the critical line, respectively, or the contours \( \gamma_1 \) and \( \gamma_2 \) are circles about the origin that are located inside and outside the unit circle, respectively. In a next step, apply the functional equation to the integrand corresponding to the contour \( \gamma_2 \), which allows us to situate both contours to the right of the critical line/inside the unit circle. Now, the explicit formula is a consequence of the Euler product/Theorem 5.3.

The underlying structure of these derivations of explicit formulae might be the same, but the resulting formulas look quite different. The principal reason for this difference is that we have substituted the Euler product by an equality that does not carry any arithmetic information. Another obvious difference is that our explicit formula for eigenvalues provides an asymptotic expression for the sums over all \( n \)-tuples of eigenvalues (for \( n \geq 1 \)), whereas Rudnick and Sarnak’s explicit formula for zeros of \( L \)-functions provides an exact expression for the sum over all 1-tuples of zeros. It would be very interesting to investigate explicit formulae for sums of \( n \)-tuples of zeros of \( L \)-functions, whose proof follows the same structure. Such a proof would be based on an arithmetic expression for

\[ \prod_{\varepsilon \in E} \frac{\Lambda' (\varepsilon, \pi)}{\Lambda (\varepsilon, \pi)} \prod_{\varphi \in F} \frac{\Lambda' (\varphi, \tilde{\pi})}{\Lambda (\varphi, \tilde{\pi})} \]

where \( E \) and \( F \) are sets of complex numbers that lie sufficiently far to the right of the critical line. This arithmetic expression might even display the same combinatorial structure as our combinatorial formula for the average of products of logarithmic derivatives of completed characteristic polynomials (stated in Theorem 5.3).

REFERENCES

[BC14] A. Borodin and I. Corwin. Macdonald processes. *Probability Theory and Related Fields*, 158(1-2):225–400, 2014.

[BCC17] J. Bouttier, G. Chapuy, and S. Corteel. From Aztec diamonds to pyramids: steep tilings. *Transactions of the American Mathematical Society*, 369(8):5921–5959, 2017.

[BG06] D. Bump and A. Gamburd. On the averages of characteristic polynomials from classical groups. *Commun. Math. Phys.*, 265(1):227–274, 2006.

[BR85] A. Berele and J. B. Remmel. Hook flag characters and their combinatorics. *Journal of Pure Applied Algebra*, 35:225–245, 1985.

[BR87] A. Berele and A. Regev. Hook Young diagrams with applications to combinatorics and representations of Lie superalgebras. *Advances in Mathematics*, 64(2):118–175, 1987.

[Bre93] F. Brenti. Determinants of super-Schur functions, lattice paths, and dotted plane partitions. *Advances in Mathematics*, 98:27–64, 1993.

[Bum13] D. Bump. *Lie Groups*, volume 225 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2013.

[CFK+05] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith. Integral moments of \( L \)-functions. *Proc. London Math. Soc.*, 91(1):33–104, 2005.

[CFS05] J. B. Conrey, P. J. Forrester, and N. C. Snaith. Averages of ratios of characteristic polynomials for the compact classical groups. *International Mathematics Research Notices*, 2005(7):397–431, 2005.

[CS08] J. B. Conrey and N. C. Snaith. Correlations of eigenvalues and Riemann zeros. *Communications in Number Theory and Physics*, 2(3):477–536, 2008.

[Deh08] P.-O. Dehaye. Joint moments of derivatives of characteristic polynomials. *Algebra and Number Theory*, 2(1):31–68, 2008.

[KSS00] J. P. Keating and N. C. Snaith. Random matrix theory and \( \zeta (1/2 + it) \). *Commun. Math. Phys.*, 214(1):57–89, 2000.
D. E. Littlewood. Some properties of $S$-functions. *Proc. London Math. Soc.*, 40(1):49–71, 1936.

N. Levinson and H. L. Montgomery. Zeros of the derivatives of the Riemann zeta-function. *Acta Math.*, 133:49–65, 1974.

I. G. Macdonald. *Symmetric Functions and Hall Polynomials*. Oxford Mathematical Monographs. Oxford University Press, New York, second edition, 1995.

E. M. Moens and J. Van der Jeugt. A determinantal formula for supersymmetric Schur polynomials. *Journal of Algebraic Combinatorics*, 17(3):283–307, 2003.

N. Metropolis, G. Nicoletti, and G.-C. Rota. A new class of symmetric functions. *Math. Anal. Appl. Part B, Adv. in Math. Suppl. Stud.*, 7:563–575, 1981.

H. L. Montgomery. The pair correlation of zeros of the zeta function. *Proc. Sym. Pure Math.*, 24:181–193, 1973.

F. D. Murnaghan. The characters of the symmetric group. *American Journal of Mathematics*, 59(4):739–753, 1937.

T. Nakayama. On some modular properties of irreducible representations of the symmetric group I and II. *Japan. J. Math.*, 17:165–184, 411–423, 1940.

Helen Riedtmann. *Combinatorial Structures in Random Matrix Theory Predictions for $L$-Functions*. PhD thesis, Universität Zürich, arXiv, 2018.

Helen Riedtmann. Overlap identities for Littlewood-Schur functions. arXiv, 2018.

Z. Rudnick and P. Sarnak. Zeros of principal $L$-functions and random matrix theory. *Duke Math. J.*, 81(2):269–322, 1996.

B. E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001.