Plethysm of Schur Functions and Irreducible Polynomial Representations of the Complex General Linear Group

Rediet Abebe
Harvard College
Department of Mathematics

March 25, 2013

Contents

Abstract

We show the intimate connection Schur functions have with the irreducible polynomial representations of $GL(n, \mathbb{C})$. Following Schur's thesis, in Section 1, we prove that the characters of these representations are given by Schur functions.

In the second section, we elaborate on how a specific combinatorial operation on symmetric functions called plethysm gives the characters of the composition of irreducible representations. There is no general combinatorial description such as the Littlewood-Richardson rule, which is used to multiply Schur functions, that is known for the plethysm of Schur functions. There are, however, numerous algorithms and a few combinatorial formulas for special cases. In hopes of inspiring future work, we give a survey of the most notable results.

1 Irreducible Polynomial Representations of the $GL(n, \mathbb{C})$

The overarching goal of this section is to show that the characters of the irreducible polynomial representations of $GL(n, \mathbb{C})$ are given by Schur functions. We build up to this through a series of results, some of which hold in greater generality and are fundamental in the representation theory of finite groups. The exposition here follows the outline provided by Appendix 2 of [21], while the proofs are from [12] and [19].

1.1 Basic Representation Theory of $GL(V)$

We begin by introducing a few definitions and theorems from representation theory and algebraic combinatorics that will come in handy for the remainder of the section.

Definition 1. An **integer partition** is a weakly decreasing sequence of positive integers $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_k)$ such that $\sum_{i=1}^{k} \lambda_i = n$.

Here, we say that λ is a partition of n, which we denote by $\lambda \vdash n$. For each partition λ of n, there are corresponding polynomials known as *Schur polynomials*.

Definition 2. Schur polynomials are defined using the Jacobi-Trudi formula,

$$s_\lambda(x_1, \cdots, x_m) = \frac{\det((x_j)^{\lambda_i + m - i})_{1 \le i,j \le m}}{\det((x_j)^{m-i})_{1 \le i,j \le m}}.$$

The reader may recognize the denominator as the Vandermonde determinant. While this is the original and frequently cited definition of Schur polynomials, there are multiple definitions that are used interchangeably. We will see two of them in this exposition.[1]

Example 1.

$$\begin{aligned} s_{2,1,1}(x_1, x_2, x_3) &= \frac{1}{\Delta} \det \begin{bmatrix} x_1^4 & x_2^4 & x_3^4 \\ x_1^2 & x_2^2 & x_3^2 \\ x_1 & x_2 & x_3 \end{bmatrix} \\ &= x_1 x_2 x_3 (x_1 + x_2 + x_3) \\ &= e_1(x_1, x_2, x_3) e_3(x_1, x_2, x_3) \end{aligned}$$

Schur polynomials are a special class of symmetric polynomials. Recall that a symmetric polynomial is a polynomial invariant under interchanging its variables. There are various equivalent definitions of Schur polynomials. A more combinatorial one, which is presented in Appendix A, is used throughout Section 2. We will also prove that Schur polynomials form a $\mathbb{Z}$-basis for the ring of symmetric polynomials in the same section.

In the definition above, if $m < k$, then $s_\lambda(x_1, \cdots, x_m) = 0$. That is, the Schur polynomials vanish if the number of variables is smaller than the number of parts of the partition. Moreover, several important results about Schur polynomials, such as the fact that they are symmetric, are independent of the number of indeterminates. It is therefore important to consider Schur polynomials in sufficiently large variables.

We define a symmetric function of degree n to be a set of symmetric polynomials $p_m(x_1, \cdots, x_m)$ of degree n, one for each m, that satisfy,

$$p_m(x_1, \cdots, x_\ell, 0, \cdots, 0) = p_\ell(x_1, \cdots, x_\ell), \forall \ell < m.$$

[1] See Appendix A for a tableaux-theoretic definition.

Let Λ_n be the $\mathbb{Z}$-module of all such functions with integer coefficients. Schur functions are the symmetric functions corresponding to Schur polynomials. We also denote them by s_λ. We set,

$$\Lambda = \oplus_{n=0}^{\infty} \Lambda_n$$

to be the graded ring of symmetric functions. Therefore, identities that we prove for the ring of symmetric polynomials, such as $s_1(x_1, \cdots, x_\ell) = e_1(x_1, \cdots, x_\ell)$, can be extended to identities in Λ. Schur functions form a $\mathbb{Z}$-basis for the ring of symmetric functions. The polynomial case is what is proved in Appendix A. For a more complete treatment of symmetric functions, see [17].

The group $GL(n, \mathbb{C})$ is the set of all invertible $n \times n$ matrices with complex entries, which we identify with $GL(V)$, the group of invertible linear transformations of an n-dimensional complex vector space V.

Definition 3. A **linear representation** of $GL(V)$ is a vector space W together with a homomorphism,

$$\phi : GL(V) \to GL(W).$$

For convenience, we will refer to ϕ as the representation of $GL(V)$. The *dimension* of this representation is $\dim W$. In this thesis, we will only consider cases where the representation is finite-dimensional.

Definition 4. A linear representation is said to be a **polynomial representation** if, after choosing ordered bases of V and W, the entries of $\phi(A)$ are polynomials in the entries of A, for all $A \in GL(V)$.

Note that this is independent of the choice of bases since polynomials are preserved under linear change of variables. We define rational representations in a similar manner.

Definition 5. [2] A representation is **homogeneous** of degree m if $\phi(\alpha A) = \alpha^m \phi(A)$, where α is a non-zero, complex number.

Thus, given a representation ϕ of $GL(V)$, which is polynomial and homogeneous of degree m, each entry of $\phi(A)$ is a homogeneous polynomial of degree m. A few examples of homogeneous polynomial representations are:

- The *trivial representation* given by $\phi(A) = 1$, which has dimension 1 and degree 0,
- The *determinant representation* given by $\phi(A) = \det A$, which has dimension 1 and degree n and,
- The *defining representation* given by $\phi(A) = A$, which has dimension n and degree 1.

where $A \in GL(V)$ and $\dim V = n$ in each case. For a less trivial example, consider the following representation:

Example 2. Take the representation $\phi : GL(2, \mathbb{C}) \to GL(3, \mathbb{C})$ given by,

$$\phi \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 & 2ab & b^2 \\ ac & ad + bc & bd \\ c^2 & 2cd & d^2 \end{pmatrix}.$$

We can easily check that this is a homomorphism. This representation is of degree 2 and dimension 3.

[2] Note that some texts use the term "degree" to mean the dimension of the representation.

Example 3. The map $\phi(A) = (A^{-1})^t$ also defines a representation. We can see that this is a representation since the $(I_n^{-1})^t = I_n$, where I_n is the $n \times n$ identity matrix and,

$$\phi(AB) = \left((AB)^{-1}\right)^t = \left(B^{-1}A^{-1}\right)^t = \left(A^{-1}\right)^t \left(B^{-1}\right)^t = \phi(A)\phi(B)$$

This is a homogeneous representation of dimension n and degree -1, since $AA^{-1} = I$, and so $\phi(\alpha A) = \alpha^{-1}\left(A^{-1}\right)^t$. However, it is a rational representation, not a polynomial representation.

Definition 6. Given a representation $\phi : GL(V) \to GL(W)$, a **subrepresentation** of ϕ is a subspace $U \subset W$, which is invariant under the action of $GL(V)$.

If such a subrepresentation exists, we denote the restriction of ϕ to U by $\phi\,|_U$.

Definition 7. A representation $\phi : GL(V) \to GL(W)$ is said to be **irreducible** if there is no proper, nonzero subspace of W invariant under the action of $GL(V)$. It is said to be reducible otherwise.

We have already seen irreducible representations in the examples above. For instance, the determinant and trivial representation are clearly irreducible since they have dimension one.

A representation is said to be *completely reducible* if it can be decomposed as a direct sum of irreducible representations

$$\phi = \phi\,|_{W_1} \oplus \cdots \phi\,|_{W_\ell},$$

where $W = W_1 \oplus \cdots \oplus W_\ell$.

Proposition 1. *Every rational representation of $GL(V)$ is completely reducible. That is, given a representation $\phi : GL(V) \to GL(W)$, every $GL(V)$-invariant subspace U of W has a $GL(V)$-invariant complement U^c such that $W = U \oplus U^c$.*

This is an essential result in representation theory, which we state without proof. In fact, it holds in greater generality, for all semi-simple algebraic groups and finite groups. The proof of this, which makes use of Schur's lemma (below), can be found in [9].

Definition 8. Two representations $\phi_1 : GL(V) \to GL(W_1)$ and $\phi_2 : GL(V) \to GL(W_2)$ are said to be **equivalent** if there exists a vector space isomorphism $\alpha : W_1 \to W_2$ such that,

$$\alpha \circ \phi_1(A) \circ \alpha^{-1} = \phi_2(A),$$

for all $A \in GL(V)$.

Lemma 1 (Schur's Lemma)**.** *If $\phi_1 : GL(V) \to GL(W_1)$ and $\phi_2 : GL(V) \to GL(W_2)$ are irreducible representations of $GL(V)$ and $\psi : W_1 \to W_2$ is a linear map, which respects the action of G, then either ψ is an isomorphism or $\psi = 0$. Furthermore, if $W_1 = W_2$, then $\psi = \lambda \cdot I$ for some $\lambda \in \mathbb{C}$ and I the identity.*

Given two equivalent representations ϕ_1 and ϕ_2 of $GL(V)$, by definition $\operatorname{tr}\phi_1(A) = \operatorname{tr}\phi_2(A)$ for $A \in GL(V)$. This quantity only depends on the conjugacy class of an element A and the isomorphism class of ϕ. It is a crucial invariant, which characterizes the representation.

Definition 9. The function $\chi : GL(V) \to \mathbb{C}$, given by $\chi(A) = \operatorname{tr}\phi(A)$ is called the **character** of the representation.

The characters of the irreducible polynomial representations of $GL(V)$ are of interest since they are Schur functions. This offers a fascinating connection between representation theory of linear groups and algebraic combinatorics and is the main result of Section 1.

We have already noted that every rational representation of the general linear group is a direct sum of irreducible representations; thus, the character of such a representation is the sum of the characters of the irreducible components. Other operations on irreducible representations include tensor product or composition. Tensor products of representations correspond to the products of characters. Products of Schur functions have a complete combinatorial description via the Littlewood-Richardson coefficients. This is presented in Appendix B. The operation needed to obtain composition, called plethysm, is the main focus of this thesis and will be discussed in Section 2.

1.2 Characters of the Representations of $GL(V)$

The goal of this section is to construct and understand the characters of the irreducible polynomial representations of the complex general linear group. We will illustrate the intricate relationship these characters have with Schur functions.

Proposition 2. *Let $\phi : GL(V) \to GL(W)$ be a homogeneous rational representation such that $\dim V = n$ and $\dim W = m$. Then, there exists a multiset of m Laurent monomials in n variables,*

$$\mathcal{M}_\phi = \{f_1(x_1, \cdots, x_n), f_2(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n)\}$$

such that for $A \in GL(V)$, if A has eigenvalues $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, then $\phi(A)$ has eigenvalues

$$\{f_1(\lambda_1, \lambda_2, \cdots, \lambda_n), \cdots, f_m(\lambda_1, \lambda_2, \cdots, \lambda_n)\}.$$

These monomials are unique (up to reordering). If ϕ is a polynomial representation, then the Laurent monomials do not have negative exponents.

Before going on to prove this proposition, let us consider a quick example. Take the representation in Example 2. Let $A \in GL(2, \mathbb{C})$ have eigenvalues λ_1 and λ_2. Then, our multiset of Laurent monomials is,

$$\mathcal{M}_\phi = \{x_1^2, x_1x_2, x_2^2\}.$$

Thus, the eigenvalues of $\phi(A)$ are $\{\lambda_1^2, \lambda_1\lambda_2, \lambda_2^2\}$. This is something that we can check easily from the matrix on the right hand side of the homomorphism.

Proof. Fix a basis of the vector space V and let D be the commuting set of all diagonal matrices with respect to this basis. We claim that for $d \in D$, $\phi(d)$ is a diagonalizable element of $GL(W)$. If not, then by Jordan normal form we can find a generalized eigenspace for $\phi(d)$ with eigenvalue θ containing two linearly independent vectors w_1, w_2 such that $\phi(d)$ preserves their span and the matrix with respect to $\{w_1, w_2\}$ is,

$$\theta \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

for some $\theta \in \mathbb{C}$.

Now, for all $z \in \mathbb{Z}$, the matrix with respect to w_1, w_2 of $\phi(d^z)$ is,

$$\theta^z \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}.$$

However, this cannot be the case for a rational representation ϕ. If $d = \operatorname{diag}(d_1, \cdots, d_n)$, then rationality means that there is a fixed rational function P such that $P(d_1^z, \cdots, d_n^z)$ equals the matrix coefficient $z\theta^z$. So, $P(d_1^z, \cdots, d_n^z)/\theta^z = z$, which cannot be the case for a fixed rational function P, giving us a contradiction.

Now, $\phi(D)$ is a commuting set of diagonalizable matrices, hence they are all simultaneously diagonalizable. That is, there is a basis of W with respect to which all elements of $\phi(D)$ are diagonal. That is, there are functions $f_i : D \to \mathbb{C}^*$, for $1 \leq i \leq m$ such that,

$$\phi(\operatorname{diag}(\lambda_1, \cdots, \lambda_n)) = \operatorname{diag}(f_1, \cdots, f_m),$$

for all $\operatorname{diag}(\lambda_1, \cdots, \lambda_n) \in D$. Now, f_i is a rational function, since ϕ is rational, and is multiplicative in $\operatorname{diag}(\lambda_1, \cdots, \lambda_n)$, since ϕ is a representation. It must therefore be a Laurent monomial in n variables.

Now, let B be the set of elements $A \in GL(V)$ such that the eigenvalues of $\phi(A)$ are $f_i(\lambda_1, \cdots, \lambda_n)$, for $1 \leq i \leq m$. Then, B contains all diagonal elements, hence all diagonalizable elements, which are dense in $GL(V)$. But, B is also closed in $GL(V)$ so B is all of $GL(V)$. This proves the eigenvalues of $\phi(A)$ are the Laurent monomials as claimed.

□

These multisets give us a more convenient definition of the characters of the polynomial representations of $GL(V)$, which we will use for the remainder of the section.

$$\operatorname{char} \phi = \operatorname{tr} \phi(A) = f_1 + \cdots + f_m$$

The character of any polynomial representation is a symmetric function in n variables. It is known that the ring of symmetric functions has a $\mathbb{Z}$-basis in the monomial symmetric functions, the elementary symmetric functions, complete homogeneous functions and Schur functions, and a $\mathbb{Q}$-basis in the power-sum functions. The proof of each of these can be found in [8]. Thus, we can write these characters as a linear combination of Schur functions (or any of the other basis). As it turns out, the Schur functions form a convenient basis since the Schur functions themselves correspond to the irreducible polynomial representations of $GL(V)$.

Proposition 3. *An irreducible polynomial representation $\phi : GL(V) \to GL(W)$ is equivalent to a homogeneous polynomial representation in n variables and degree m, where* $\dim V = n$ *and* $\dim W = m$.

Proof. Take $x \in \mathbb{C}$. Then, $\phi(xI_n) = x^k C_o + x^{k-1} C_1 + \cdots + C_k$, where each C_i is an $m \times m$ matrix with complex entries. Because ϕ is a homomorphism,

$$\phi(xI_n)\phi(yI_n) = \phi(xyI_n).$$

Thus, $C_i \cdot C_j = 0$ if $i \neq j$ and $C_i \cdot C_j = C_i$ if $i = j$. These C_i's are commuting matrices. Matrices are simultaneously diagonalizable if and only if they commute.[3] From this, it follows that there exists a matrix M such that $M^{-1} \cdot C_i \cdot M$ is a block matrix with block entries from left to right being $(O_{r_0}, O_{r_1}, \cdots, O_{r_{i-1}}, I_{r_i}, O_{r_{i+1}}, \cdots, O_{r_k})$, where each of the O_ℓ's are $\ell \times \ell$ matrix of all zeros.

Thus, $M^{-1} \cdot \phi(xI_n) \cdot M$ is the block matrix of where the block entries are $(x^k I_{r_0}, x^{k-1} I_{r_1}, \cdots, I_{r_k})$. Now, for any $A \in GL(V)$, write the matrix as follows:

$$\begin{bmatrix} \phi_{0,0}(A) & \cdots & \phi_{0,k}(A) \\ \vdots & \ddots & \vdots \\ \phi_{k,0}(A) & \cdots & \phi_{k,k}(A) \end{bmatrix}$$

[3] The proof of this can be found in [3].

where $\phi_{i,j}$ is an $r_i \times r_j$ matrix. Now, $\phi(A)\phi(xI_n) = \phi(xI_n)\phi(A)$. Thus,

$$x^{k-i}\phi_{i,j}(A) = x^{k-j}\phi_{i,j}(A), i \neq j$$

and thus, $\phi_{i,j}(A) = 0, i \neq j$. Moreover, $\phi_{i,i}(xA) = x^{k-i}\phi_{i,i}(A)$, and so $\phi_{i,i}(x) : GL(V) \to GL(W_{r_i})$, for W_{r_i} an r_i-dimensional vector space, is a homogeneous polynomial representation of degree $k-i$ and ϕ is a direct sum of the $\phi_{i,i}$'s ignoring the zero terms.

□

Lemma 2. *Given a homogeneous polynomial representation $\phi : GL(n, \mathbb{C}) \to GL(m, \mathbb{C})$ of order k, $\operatorname{char}\psi = ch\ \chi$ where $ch\ \chi$ is the Frobenius characteristic.*

Proof. Let V be an n-dimensional vector space. The group $GL(V)$ acts on the tensor space $V^{\otimes k}$ by

$$A \cdot (v_1 \otimes \cdots \otimes v_k) = A \cdot v_1 \otimes \cdots \otimes A \cdot v_k, \text{ where } A \in GL(V),$$

and S_k acts on it by

$$\omega \cdot (v_1 \otimes \cdots v_k) = v_{\omega^{-1}(1)} \otimes \cdots v_{\omega^{-1}(k)}, \text{ where } \omega \in S_k$$

These actions by $GL(V)$ and S_k commute. Therefore, we have an action of $S_k \times GL(V)$ on the tensor space $V^{\otimes k}$. □

Lemma 3. *Let $\omega \in S_k$, $A \in GL(V)$ with eigenvalues $\lambda = \{\lambda_1, \cdots, \lambda_n\}$, and ω has cycle type $\rho = (\rho_1, \rho_2, \cdots)$. Thus, $(\omega, A) \in S_k \times GL(V)$ and $\operatorname{tr}(\omega, A)$ is the trace of this action. Then,*

$$\operatorname{tr}(\omega, A) = p_\rho(\lambda)$$

where p_ρ denotes the power sum polynomial.

Proof. We will prove this theorem for a particular $\omega = (i_1, i_2, \cdots, i_{\rho_1})(i_{\rho_1+1}, \cdots, i_{\rho_1+\rho_2})\cdots$. The proof works in general since we are working with class functions and it doesn't matter what ω you choose.

Let $A = (a_{ij})$. Then, the $((i_1, \cdots, i_k), (j_1, \cdots, j_k))$ entry of the matrix of A acting on $V^{\otimes k}$ is,

$$a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_k j_k}, \text{ such that } 1 \leq i_\ell, j_\ell \leq n.$$

Therefore, the corresponding entry of (ω, A) is

$$a_{i_1 j_2} a_{i_2 j_3} \cdots a_{i_{\rho_1} j_1} a_{i_{\rho_1+1} j_{\rho_1+2}} \cdots a_{i_{\rho_1+\rho_2}, j_{\rho_1+1}} \cdots$$

Hence,

$$\begin{aligned} \operatorname{tr}(\omega, A) &= \sum_{i_1, \cdots, i_k} (a_{i_1 i_2} a_{i_2 i_3} \cdot a_{i_{\rho_1} i_1})(a_{i_{\rho_1+1} i_{\rho_1+2}} \cdots a_{i_{\rho_1+\rho_2} i_{\rho_1+1}}) \cdots \\ &= (\operatorname{tr} A^{\rho_1})(\operatorname{tr} A^{\rho_2}) \cdots \\ &= p_\rho(\lambda). \end{aligned}$$

□

Lemma 4. *Let*

$$q = \frac{1}{k!} \sum_{\omega \in S_k} (n^{c(\omega)})^2 = \frac{1}{k!} \sum_{\omega \in S_k} (\operatorname{tr}(\omega, I_n))^2$$

where $c(\omega)$ is the number of cycles of ω. Then, $q = \binom{n^2+k-1}{k}$.

Proof. This is a direct consequence of the following identity, which can be found in [21]:

$$x(x+1)\cdots(x+k-1) = \sum_{\omega \in S_k} x^{c(\omega)}.$$

□

Let ξ be the character of S_k acting on the tensor space $V^{\otimes k}$ by,

$$\omega(v_1 \otimes \cdots \otimes v_k) = v_{\omega^{-1}(a)} \otimes \cdots v_{\omega^{-1}(k)}$$

Let $\xi = \sum_{\lambda \vdash k} g_\lambda \chi^\lambda$, where $g_\lambda \in \mathbb{N}$. Therefore g_λ is the multiplicity of the irreducible representations. Let $[\omega A]$ be the matrix of $(\omega, A) \in S_k \times GL(V)$ acting the tensor space $V^{\otimes k}$. We assume that,

$$[\omega] = \amalg_{\lambda \vdash k} \begin{bmatrix} R^\lambda & & & \\ & R^\lambda & & \\ & & \ddots & \\ & & & R^\lambda \end{bmatrix}$$

where $\omega \mapsto R^\lambda$ is the irreducible representation of S_k with character χ^λ.

Now, let M^λ be an S_k irreducible component of the tensor space $V^{\otimes k}$ with character χ^λ. Likewise, let N^λ be the λ-isotypic component. For $A \in GL(V)$, $A : M^\lambda \to V^{\otimes k}$. Since $A(\omega v) = \omega(Av)$ it follows that $A \cdot M^\lambda$ is invariant under the action of S_k and so is isomorphic to M^λ. Thus, $A \cdot M^\lambda \subseteq N^\lambda$, and so $A \cdot N^\lambda = N^\lambda$.

Consider the case where A is restricted to N^λ. We write,

$$[A]_\lambda = \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,g_\lambda} \\ A_{2,1} & A_{2,2} & \cdots & \\ & & \ddots & \\ & & & \ddots \end{bmatrix}$$

Then, $[A\omega] = [\omega A]$.

$$\begin{bmatrix} R^\lambda & & \\ & R^\lambda & \\ & & \ddots \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} & \\ A_{21} & A_{22} & \\ & & \ddots \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \\ A_{21} & A_{22} & \\ & & \ddots \end{bmatrix} \begin{bmatrix} R^\lambda & & \\ & R^\lambda & \\ & & \ddots \end{bmatrix}$$

Thus, $R^\lambda A_{i,j} = A_{i,j} R^\lambda$. But the only matrices commuting with an irreducible representation are scalars, which is shown in Proposition 4 of [20]. Therefore, $A_{i,j} = \alpha_{i,j} I_{f_\lambda}$ for some $f_{i,j} \in \mathbb{C}$, where f_λ gives the dimension of the space R^λ. Thus,

$$[A]_\lambda = \begin{bmatrix} f_{1,1}I & f_{1,2}I & \cdots \\ f_{2,1}I & \ddots & \\ & & \ddots \end{bmatrix} \cong [f_{i,j}] \otimes I \cong I \otimes [f_{i,j}]$$

Therefore, the map $A \mapsto [f_{i,j}]$ is a polynomial representation of $GL(V)$ of degree g_λ and

$N^\lambda \cong M^\lambda \otimes F^\lambda$, where F^λ affords $\omega A(v \otimes v') = \omega v \otimes Av'$.

We are now ready to state and prove the theorem that is at the heart of this section.

Theorem 1 (The Main Result)**.** *The irreducible polynomial representations, ϕ^λ of $GL(V)$, can be indexed by partitions λ of length at most $n := \dim V$, such that*

$$\operatorname{char} \phi^\lambda = s_\lambda(x_1, \cdots, x_n).$$

Using this theorem, one can find the decomposition of any polynomial representations of the general linear group into its irreducible components by writing its character as a sum of corresponding Schur functions, which is a phenomenal result. We assume that the irreducible representations are inequivalent. Using this, we will prove the theorem as a series of claims.

Proposition 4. *Let G be a group and K be an algebraically closed field. Let $V_1, \cdots, V_k$ be pairwise non-isomorphic irreducible matrix representations of G over this field. Take any $g \in G$, and let*

$$V_r(g) = (f_{ij}^{(r)}(g))_{1 \le i,j \le n_r}.$$

Then, the coordinate functions, $\{f_{ij}^{(r)} : 1 \le i, j \le n_r, 1 \le r \le k\}$, are all linearly independent.

The proof of this proposition can be found in [4] (Theorem 27.8).

Claim 1. *If ϕ is an irreducible polynomial representation of $GL(V)$ of order k, then $\phi = \phi^\lambda$ for some λ.*

Proof. Let ϕ be of degree ℓ. Now, assume the claim does not hold. Then, the matrix entries of the ϕ^λ's and ϕ would be linearly independent by the theorem above. But the total number of such entries is $\sum (g^\lambda)^2 + \ell^2 = \binom{n+k-1}{k} + \ell^2$, which is greater than the dimension of homogeneous polynomials of degree k in n^2 variables. □

Claim 2. $\operatorname{char} \phi^\lambda = s_\lambda$

Proof. Finally, we are ready to compute the character of ϕ^λ. Take $(\omega, A) \in S_k \times GL(V)$ and $\operatorname{tr}(\omega, A)$ as in the above case. Since $V = \amalg_\lambda S^\lambda \otimes F^\lambda$, we find that

$$\operatorname{tr}(\omega, A) = \sum_\lambda \chi^\lambda(\omega) \operatorname{tr}(\phi^\lambda(A))$$

Let $\theta = (\theta_1, \cdots, \theta_n)$ be the eigenvalues of A. We have already shown that $\operatorname{tr}(\omega, A) = p_{\rho(\omega)}(\lambda)$. It then follows that,

$$p_{\rho(\omega)} = \sum_\lambda \chi^\lambda(\omega)(\operatorname{char} \phi^\lambda)(\theta)$$

But, we know that

$$p_{\rho(\omega)} = \sum_\lambda \chi^\lambda(\omega) s_\lambda,$$

which uses a standard fact of symmetric functions from Chapter 7 of [21]. Since the χ^λ are linearly independent, $\operatorname{char} \phi^\lambda = s_\lambda(x_1, \cdots, x_n)$. □

Corollary 1. *Let $\lambda \vdash k$, then $g_\lambda \ne 0$, i.e., χ^λ appears in the S_k action on $V^{\otimes k}$, if and only if λ is a partition of length at most n.*

This completes the proof of the main theorem. Finding the irreducible polynomial representations of $GL(V)$ is of paramount importance since it also gives us the irreducible rational representations. In particular, every irreducible rational representation ϕ of $GL(V)$ is of the form $\phi(A) = (\det A)^m \phi'(A)$, for some $m \in \mathbb{Z}$ and some irreducible polynomial representation ϕ' of $GL(V)$. The corresponding characters are hence related by $\operatorname{char}\phi = (x_1 \cdots x_n)^m \operatorname{char}\phi'$. This is also discussed in [21].

1.3 Direct Sum, Tensor Products and Composition

Given two irreducible representations of $GL(V)$, we can perform several operations to get new representations. In this section, we are concerned with how we can obtain the characters of these new representations. We have already mentioned that the character of a representation is the sum of the characters of its irreducible components. To see this, consider the definition of the characters for the representations of $GL(V)$ via the multiset of Laurent monomials. The following lemma is an immediate consequence.

Lemma 5. *If a representation ϕ can be decomposed as the direct sum of rational representations, $\phi = \phi_1 \oplus \phi_2 \oplus \cdots \oplus \phi_\ell$ then,* $\operatorname{char}\phi = \sum_{i=1}^{\ell} \operatorname{char}\phi_i$.

Likewise, given two polynomial representations, $\phi_i : GL(V) \to GL(W_i)$, for $i = 1, 2$, we can take the tensor product representation,

$$\phi_1 \otimes \phi_2 : GL(V) \to GL(W_1 \otimes W_2)$$

defined by $A \cdot (w_1 \otimes w_2) = (A \cdot w_1) \otimes (A \cdot w_2)$. This extends to all of $W_1 \otimes W_2$ by linearity. Take the following linear transformations $B_1 : W_1 \to W_1$ and $B_2 : W_2 \to W_2$, with eigenvalues $\{\lambda_1^{B_i}, \lambda_2^{B_i}, \cdots, \lambda_n^{B_i}\}, i = 1, 2$. The eigenvalues of $B_1 \otimes B_2$ are $\{\lambda_i^{B_1}\lambda_j^{B_2}, 1 \le i, j \le n\}$. Thus,

$$\operatorname{char}(\phi_1 \otimes \phi_2) = \operatorname{char}\phi_1 \cdot \operatorname{char}\phi_2.$$

In the previous subsection, we have seen that the characters of the irreducible polynomial representations are Schur functions. In Appendix B, we show how to multiply the Schur functions using the Littlewood-Richardson rule.[4] In particular, if we are given that λ, μ, ν are partitions of length at most $n = \dim V$, then the multiplicity of ϕ^λ in $\phi^\mu \otimes \phi^\nu$ is the Littlewood-Richardson coefficient,

$$\langle s_\lambda, s_\mu s_\nu \rangle = c_{\mu\nu}^\lambda.$$

The above two operations on the representations of $GL(V)$ have corresponding operations on the characters that are straightforward to compute. The third operation, composition of representations, however, is not as easy to compute. Consider the following polynomial representations:

$$GL(V) \xrightarrow{\phi} GL(W) \xrightarrow{\psi} GL(Y).$$

The composition $\psi\phi : GL(V) \to GL(Y)$ also defines a polynomial representation.

Suppose that $A \in GL(V)$ has eigenvalues $\lambda_1, \cdots, \lambda_n$. We know that the eigenvalues of $\phi(A)$ are monomials λ^a for $x^a \in \mathcal{M}_\phi$. denote the monomials λ^a by $\lambda^{a^1}, \cdots, \lambda^{a^n}$. We note a similar result for ψ. That is, if the eigenvalues of B are $\xi_1, \cdots \xi_N$, then the eigenvalues of $\psi(B)$ are the monomials ξ^b for $x^b \in \mathcal{M}_\psi$. Thus, the eigenvalues of $\psi\phi(A)$ are the monomials $x^b|_{x_i = \lambda^{a^1}}$, that is x^b evaluated at $x_i = \lambda^{a^i}$, where x^b is an element in $\mathcal{M}_\psi$. Therefore, if f is the characteristic

[4] This gives one of many proofs that the Littlewood-Richardson coefficients are non-negative.

function of ϕ and g is the characteristic function of ψ, then,

$$\mathrm{char}(\psi\phi) = g(\lambda^{a^1}, \cdots, \lambda^{a^N}).$$

This operation is, in fact, not limited to characters of representations. We define such an operation for all symmetric functions.

Definition 10. Let f be a symmetric function, which is a sum of monomials, $f = \sum_{i\geq 1} x^{a^i}$, and g be another symmetric function. Then, the **plethysm** $g[f]$ (sometimes denoted by $f \circ g$) is defined as

$$g[f] = g(x^{a^1}, x^{a^2}, \cdots)$$

From the definition of plethysm, it is not hard to see that,

$$(af + bg)[h] = af[h] + bg[h], a, b \in \mathbb{Q}$$
$$(fg)[h] = f[h] \cdot g[h]$$

Using this definition we can simplify the above formula regarding the characters of the composition of representations as

$$\mathrm{char}(\psi\phi) = (\mathrm{char}\,\psi)[\mathrm{char}\,\phi]$$

Example 4. A simple example of this is to consider any $f \in \Lambda$ and the elementary symmetric function $e_1(x_1, x_2, \cdots) = x_1 + x_2 + \cdots$. Here, we get the identity $f[e_1] = f$. Assume that we have the following representations:

$$GL(V) \xrightarrow{\phi} GL(W) \xrightarrow{\psi} GL(Y)$$

where $\mathrm{char}(\phi) = e_1$ and $\mathrm{char}(\psi) = f$. Then, $(\mathrm{char}\,\psi)[\mathrm{char}\,\phi] = f[e_1] = f$. This result is clear from a representation theory perspective since the character, e_1, corresponds to the defining representation.

Theorem 2. *Let f and g be an $\mathbb{N}$-linear combination of Schur functions. Then, the plethysm $g[f]$ is also an $\mathbb{N}$-linear combination of Schur functions.*

Proof. Given a positive integer m, $f(x_1, x_2, \cdots, x_m)$ is the character of the polynomial representation $\phi : GL(V) \to GL(W)$, where V is of dimension m and W is of dimension n. Likewise, $g(x_1, x_2, \cdots, x_n)$ is the character of a polynomial representation, $\psi : GL(W) \to GL(Y)$. Therefore, the plethysm $g[f](x_1, \cdots, x_n)$ is the character of the composition $\psi\phi$, which we know has to be an $\mathbb{N}$-linear combination of Schur functions. Now, take $n \to \infty$ to see the desired result. □

While we have this representation theoretic proof that falls out of the above definition, there is no known combinatorial proof. i.e. we do not have a combinatorial rule (such as the Littlewood-Richardson rule) for expanding the plethysm of $s_n[s_m]$. In the next section, we will present a few cases that are known in hopes of understanding the composition of such representations better.

2 Plethysm of Schur Functions

It is in the context provided in Section 1 that plethysm arose; it was introduced by Littlewood while discussing the composition of the polynomial representations of $GL(n, \mathbb{C})$ in [14]. The problem of expressing the coefficients of the plethysm of Schur functions using combinatorial rules remains one of the most fundamental open problems in the theory of symmetric functions. While a complete combinatorial rule has yet to be found, there are numerous algorithms and a few combinatorial descriptions for some special cases. We focus on combinatorial results and give a survey of the most notable ones in this section.

2.1 Basic Combinatorial Tools

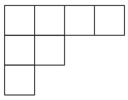
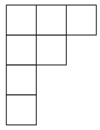
We begin by introducing some definitions, notations, and combinatorial objects that will be handy for the remainder of the section.

Definition 11. A **Young diagram** is a finite collection of n boxes arranged in left-justified rows such that the row lengths are weakly decreasing.

We denote the number of boxes in a Young diagram λ by $|\lambda|$. Young diagrams give a convenient way to represent partitions of n. From here on, λ may refer to either the partition of n or the Young diagram corresponding to this partition, which will be clear from the context.

Given a partition λ we obtain its *conjugate*, denoted by λ^T, by reflecting the corresponding Young diagram along the line $y = -x$.

Example 5. Let $n = 7$ and $\lambda = (4, 2, 1)$. Then,

$\lambda =$  $\lambda^T =$ 

Definition 12. A **Young tableau** is a Young diagram whose boxes are filled with elements of a totally ordered set, usually positive integers.

A Young tableau is said to be *semi-standard* if its entries are strictly increasing in the columns, and weakly increasing in the rows. In this paper, all the Young tableaux we work with have this property, so we will just drop this description. The set of semi-standard Young tableaux of shape λ with entries from a set S is denoted by $SSYT^{\lambda}(S)$.

Definition 13. A **standard Young tableau** is a semi-standard Young tableau of size n that is obtained by filling in the boxes with the numbers from the set $\{1, 2, \cdots, n\} = [n]$, where each number is used exactly once.

Example 6. Two examples of Young tableaux that we can get from the Young diagram above are,

$$\lambda = \begin{array}{|c|c|c|c|}\hline 1 & 2 & 2 & 3 \\ \hline 2 & 3 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array} \qquad \mu = \begin{array}{|c|c|c|c|}\hline 1 & 2 & 5 & 6 \\ \hline 3 & 4 \\ \cline{1-2} 7 \\ \cline{1-1}\end{array}$$

λ is a not a standard Young tableau while μ is.

Definition 14. Given two partitions, $\lambda = (\lambda_1, \lambda_2, \cdots)$ and $\mu = (\mu_1, \mu_2, \cdots)$, we say that λ **majorizes** μ if $\sum_{i=1}^{\ell} \lambda_i \geq \sum_{i=1}^{\ell} \mu_i$, for all ℓ.

Definition 15. A **skew (Young) diagram**, denoted as λ/μ, is a diagram obtained using two partitions λ and μ and taking the boxes of λ not contained in μ.

Note that for this definition to work, λ must majorize μ. Moreover, unless μ is the empty partition, the rows of the skew diagram will no longer be left-justified. The above definitions (Young tableau, semi-standardness, and standardness) generalize in the obvious way to the case of skew diagrams.

Example 7. The following is a skew standard Young tableaux of shape $(4, 2, 1)/(2)$:

$$\begin{array}{cccc} & & \boxed{4} & \boxed{5} \\ \boxed{1} & \boxed{3} & & \\ \boxed{2} & & & \end{array}$$

Definition 16. An **integer composition** is any sequence of positive integers $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k)$ such that $\sum_{i=1}^{k} \alpha_i = n$.

Definition 17. The **weight** of a tableau T, denoted by $\mathrm{wt}(T)$, is the sequence $(\mu_1, \mu_2, \cdots)$, where μ_k is the number of boxes in T that contain k.

We will mostly work with tableaux that are filled in with positive integers. Thus, we assume that k is minimal. That is, we are using all the integers in $[k]$ for some k. The tableaux in Example 6 have weights $(1, 3, 2, 1)$, and $(1, 1, 1, 1, 1, 1, 1)$, respectively. In fact, standard Young tableaux are precisely semi-standard Young tableaux of weight $(1, 1, \cdots, 1)$.

Given a monomial $x_1^{\mu_1} x_2^{\mu_2} \cdots x_k^{\mu_k}$, its weight is $\mu = (\mu_1, \mu_2, \cdots, \mu_k)$. Thus, we have a correspondence between tableaux of weight μ and monomials $x^T = x_1^{\mu_1} x_2^{\mu_2} \cdots x_k^{\mu_k}$. We use this to provide a more convenient, tableaux-theoretic definition of Schur polynomials:

$$s_\lambda(x_1, \cdots, x_m) = \sum_{T \in SSYT^\lambda([m])} x^T,$$

where $|\lambda| = n$ and $x^T = x^{\mathrm{wt}(T)} = x_1^{\mu_1} x_2^{\mu_2} \cdots x_m^{\mu_m}$, where $(\mu_1, \mu_2, \cdots, \mu_m)$ is a composition of n.

Definition 18. Given a tableau T, the **word** of the tableau, which we denote by $w(T)$, is obtained by reading the entries from left to right then bottom to top.

For instance, tableaux in Example 6 have words $(4, 2, 3, 1, 2, 2, 3)$ and $(7, 3, 4, 1, 2, 5, 6)$. If the shape of the tableau is known, then T is uniquely recoverable from its word by looking at where the sequence of alphabets decreases. Therefore, we can identify $SSYT^\lambda$ with a set of words, which we denote by $\mathcal{W}(\lambda)$.

After fixing a shape λ, we can put a total ordering on the set of all Young tableaux of this shape using reverse lexicographic ordering of the corresponding words.

Definition 19. Given two partitions λ and μ, the **Kostka coefficient** $K_{\lambda\mu}$ is a number is the number of semi-standard Young tableaux of shape λ and weight μ. It is can also be expressed as,

$$s_\lambda = \sum_{\mu \vdash n} K_{\lambda\mu} m_\mu.$$

This is a well known number, and one can derive the first part of the definition from the second part. Clearly, if $\lambda < \mu$ in lexicographic order, then this coefficient is zero. This number is always non-negative, which we note using the first definition.

Similarly, we can define the *inverse Kostka coefficient* $K^{-1}_{\lambda,\mu}$ as,

$$m_\lambda = \sum_{\mu \vdash n} K^{-1}_{\lambda,\mu} s_\mu.$$

This also has (albeit a more complicated) combinatorial description shown in [2]. From this equality, we note that,

$$h_\lambda = \sum_{\mu \vdash n} K_{\lambda,\mu} s_\mu$$

which is also 0 if $\lambda < \mu$ in lexicographic order.

2.2 Quasisymmetric Expansion of Plethysm

We will give an explicit expansion of the plethysm of Schur functions in terms of Gessel's fundamental quasisymmetric functions as presented in [16]. Using this, we are able to compute plethysms of the form $s_2[s_n]$. We begin by introducing a few tools and definitions relevant to this section.

Given a composition $\alpha = (\alpha_1, \alpha_2, \cdots \alpha_s)$ of a positive integer k, the gap of the composition is defined as,

$$\text{gap}(\alpha) = (\alpha_1, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2 + \alpha_3, \cdots, \alpha_1 + \cdots + \alpha_{s-1}).$$

This gives us a one-to-one correspondence between compositions of k and subsets of the set $\{1, 2, \cdots, k-1\}$, where the forward direction is given by the definition and the backward direction sends a subset $\{i_1 < i_2 < \cdots < i_s\}$ to the composition $(i_1, i_2 - i_1, \cdots, i_s - i_{s-1}, k - i_s)$.

Definition 20. Gessel's fundamental **quasisymmetric polynomial**, which is indexed by a composition α and with coefficients in $\mathbb{R}$, is defined as,

$$Q_\alpha(x_1, x_2, \cdots, x_n) = Q_{\text{gap}(\alpha)}(x_1, x_2, \cdots, x_n) = \sum_{1 \le i_1 \le i_2 \le i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k}$$

such that the inequalities in the summand are strict at positions in gap(α), i.e. $i_j < i_{j+1}$ for all $j \in \text{gap}(\alpha)$.

Example 8. Let α be the composition $(2, 1)$. Then $\text{gap}(\alpha) = (2, 3)$. Thus,

$$Q_\alpha(x_1, x_2, x_3, x_4) = Q_{\text{gap}(\alpha)}(x_1, x_2, x_3, x_4) = x_1^2 x_2 x_3 + x_1^2 x_2 x_4 + x_1^2 x_3 x_4 + x_2^2 x_3 x_4.$$

This is clearly not symmetric. In fact, the smallest symmetric function that contains it, which we can find by 'completing' each of the monomials is,

$x_1^2x_2x_3 + x_1^2x_2x_4 + x_1^2x_3x_4 + x_2^2x_3x_4 + x_1x_2^2x_3 + x_1x_2^2x_4 + x_1x_3^2x_4 + x_2x_3^2x_4 + x_1x_2x_3^2 + x_1x_2x_4^2 + x_1x_3x_4^2 + x_2x_3x_4^2$.

We define Gessel's fundamental quasisymmetric function (*quasisymmetric function*, for short) by dropping the condition that $i_k \le n$ and considering the definition over infinitely many variables. Quasisymmetric functions generalize symmetric functions. They form a ring, which we denote by QSym, over any commutative ring R.

Given a composition α, let $E(\alpha, n) = E(\text{gap}(\alpha), n)$ to be the set of sequences $\overrightarrow{i} = i_1 i_2 \cdots i_k$, which satisfy the condition of strict ascendancy, $i_j < i_{j+1}$ for $j \in \text{gap}(\alpha)$, as in the definition above. We can define the weight of such a sequence in the obvious manner. The monomial corresponding to this weight is $x_{i_1} \cdots x_{i_k}$.[5] When working with quasisymmetric functions, we instead consider the set $E(\alpha) = E(\text{gap}(\alpha))$, which is defined in a similar manner.

In the previous subsection, we noted a total ordering on the set of tableaux of a given shape. Let $\mathcal{W}(\mu)$ be the set of words of all semi-standard Young tableaux of shape μ. Using this, we denote the set of tableaux of shape λ whose entries come from the set of words $\mathcal{W}(\mu)$ by $SSYT^{\lambda}(\mathcal{W}(\mu))$.

Example 9. Let $\mu = (2, 1)$. If we only allow entires from the set $[2]$, then there are only two possible tableaux:

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1}\end{array}$$

Therefore, $\mathcal{W}(\mu) = \{212, 211\}$. Now if we let $\lambda = (2)$, then $SSYT^{\lambda}(\mathcal{W}(\mu))$ is,

$$\begin{array}{|c|c|}\hline \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1}\end{array} \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array} \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array} \\ \hline\end{array}$$

It is more convenient to represent such nested-tableaux using $a \times b$ matrices, where $\lambda \vdash a$ and $\mu \vdash b$. We do this by setting each row i of the matrix A to be the word of the μ-tableau in the i'th cell of the λ-tableau in the reading word order. For instance, in the above case, the corresponding matrices are:

$$\begin{bmatrix} 2 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}$$

We call such matrices *λ, μ-matrices.* Let $M_{a,b}(\lambda, \mu)$ be the set of all such matrices of size $a \times b$. The weight of this matrix, which we denote by $\text{wt}(A)$ is defined as the monomial $\Pi_{i,j} x_{A(i,j)}$. For instance, the above nested tableaux have weights $x_1^4 x_2^2$, $x_1^2 x_2^4$, and $x_1^3 x_2^3$, respectively. Using the weights of the nested tableaux, it is not hard to see that we can redefine the plethysm of Schur functions as,

$$s_{\lambda}[s_{\mu}] = \sum_{A \in M_{a,b}(\lambda,\mu)} \text{wt}(A)$$

This tableaux-theoretic view of plethysm is more convenient to work with in this section. Before obtaining the quasisymmetric expansion of Schur functions, we must take a detour and build the notion of standardization of matrices.

2.2.1 Standardization of Words and Matrices

Recall that the set of permutations of $[n]$ can be expressed in terms of words in which each element of this set occurs exactly once. Given any word $w = w_1 w_2 \cdots w_n$, where each of the w_i's is a positive integer and they are not necessarily inequal, we can produce its standardaziation $\text{std}(w)$ as follows: say that w contains k_i number of i's. Replace the ones in w from left to right with $1, 2, \cdots, k_1$. Then, replace the twos in w from left to right with $k_1 + 1, k_1 + 2, \cdots, k_1 + k_2$, etc. Thus $\text{std}(w) = S_n$. Let $\text{sort}(w)$ be the word obtained by sorting the letters of w into weakly increasing order. For instance, given $w = 4231223$, $\text{std}(w) = 7251346$ and $\text{sort}(w) = 1222334$.

Define the *descent set* of a word w to be $\text{des}(w) = \{i < n : w_i > w_{i+1}\}$ and the *inverse descent set*, which we denote by $\text{des}(w^{-1})$, to be the set of all $j < n$ such that $j + 1$ appears somewhere to

[5]Note that when we say weight, we could be referring to either the monomial or the powers of each of the variables in the monomial, which should be clear from the context.

the left of j in the word w. Thus, if we let $v = \mathrm{std}(w)$, where w is as in the previous paragraph, then $\mathrm{des}(v) = \{1, 3\}$ and $\mathrm{des}(v^{-1}) = \{1, 4\}$. Using the bijection mentioned above, we can note that these subsets correspond to the compositions $(1, 2, 4)$ and $(1, 3, 3)$, respectively.

Given a word v, note the following map: $v \mapsto (\mathrm{std}(v), \mathrm{sort}(v))$. This map creates a bijection between the set of all words of length n and the set of all pairs $(w, \overrightarrow{i})$, where $w \in S_n$ and $\overrightarrow{i} = i_1 i_2, \cdots i_n$ is a subscript sequence in $E(\mathrm{des}(w^{-1}))$. Using this bijection,

$$\sum_{v \in X} \mathrm{wt}(v) = \sum_{w \in S_n} Q_{\mathrm{des}(w^{-1})}.$$

Applying the RSK-correspondence to the terms of the expression on the right hand side and using the Q-expansion of Schur functions, we have the equality,

$$\sum_{w \in S_n} Q_{des(w^{-1})} = \sum_{\lambda} f^{\lambda} s_{\lambda},$$

where λ is a partition of n and f^{λ} is the number of standard tableaux of shape λ.

Now, let $M_{a,b}$ be the set of all $a \times b$ matrices with positive integer entries. Such a matrix said to be standard if each of the entries in the set $[ab]$ appear exactly once in the matrix. We denote this subset of matrices by $S_{a,b}$.

Given a matrix $A \in S_{a,b}$, we define its word, which we denote by $w(A)$ using the following rule: scan the columns of the matrix from left to right. For the first $b - 1$ columns, take each column i and write down the symbols using an ordering determined by the symbols in column $i + 1$. The ordering is that the symbol corresponding to the smallest symbol in column $i + 1$ comes first and that corresponding to the largest symbol comes last. For the rightmost column, we write the symbol in order from top to bottom. Given a word of a matrix and the size of the matrix, we can easily recover the matrix. Thus, we have a bijection between $S_{a,b}$ and the set of all permutations of $[ab]$.

Given $w(A)$, we define $\mathrm{sort}(A)$ by sorting $w(A)$ such that it is in weakly increasing order. If A is a standard matrix, then we just get $123 \cdots ab$. The *ascent set* of such a matrix, which we denote by $\mathrm{asc}(A)$, is the set of entries i such that $i + 1$ appears to the left of i in the word of A. Therefore, $\mathrm{asc}(A) = \mathrm{des}(w(A)^{-1})$.

Example 10. For instance, take the matrix

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

Then, $w(A) = 247615893$. Using this word, $\mathrm{asc}(A) = \{1, 3, 5, 6\}$.

The weight of a matrix $A \in M_{a,b}$ is defined in way that we would want it to:

$$\mathrm{wt}(A) = \Pi_{i=1}^{a} \Pi_{j=1}^{b} x_{A(i,j)}.$$

If our matrix is standard, then we just get the monomial $x_1 x_2 \cdots x_{ab}$.

Let $N_{a,b}$ to be the set of pairs (A, i) for which $A \in S_{a,b}$ and $i = i_1 i_2 \cdots i_{ab} \in E(\mathrm{asc}(A))$. We define the weight of such a pair to be the weight that we obtain from the sequence. An operation on matrices, *standardization*, and its inverse operation *unstandardization*, defines a bijective correspondence between the set of all such pairs for a given a and b and the set of all $a \times b$ matrices.

Given a matrix C, we define a corresponding standard matrix, $\mathrm{std}(C)$, as follows:

1. First, denote by $N(k)$ to be the number of symbols in C that are less than or equal to k. Thus, $N(0) = 0$ and $N(k) = ab$ for k sufficiently large enough. Compute each value $N(i), 0 \leq i \leq k$ for this matrix.

2. Using this, define the *standard labels* as follows, $L_i = \{N(i-1)+1, N(i-1)+2, \cdots, N(i)\}$. Compute each value L_i for the matrix.

3. Take the rightmost column of the matrix C. We sweep through the entries from bottom up one at a time, replacing each symbol i by the largest unused label in L_i.

4. We will work through the other columns one at a time from right to left. For each column j, scan the entries of the column from the largest row to the smallest row which is determined by the column $j+1$. Replace each symbol i as it is encountered by the largest unused label in L_i.

Definition 21. The *standardization* of a matrix C, denoted by $S(C)$, is $S(C) = (\mathrm{std}(C), \mathrm{sort}(C))$.

We call the inverse of this operation *unstandardization*, $U(A, \overrightarrow{i}) \in M_{a,b}$, which we obtain by replacing a unique copy of j in A by i_j, for $1 \leq j \leq ab$.

Example 11. Let $C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 4 \\ 1 & 2 & 3 \end{bmatrix}$ Then, $\mathrm{std}(C) = \begin{bmatrix} 1 & 4 & 6 \\ 2 & 5 & 9 \\ 3 & 7 & 8 \end{bmatrix}$. Clearly, if we were given $\mathrm{std}(C)$ and the sequence 1111234, then we can retrieve the matrix C by reversing the algorithm described above.

Lemma 6. *Standardization, $S : M_{a,b} \to N_{a,b}$, and unstandardization, $U : N_{a,b} \to M_{a,b}$, are weight-preserving and mutually-inverse bijections.*

Proof. Given $U(A, \overrightarrow{i})$, we clearly always get a matrix in $M_{a,b}$. The forward direction, however, is a little more complicated. First, note that $A = \mathrm{std}(C) \in S_{a,b}$, by construction. Thus, we only need to check that $\mathrm{sort}(C) \in E(\mathrm{asc}(A))$. Let $\mathrm{sort}(C) = i_1 i_2 \cdots i_{ab}$, which is weakly increasing by definition. The algorithm gives us the symbols in A in the exact reverse order as they occur as alphabets in $w(A)$.

Moreover, since the labels in each of the L_i's are chosen from largest to smallest, the subsequence of $w(A)$ consisting of the symbols in L_i is weakly increasing. Thus, $\mathrm{asc}(A) \subseteq \{N(1), N(2), \cdots, \}$. If $N(i-1) < k < N(i)$, the the symbol k in A was used to renumber one of the i's in the matrix C, and so $i_k = i$. Thus, $k \in \mathrm{asc}(A)$ means that $i_k < i_{k+1}$, which shows the desired result.

Finally, we want to check that $U(S(C)) = C$ and that $S(U(C)) = C$. The former is routine. To see the latter, let $x = (A, i_1 i_2 \cdots i_{ab}) \in N_{a,b}$, $C = U(x)$, and $x' \in S(C) = (A', i'_1 \cdots i'_{ab})$. For instance, using the above case, if we let

$$x = \left(\begin{bmatrix} 1 & 4 & 6 \\ 2 & 5 & 9 \\ 3 & 7 & 8 \end{bmatrix}, 111111234 \right),$$

then C would $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 4 \\ 1 & 2 & 3 \end{bmatrix}$, and $x' = (A', i'_1 \cdots i'_{ab}) = \left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 4 \\ 1 & 2 & 3 \end{bmatrix}, 111111234 \right)$.

Clearly, $i_1 \cdots i_{ab} = i'_1 \cdots i'_{ab} = \mathrm{sort}(C)$. Therefore, what remains to show is that $A' = A$. Using the construction above, let $N(i)$ and L_i be determine from C. We get C from A by replacing all labels in L_i by occurrences of i; likewise, we get A from C by replacing these i's by the standard labels in L_i. Therefore, what we would like to check is that the latter replacement puts each $k \in L_i$ back in the same position in A' that k occupied in A; but, this must always be the case using the constructing of standardization and unstandardization. In particular, if we take $w(A)$ and $w(A')$, then the labels in each L_i must occur as an increasing subsequence since $i_1 \cdots i_{ab} \in E(\mathrm{asc}(A))$ and we have already seen that it holds for $w(A)$ above. Therefore, using backwards induction on the column index, $A = A'$.

□

Corollary 2. *Using the construction of standardization and unstandardization,*

$$\sum_{C \in M_{a,b}} \mathrm{wt}(C) = \sum_{A \in S_{a,b}} Q_{\mathrm{asc}(A)} = \sum_{w \in S_{ab}} Q_{w^{-1}} = \sum_{\lambda \vdash ab} f^{\lambda} s_{\lambda}$$

Proof. Since standardization and unstandardization are weight-preserving bijections,

$$\sum_{C \in M_{a,b}} \mathrm{wt}(C) = \sum_{A \in S_{a,b}} \left(\sum_{\overrightarrow{i} \in E(\mathrm{asc}(A))} \mathrm{wt}(\overrightarrow{i}) \right)$$

But the sum inside the bracket is precisely $Q_{\mathrm{asc}(A)} = Q_{\mathrm{des}(w(A)^{-1})}$. We noted earlier that the word of such a matrix creates a bijection between $S_{a,b}$ and S_{ab} and that $\sum_{w \in S_n} Q_{\mathrm{des}(w^{-1})} = \sum_{\lambda \vdash n} f^{\lambda} s_{\lambda}$, which proves the equalities in the statement of the corollary. □

2.2.2 Quasisymmetric Expansion

We are finally ready to give the full quasisymmetric expansion of the plethysm of Schur functions, which is at the crux of this section. The goal of this subsection is to prove the following theorem:

Theorem 3 (Quasisymmetric Expansion)**.** *For all partitions λ, μ with $|\lambda| = a$ and $|\mu| = b$,*

$$s_{\lambda}[s_{\mu}] = \sum_{A \in S_{a,b}(\lambda,\mu)} Q_{\mathrm{asc}(A)}$$

Let $S_{a,b}(\lambda,\mu)$ denote the set of standard λ,μ-matrices and $N_{a,b}(\lambda,\mu)$ consist of all pairs $(A, \overrightarrow{i}) \in N_{a,b}$ where $A \in S_{a,b}(\lambda,\mu)$.

Lemma 7. *Let $C \in M_{a,b}$ and $A = \mathrm{std}(C)$, then for $1 \le i \le a$ and $1 \le j < k \le b$, $C(i,j) \le C(i,k)$ if and only if $A(i,j) < A(i,k)$ and $C(i,j) > C(i,k)$ if and only if $A(i,j) > A(i,k)$.*

Proof. First, let $x = C(i,j)$ and $y = C(i,k)$. If $x > y$, then $A(i,j) > A(i,k)$ since we will have that every label in L_x will exceed every label in L_y. Likewise, $x < y$ implies that $A(i,j) < A(i,k)$. On the other hand, if $x = y$, then $A(i,j) < A(i,k)$ since standardization creates the columns of A from right to left, with occurrences of x in C replaced by elements of L_x in decreasing order, which completes the proof. □

Lemma 8. *Given $C \in M_{a,b}$ and $A = \mathrm{std}(C)$, let C_i (and respectively A_i) be row i of C (and respectively A), viewed as a word of length b. For $1 \le i < j \le a$, $C_i \le C_j$ in lexicographic order if and only if $A_i < A_j$ and $C_j < C_i$ if and only if $A_j < A_i$.*

Proof. Let $i < j$. First assume that $C_i < C_j$. Then, there exists $k \leq b$, where $C(i,s) = C(j,s)$ for $1 \leq s < k$, and $C(i,k) < C(j,k)$ (and thus $A(i,k) < A(j,k)$). Therefore, $A(i,s) < A(j,s)$, for such $1 \leq s \leq k$ by backwards induction on the s. Take $s = 1$, then $A_i < A_j$. By a similar argument, we can show that if $C_j < C_i$, then $A_i < A_j$. Finally, let $C_i = C_j$. Since $i < j$, $A(i,b) < A(j,b)$, which we note using the standardization rule for the rightmost column. Then, $A(i,s) < A(j,s)$, for $1 \leq s \leq b$ again by backwards induction on the s. Therefore, $A(i,1) < A(j,1)$, and $A_i < A_j$. □

The proof of the theorem easily follows by noting that $s_\lambda[s_\mu] = \sum_{A \in M_{a,b}(\lambda,\mu)} \mathrm{wt}(A)$ and showing that standardization and unstandardization restrict to give us a bijection between $M_{a,b}(\lambda,\mu)$ and $N_{a,b}(\lambda,\mu)$. To see the latter claim, it suffices to show that $C \in M_{a,b}(\lambda,\mu)$ if and only if $\mathrm{std}(C) \in S_{a,b}(\lambda,\mu)$. But this follows from the order-preserving property, which we note in the above lemma.

Using this theorem, we can, for instance, see that,

$$s_{(2)}[s_{(3)}] = Q_{(33)} + Q_{(222)} + Q_{(24)} + Q_{(231)} + Q_{(42)} + Q_{(6)} + Q_{(321)} + Q_{(123)} + Q_{(141)} + Q_{(132)}.$$

Using this quasisymmetric expansion, we are able to give a combinatorial description of plethysms of the form $s_{(2)}[s_{(n)}] = h_{(2)}[h_{(n)}]$. This is one of the most popular plethysms and there are several, mostly algorithmic ways, of computing it.

Lemma 9. *For all $n \geq 1$,*

$$s_2[s_n] = \sum_{c=0, c \text{ is even}}^{n} s_{2n-c,c}.$$

We do not present the construction here since the proof of the lemma would be highly technical and not very insightful. However, the idea is that we use matrices of the form $M_{2,n}((2),(n))$, which are matrices where each row is weakly increasing and the first row occurs weakly before the second row in lexicographic order. Given a matrix $A \in S_{2,n}((2),(n))$, we define the word of the matrix $w(A) = w_1 \cdots w_{2n}$ by letting each w_i equal 0 if i occurs in the first row of A and 1 if it occurs in the second row of A. This word then also encodes a lattice path from $v_0 = (0,0)$ to $v_{2n} = (n,n)$. We can now apply concepts from lattice paths, such as the idea of a lattice path being marked, to work out the expansion above.

2.3 Monomial Expansion of Plethysm

In this section, we will give a combinatorial description for the coefficients of the plethysm of Schur functions in terms of the monomial symmetric functions,

$$s_\lambda[s_\mu] = \sum_{\nu \vdash ab} Y^{\nu}_{\lambda[\mu]} m_\nu,$$

where m_ν is the monomial symmetric function corresponding to the partition ν, $\lambda \vdash a$, and $\mu \vdash b$. Using this, we are able to determine the first term (the maximal element in lexicographic order) in the Schur function expansion and show that this term always has coefficient 1. These results are from [11].

2.3.1 Monomial Symmetric Function Expansion

As discussed earlier, we can picture the plethysm of Schur functions as a nested Young tableau, where we have put a total ordering on all the semi-standard Young tableaux of a given shape using the lexicographic order on the corresponding words. Recall that we denote the set of semi-standard Young tableaux of shape λ and entries from the set S by $SSYT^{\lambda}(S)$. In this section, we denote the set of semi-standard Young tableaux of this shape with weight μ by $SSYT(\lambda;\mu)$. Thus, the set of semi-standard Young tableaux corresponding to the plethysm $s_\lambda[s_\mu]$ is $SSYT^{\lambda[\mu]}$ and the set of semi-standard Young tableaux of shape $\lambda[\mu]$ with entries from $\{1, 2, \cdots, s\} = [s]$ is $SSYT^{\lambda[\mu]}([s])$. Therefore, a Schur polynomial with m variables is,

$$s_\lambda(x_1, \cdots, x_m) = \sum_{T \in SSYT^{\lambda}([m])} X^T.$$

Definition 22. Given $T \in SSYT^{\lambda[\mu]}$ with weight $\nu = (\nu_1, \nu_2, \cdots, \nu_k)$, such that $\lambda \vdash a, \mu \vdash b$ and $\nu \vdash ab$,

$$Y^{\nu}_{\lambda[\mu]} = |SSYT(\lambda[\mu];\nu)|$$

Example 12. Take, for instance, $\lambda = (2), \mu = (1,1), \nu = (2,2)$. There are two different possible nested tableaux,

$$\boxed{\begin{array}{|c|}\hline 1\\\hline 2\\\hline\end{array}\,\begin{array}{|c|}\hline 1\\\hline 2\\\hline\end{array}} \qquad \boxed{\begin{array}{|c|}\hline 1\\\hline 1\\\hline\end{array}\,\begin{array}{|c|}\hline 2\\\hline 2\\\hline\end{array}}$$

Therefore, $Y^{\nu}_{\lambda[\mu]} = 2$.

Theorem 4 (Monomial Expansion)**.** *Given partitions $\lambda \vdash a$, $\mu \vdash b$, and $\nu \vdash ab$, $Y^{\nu}_{\lambda[\mu]}$ in the definition above gives as the coefficients in the monomial symmetric function expansion,*

$$s_\lambda[s_\mu] = \sum_{\nu \vdash ab} Y^{\nu}_{\lambda[\mu]} m_\nu.$$

Proof. Given a positive integer s, let $r = |SSYT^{\mu}([s])|$. We order these tableaux lexicographically and denote the i'th largest tableau in this set by T_i. Therefore, $SSYT^{\mu}([s]) = \{T_1, \cdots, T_r\}$. Let $y_i = x^{T_i}$.

We have the following maps:

$$\iota : SSYT^{\lambda}([r]) \to SSYT^{\lambda}(SSYT^{\mu}([s])) \to SSYT^{\lambda[\mu]}([s]))$$

Let us consider an example where $\mu = (2)$ and $s = 2$. Then,

$$T_1 = \begin{array}{|c|}\hline 2\\\hline 2\\\hline\end{array} \qquad y_1 = x_2^2$$

$$T_2 = \begin{array}{|c|}\hline 1\\\hline 2\\\hline\end{array} \qquad y_2 = x_1x_2$$

$$T_3 = \begin{array}{|c|}\hline 1\\\hline 1\\\hline\end{array} \qquad y_3 = x_1^2$$

Let $\lambda = (1,1)$, then the map above claims that there is a bijection between the set $\left\{\begin{array}{|c|}\hline 1\\\hline 2\\\hline\end{array}, \begin{array}{|c|}\hline 1\\\hline 3\\\hline\end{array}, \begin{array}{|c|}\hline 2\\\hline 3\\\hline\end{array}\right\}$ and the following set of nested tableaux,

$$\boxed{\begin{array}{c}\begin{array}{|c|c|}\hline 1&1\\\hline\end{array}\\ \begin{array}{|c|c|}\hline 1&2\\\hline\end{array}\end{array}} \qquad \boxed{\begin{array}{c}\begin{array}{|c|c|}\hline 1&1\\\hline\end{array}\\ \begin{array}{|c|c|}\hline 2&2\\\hline\end{array}\end{array}} \qquad \boxed{\begin{array}{c}\begin{array}{|c|c|}\hline 1&2\\\hline\end{array}\\ \begin{array}{|c|c|}\hline 2&2\\\hline\end{array}\end{array}}$$

From this example, it is not hard to see that the maps given above are, in fact, bijections. Now, using the tableaux-theoretic definition of Schur functions,

$$s_\lambda(x_1, \cdots, x_s) = \sum_{T \in SSYT^\lambda([s])} x^T = y_1 + y_2 + \cdots + y_r.$$

For instance, in the above case, $s_{1,1}[s_2] = x_1^3 x_2 + x_1^2 x_2^2 + x_1 x_2^3$, which we can see by computing both $\sum_{T \in SSYT^\lambda([s])} x^T$ and $y_1 + y_2 + \cdots + y_r$, independently. Therefore,

$$\begin{aligned}
s_\lambda[s_\mu](x_1, \cdots, x_s) &= s_\lambda(x^{T_1}, \cdots, x^{T_r}) \\
&= s_\lambda(y_1, \cdots, y_r) \\
&= \sum_{U \in SSYT^\lambda([r])} y^U \\
&= \sum_{\iota(U) \in SSYT^{\lambda[\mu]}([s])} x^{\iota(U)}
\end{aligned}$$

We get the desired equality above by considering what happens as $s \to \infty$,

$$s_\lambda[s_\mu] = \sum_{T \in SSYT^{\lambda[\mu]}} x^T = \sum_{\nu \vdash ab} Y^\nu_{\lambda[\mu]} m_\nu$$

□

2.3.2 The First Term

As promised, we are able to find the maximal element in the Schur function expansion of our plethysm using the monomial expansion description above.

Theorem 5 (First Term)**.** *Let ℓ_λ be the length of the partition λ and ℓ_μ be the length of the partition μ. Then, the first term of the plethysm $s_\lambda[s_\mu]$ is s_{ν_0}, where*

$$\nu_0 = (a\mu_1, a\mu_2, \cdots, a\mu_{\ell_\mu} - 1 + \lambda_1, \lambda_2, \cdots, \lambda_{\ell_\lambda}).$$

Moreover, this maximal term always has coefficient 1.

Proof.

$$\begin{aligned}
a^\nu_{\lambda[\mu]} &= \langle s_\lambda[s_\mu], \sum_\kappa K^{-1}_{\kappa,\nu} h_\kappa \rangle \\
&= \sum_\kappa K^{-1}_{\kappa,\nu} \langle s_\lambda[s_\mu], h_\kappa \rangle \\
&= \sum_\kappa K^{-1}_{\kappa,\nu} Y^\kappa_{\lambda[\mu]}, \text{ by the monomial expansion theorem} \\
&= Y^\nu_{\lambda[\mu]} + \sum_{\kappa > \nu} K^{-1}_{\kappa,\nu} Y^\kappa_{\lambda[\mu]}, \text{ using the definition of Kostka coefficients.}
\end{aligned}$$

To see that the partition ν_0 does, in fact, give us the maximal such partition such that

$Y^{\nu}_{\lambda[\mu]} \neq 0$, consider the following series of tableaux T_i:

$$\begin{aligned}
\mathrm{wt}(T_1) &= (\mu_1, \mu_2, \cdots, \mu_{\ell_\mu}) \\
\mathrm{wt}(T_2) &= (\mu_1, \mu_2, \cdots, \mu_{\ell_\mu} - 1, 1) \\
\mathrm{wt}(T_3) &= (\mu_1, \mu_2, \cdots, \mu_{\ell_\mu} - 1, 0, 1) \\
&\vdots \\
\mathrm{wt}(T_\ell) &= (\mu_1, \mu_2, \cdots, \mu_{\ell_\mu} - 1, 0, \cdots, 0, 1)
\end{aligned}$$

For each of the T_i's all the 1's appear in the first row, all the 2's appear in the second row, etc until we get to row ℓ_μ. The last row has two entries, the left most of which is ℓ_μ and the right most of which is $\ell_\mu + \ell_\lambda - 1$. The $T_1, T_2, \cdots T_\ell$ are ordered in lexicographic order by their weight from largest to smallest.

This construction is particularly useful since the sequence $T^1 \leq T^2 \leq \cdots \leq T^a$ that has maximal weight under the condition that $Y^{\nu}_{\lambda[\mu]} \neq 0$ is,

$$\begin{aligned}
T_1 &= T^1 = \cdots T^{\lambda_1} \\
T_2 &= T^{\lambda+1} = \cdots T^{\lambda_1+\lambda_2} \\
&\vdots \\
T_\ell &= T^{a-\lambda_\ell+1} = \cdots = T^a
\end{aligned}$$

where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_\ell)$ in the plethysm $s_\lambda[s_\mu]$. Therefore, the maximal weight is,

$$\mathrm{wt}(T^1) + \cdots + \mathrm{wt}(T^a) = \lambda_1 \,\mathrm{wt}(T_1) + \cdots + \lambda_\ell \,\mathrm{wt}(T_\ell) = (a\mu_1, a\mu_2, a(\mu_{\ell_\mu} - 1) + \lambda_1, \lambda_2, \cdots, \lambda_{\ell_\lambda}),$$

which concludes the proof. $\square$

Example 13. Take the plethysm $s_{1,1}[s_3] = s_3 + s_{5,1}$. The term $s_{5,1}$ is the maximal element. We can also obtain this using the theorem above since $\{5, 1\} = \{2(3 - 1) + 1, 1\}$.

This theorem also gives us another combinatorial expression for $a^{\nu}_{\lambda[\mu]}$ using $Y^{\kappa}_{\lambda[\mu]}$, which we have seen has a combinatorial coefficient and the inverse Kostka coefficients, which are also shown to have a combinatorial description in [5]. However, this is not always so easy to compute.

Corollary 3. $a^{\nu}_{\lambda[\mu]} = \sum_{\kappa \vdash ab} K^{-1}_{\kappa,\mu} Y^{\kappa}_{\lambda[\mu]}$.

2.4 Plethysm at Hook Shapes

We consider hook shaped partitions, i.e. partitions of the form $\nu = (1^a, b)$, and show that such partitions occur in the expansion $s_\lambda[s_\mu] = a^{\nu}_{\lambda,\mu} s_\nu$ only when both λ and μ are hook shaped. We will also show that a single hook shape occurs in the expansion of the plethysm $s_{(1^c,d)}[s_{(1^a,b)}]$. These results are from [13].

Given a shape λ calculate its corresponding Schur function over the variables $X = \{x_1, x_2, \cdots\}$ and over the variables $Y = \{y_1, y_2, \cdots\}$. Using [17],

$$s_\lambda(x_1, x_2, \cdots, y_1, y_2, \cdots) = s_\lambda(X + Y) = \sum_{\mu \subseteq \lambda} s_\mu(X) s_{\lambda/\mu}(Y).$$

Similarly,

$$s_\lambda(X - Y) = \sum_{\mu\subseteq\lambda} s_\mu(X)(-1)^{|\lambda/\mu|} s_{(\lambda/\mu)^T}(Y).$$

From this, we can note that,

$$s_{\lambda/\mu}(-X) = (-1)^{|\lambda/\mu|} s_{(\lambda/\mu)^T}(X).$$

Example 14. We have noted that $s_{1,1}(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_3$. Now let $X = \{x_1, x_2\}$ and $Y = \{x_3\}$. Then, using the above formula,

$$\begin{aligned} s_{(1,1)}(x_1, x_2, x_3) &= s_{1,1}(X)s_\emptyset(Y) + s_{(1)}(X)s_{(1)}(Y) + s_\emptyset(X)s_{1,1}(Y) \\ &= x_1x_2 + (x_1 + x_2)x_3 \end{aligned}$$

as desired. To work out an example for the second equality, let $X = \{x_1, x_2, x_3\}$ and $Y = \{x_3\}$. Then, we expect that $s_{(1,1)}(X - Y) = x_1x_2$. To note this using the equality above,

$$\begin{aligned} s_{(1,1)}(X - Y) &= s_{1,1}(X)s_\emptyset(Y) - s_{(1)}(X)s_{(1)}(Y) + s_\emptyset(X)s_2(Y) \\ &= x_1x_2 + x_1x_3 + x_2x_3 - (x_1 + x_2 + x_3)(x_3) - x_3^2 = x_1x_2 \end{aligned}$$

again as desired.

We will first show that no Schur function corresponding to a hook shape can show up in an expansion of the plethysm of Schur functions unless both Schur functions are hook shaped. Let

$$s_\lambda[s_\mu]\ |_{\text{hooks}} = \sum_{\nu \text{ a hook}} a_\nu s_\nu$$

.

Lemma 10. $s_\lambda[s_\mu]\ |_{\text{hooks}} = 0$ *unless both* λ *and* μ *have hook shapes.*

Proof. Let $s_\lambda[s_\mu] = \sum_\nu a_\nu s_\nu$. We consider this plethysm with variable $(1 - x)$. i.e.,

$$s_\lambda[s_\mu](1 - x) = \sum_\nu a_\nu s_\nu(1 - x).$$

If we let $1 = X$ and $x = Y$, then, by the above theorem,

$$s_\nu(1 - x) = \sum_{\rho\subseteq\nu} s_\rho(1)(-1)^{|\nu/\rho|} s_{(\nu/\rho)^T}(x).$$

Note that we have only one parameter, x, in this evaluation. Using the tableaux-theoretic definition of Schur functions, we can see that the only time such a Schur function can be non-zero is when it does not have a column of height greater than one. If we are considering a non-skew tableau, then we get a single row and if we have a skew-tableau, then we get a skew-row. This can only happen if $\nu = (1^a, b)$ and $\rho = (b)$ or $(b-1)$. Thus, $s_\lambda[s_\mu](1-x) = \sum_{\nu \text{ a hook}} a_\nu s_\nu(1-x)$. Using the above equations,

$$\begin{aligned} s_{(1^a,b)}(1 - x) &= \sum_{\rho\subseteq(1^a,b)} s_\rho(1)(-1)^{|(1^a,b)/\rho|} s_{((1^a,b)/\rho)^T}(x) \\ &= s_{b-1}(1)(-1)^{a+1}(s_1s_a)(x) + s_b(1)(-1)^a s_1(x) \\ &= (-1)^{a+1}x^{a+1} + (-1)^a x^a \end{aligned}$$

Therefore, we want to look for sums of this form in the expansion $s_\lambda[s_\mu](1-x)$. First, since $s_\lambda[s_\mu](1-x) = s_\lambda(s_\mu(1-x))$, this former is zero unless μ is hook shaped.

Now, if we let $\mu = (1^a, b)$, then $s_\lambda(s_{(1^a,b)}(1-x)) = s_\lambda((-1)^{a+1}x^{a+1} + (-1)^a x^a)$. But the expression on the right hand side has one positive and one negative term. Therefore, using the same argument as the one for $s_\nu(1-x)$, the plethysm ${}_\lambda(s_{(1^a,b)}(1-x))$ is zero unless λ is also hook shaped. □

In light of the above lemma, we are now able to give an explicit formula for computing the hook shaped Schur function that shows up in the plethysm of two hook shaped Schur functions. This description is nice since it gives us a direct formula that allows us to compute a term in the Schur function expansion of the plethysm.

Theorem 6. *Let* $\lambda = (1^c, d)$ *and* $\mu = (1^a, b)$*, then*

$$s_{(1^c,d)}[s_{(1^a,b)}]\,|_{\text{hooks}} = s_{1^{a(c+d)+c},b(c+d)-c}, \text{ if } a \text{ is even, and}$$
$$s_{(1^c,d)}[s_{(1^a,b)}]\,|_{\text{hooks}} = s_{1^{a(c+d)+d-1},b(c+d)-d+1}, \text{ if } a \text{ is odd}$$

Proof. The proof of this is almost purely computational and not interesting, so we will only outline the steps. Take the plethysm $s_{(1^c,d)}[s_{(1^a,b)}](1-x) = s_{(1^c,d)}((-1)^{a+1}x^{a+1} + (-1)^a x^a)$. First, we assume that a is odd. Then, the expansion of $s_{(1^c,d)}(x^{a+1} - x^a)$ gives is $x^{a(c+d)+d-1}(-1)^{c+1} + x^{a(c+d)+d}(-1)^c$. (This involves multiple steps, and the reader may want to verify this for themselves.) This is almost in the form that we want it from using the above lemma, except that we have to verify that $a(c+d)+d$ and c have the same parity. But $a(c+d)+d = ac + d(a+1)$, which shows that they in fact do. Therefore,

$$s_{(1^c,d)}[s_{(1^a,b)}](1-x) = x^{a(c+d)+d-1}(-1)^{a(c+d)+d-1} + x^{a(c+d)+d}(-1)^{a(c+d)+d},$$

which, using the above lemma,

$$s_{(1^c,d)}[s_{(1^a,b)}](1-x) = s_{(1^{a(c+d)+d-1},\ell)}(1-x)$$

for some ℓ. But since $\lambda \vdash a, \mu \vdash b$, and $\nu \vdash ab$, $(c+d)(a+b) = a(c+d) + d - 1 + \ell$, and so $\ell = b(c+d) - d + 1$. Therefore, we get the desired result for when a is odd. The case of a even should also follow similarly. □

Example 15. Using SAGE or another computing software, we can see that,

$$s_{1,1}[s_{2,1}] = s_{2,2,1,1} + s_{3,2,1} + s_{3,3} + s_{4,1,1}.$$

Indeed, we can get the term $s_{4,1,1}$ using this formula provided in the theorem since $a, c, d = 1$ and $b = 2$.

The paper that discusses these results, [13], also considers an extension to the case of near-hook shapes, which are shapes of the form $(1^a, b, c)$ and $(1^a, 2^b, c)$. In this case, we have to consider $s_\nu(1 + x - y)$. While they were able to show that this Schur function has value 0 unless ν is contained in a hook plus a row, proving the generalized form of the lemma and theorem above involves extensive computations and an explicit formula for near-hook shapes has not been found.

3 Conclusion and Further Work

Finding a full combinatorial description for the expansion of the plethysm of Schur functions remains one of the most fundamental open problems in the theory of symmetric functions. As discussed in Section 1, this would give us a direct way of calculating the characters of the compositions of irreducible polynomial representations of $GL(n, \mathbb{C})$.

Schur functions appear in various areas of mathematics and mathematical physics. For instance, they are useful in the representation theory of the symmetric group and the general linear group over $\mathbb{C}$ or over finite fields. The cohomology ring of the Grassmann variety can be described in terms of Schur functions, providing a deep connection with Schubert calculus. They are also ubiquitous in mathematical physics, as shown in [10]. Therefore, progress made on this problem will have wide-spread impact.

In this exposition, we note different expansions of Schur functions in terms of quasisymmetric functions, monomial symmetric functions and geometrically as nested Young tableaux. These are used to give combinatorial descriptions of plethysms of special classes of Schur functions, such as those involving hook shapes and plethysms of the form $s_{(2)}[s_{(n)}]$.

While a general combinatorial description has been elusive, it is possible that some of the results given here can generalize further. For instance, the description in Corollary 3 as well as coefficients in the monomial expansions offer a perspective that have not been explored as widely. Both of these coefficients have combinatorial descriptions, which may inspire further work. Likewise, the geometric restriction imposed by considering Schur functions of hook shape significantly simplifies their plethysm. In addition to considering more combinatorial approaches for plethysm of Schur functions of near-hook shapes, it might also be worthwhile to attempt to find a full expansion for hook-shaped ones.

Acknowledgements

This thesis is a result of a two year reading experience with Professor Richard Stanley, where my interest in enumerative and algebraic combinatorics came to fruition and a two semester reading course with Professor Michael Hopkins, where I feel in love with Young tableaux. I express my deepest gratitude to both Professor Hopkins and Professor Stanley for supervising my reading, nurturing my interests and supporting my growth as a mathematician. I could not have asked for better advisors and mentors.

I would also like to thank Dr. Paul Bourgade for teaching a course in algebraic combinatorics where my interest took root, Atanas 'Nasko' Atanasov for discussing various mathematical topics with me, Sam Raskin for leading a representation theory tutorial during the summer of 2010, which further motivated my interest in algebraic combinatorics, and other members of the Harvard University Department of Mathematics who have always been open and willing to answer my questions and provide suggestions.

Finally, I would like to dedicate this thesis to my biological and chosen family–the Abebes and Neale Parises–for all the love and support through the years, and friends who have made my time in college phenomenal.

A Schur Polynomials

We have made extensive use of certain basic properties of Schur polynomials and Schur functiosn throughout this thesis. We discuss a few of these in this section. We begin with a tableaux-theoretic definition of Schur polynomials.

Definition 23. For $\lambda \vdash n$, the **Schur polynomial** $s_\lambda(x_1, x_2, \cdots, x_m)$ is the symmetric polynomial given by,

$$s_\lambda(x_1, x_2, \cdots, x_m) = \sum_{T \in SSYT^\lambda([m])} x_1^{\mu_1} x_2^{\mu_2} \cdots x_m^{\mu_m}$$

where the sum runs over all semi-standard Young tableaux with shape λ and entries from $[m]$. The μ_i are the number of entries of i in T, so that $\sum \mu_i = n$.

Recall that the monomial $x_1^{\mu_1} x_2^{\mu_2} \cdots x_m^{\mu_m}$ is said to be of weight $\mu = (\mu_1, \cdots, \mu_m)$.

Example 16. If $\lambda = (2, 1)$, the Schur polynomial in 3 variables is,

$$s_\lambda(x_1, x_2, x_3) = x_1^2 x_2 + x_1^2 x_3 + x_1 x_2^2 + x_1 x_3^2 + x_2^2 x_3 + x_2 x_3^2 + 2x_1 x_2 x_3$$

Lemma 11. *Schur polynomials are symmetric.*

Proof. It is sufficient to prove that s_λ is invariant under interchanging x_i and x_{i+1}, for $1 \leq i < m$. Let $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \cdots)$ and $\tilde{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_{i-1}, \alpha_{i+1}, \alpha_i, \cdots)$. Consider the set of $SSYT^\lambda$ with weight α and those with weight $\tilde{\alpha}$, which we denote by $\mathcal{T}_{\lambda,\alpha}$ and $\mathcal{T}_{\lambda,\tilde{\alpha}}$, respectively. We want a bijection $\phi : \mathcal{T}_{\lambda,\alpha} \to \mathcal{T}_{\lambda,\tilde{\alpha}}$.

Take $T \in \mathcal{T}_{\lambda,\alpha}$, and consider the parts of T equal to i and $i+1$. We ignore the columns that contain no or two such parts. For the remaining columns, either i or $i+1$ occur once in each column. For a row with number of i entries equal to r and number of $i+1$ entries equal to s, convert this into a row with number of i and $i+1$ entries equal to s and r respectively. The resulting array clearly belongs to $\mathcal{T}_{\lambda,\tilde{\alpha}}$ since we have not changed the shape of the tableau and the type is precisely what we get when performing these switches. This establishes the desired bijection. □

As λ varies over all partitions of n into at most m parts, these Schur polynomials form a basis for the symmetric polynomials of degree n in m variables.

Lemma 12. *The Schur polynomials form a $\mathbb{Z}$-basis for the ring of all symmetric polynomials.*

Proof. Take a symmetric polynomial $p(x_1, \cdots, x_m)$ and order the monomials in the variables x_i for $1 \leq i \leq m$ in lexicographic order, where each of the individual terms are ordered $x_1 > x_2 > \cdots > x_m$. That is, the dominant term of p is the one with the highest occurring power of x_1, and among those the one with the highest power of x_2, and so on. The first term is called the leading monomial. Take $p_1 := p - a \cdot s_\lambda(x)$, where a is the coefficient of the leading monomial and λ is given by the weight of the leading monomial. Note that p_1 is symmetric since the sum of symmetric polynomials is symmetric, and it is smaller with respect to this ordering. We can iterate this process, which terminates in a finite number of steps. Therefore, the Schur polynomials span the ring of all symmetric polynomials.

To see that they are linearly independent, assume otherwise. For a non-trivial relation of linear dependence, $\sum a_\lambda s_\lambda = 0$, let λ be maximal such that $a_\lambda \neq 0$. Then, the left hand side has a leading monomial x^λ with coefficient a_λ, so cannot be zero, giving us a contradiction. □

Example 17. The polynomial $p = 2x_1^2x_2 + 2x_1x_2^2 + 5x_1^3 + 5x_2^3$ is symmetric and has leading monomial $5x_1^3$. We demonstrate the process given above as follows:

$$\begin{aligned}
p &= 2x_1^2x_2 + 2x_1x_2^2 + 5x_1^3 + 5x_2^3 \\
p_1 &= (2x_1^2x_2 + 2x_1x_2^2 + 5x_1^3 + 5x_2^3) - 5s_{(3)}(x_1, x_2) \\
&= -3x_1^2 - 3x_1x_2^2 \\
p_2 &= -3(x_1^2x_2 + x_1x_2^2) + 3s_{(2,1)}(x_1x_2) = 0
\end{aligned}$$

There are other classes of symmetric polynomials, such as the elementary symmetric polynomials and the monomial symmetric polynomials, that also form such a basis for the ring of symmetric polynomials.

These results and other properties involving different classes of symmetric polynomials are not dependent on the number of variables. We have also seen that having too few variables can lead to loss of data; if the number of parts of the partition of n is greater than the number of variables, then the Schur polynomial vanishes.

We therefore work in the ring of symmetric functions, the "direct limit" of the ring of symmetric polynomials in m indeterminates as m goes to infinity. The ring of symmetric functions serves as a structure where the relations between the different symmetric polynomials and their properties can be expressed in a way that is independent of the number of variables. Schur functions are the elements of this ring that correspond to Schur polynomials. A complete treatment of this topic can be found in [17] and in [21].

B Littlewood-Richardson Rule

Throughout this thesis, we stated that the product of Schur functions can be computed explicitly using a combinatorial rule. This gives us the characters of the tensor product representations. We present this rule in this section.

Consider the product of two Schur functions s_λ and s_μ. We would like to write this product as a linear combination of Schur functions. Let $c^\nu_{\lambda,\mu}$ be the coefficient of s_ν in this expansion, where $|\lambda| + |\mu| = |\nu|$. That is,

$$s_\lambda s_\mu = \sum_\nu c^\nu_{\lambda,\mu} s_\nu,$$

or equivalently,

$$s_{\nu/\lambda} = \sum_\mu c^\nu_{\lambda,\mu} s_\mu.$$

We call $c^\nu_{\lambda,\mu}$ the Littlewood-Richardson coefficient.

The Littlewood-Richardson rule, which gives us these coefficients, shows that the value of these coefficients depends only on these shapes and not the entries of the tableaux. It is therefore independent on the number of variables.

Definition 24. A lattice word is a sequence of integers such that when reading it from left to right there are as many instances of i as $i + 1$. A **Yamanouchi word** is a sequence whose reversal is a lattice word.

Recall that a word of a tableaux is obtained by reading it from left to right then bottom to top.

Definition 25. A skew tableau is called a **Littlewood-Richardson skew tableau** if its word is a Yamanouchi word.

Proposition 5 (Littlewood-Richardson Rule)**.** *The Littlewood-Richardson coefficient $c^\nu_{\lambda,\mu}$ counts the number of Yamanouchi skew tableau of shape ν/λ and weight μ.*

There are two main restrictions here. The first is that ν must contain λ, (or alternatively μ, since Schur function products are commutative). The second is that the word of such a tableau must be a Yamanouchi word.

Example 18. Let $\lambda, \mu = (2, 1)$. The possible shapes of ν are,

$$(4, 2), (4, 1, 1), (3, 3), (3, 2, 1), (3, 1, 1, 1, (2, 2, 1), (2, 2, 1, 1).$$

First, ν cannot be the partition (6) since this partition does not contain $(2, 1)$. If we let $\nu = (5, 1)$, then the only possible tableaux of shape $(5, 1)/(2, 1) = (3)$ of weight $(2, 1)$ is,

$$\begin{array}{|c|c|c|}\hline 1 & 1 & 2 \\ \hline\end{array}$$

But, the corresponding word for this tableau is 112, which is not a Yamanouchi word.

Let $\nu = (4, 2)$. Then, the following is the corresponding Littlewood Richardson skew tableaux:

$$\begin{array}{|c|c|c|c|}\hline \bullet & \bullet & 1 & 1 \\ \hline \bullet & 2 \\ \cline{1-2}\end{array}$$

where the boxes containing $\bullet$ indicate the shape of λ. For $\nu = (3, 2, 1)$, we have two possibilities,

$$\begin{array}{|c|c|c|}\hline \bullet & \bullet & 1 \\ \hline \bullet & 2 \\ \cline{1-2} 1 \\ \cline{1-1}\end{array} \quad \text{and} \quad \begin{array}{|c|c|c|}\hline \bullet & \bullet & 1 \\ \hline \bullet & 1 \\ \cline{1-2} 2 \\ \cline{1-1}\end{array}$$

We proceed in this manner to note that,

$$s_{(2,1)}s_{(2,1)} = s_{(4,2)} + s_{(4,1,1)} + s_{(3,3)} + 2s_{(3,2,1)} + s_{(3,1,1,1)} + s_{(2,2,2)} + s_{(2,2,1,1)}.$$

We can therefore use this rule to decompose the product of Schur functions into linear combinations of Schur functions, which corresponds to decomposing the tensor product of representations into (the direct sum of) irreducible representations.

This rule was given by D.E. Littlewood and A.R. Richardson in [15], although a complete proof was not found until four decades later. Several shorter proofs have been found in recent decades. We do not present any here, but direct the reader to [22] for one proof.

Using the Littlewood-Richardson rule, we find a simple formula for two special cases. Let (p) denote the Young diagram with a single row and (1^p) denote the Young diagram with a single column, both of length p. Then,

$$s_\lambda s_{(p)} = \sum_\mu s_\mu,$$

where the sum is taken over all μ that are obtained from λ by adding p boxes with no two in the same column; and

$$s_\lambda s_{(1^p)} = \sum_\mu s_\mu,$$

where the sum is taken over all μ that are obtained from λ by adding p boxes with no two in the same row.

This product is analogous to Pieri's formula, which tells us how to intersect an arbitrary Schubert class with a special type of Schubert classes corresponding to partitions of the form (p) and (1^p). In fact, it is known that we multiply Schubert classes in the Grassmann using the same rule as above. This is used to solve various problems in enumerative geometry. A complete exposition of this can be found in [7].

References

[1] C. Carré and B. Leclerc. Splitting the square of a Schur function into its symmetric and antisymmetric parts. *Journal of Algebraic Combinatorics*, 4(3): pp 201-231, 1995.

[2] Y. M. Chen, A. M. Garsia, and J. B. Remmel. Algorithms for plethysms. *Combinatorics and Algebra (Boulder, Colo., 1983)*, vol. 34 of *Contemporary Mathematics*: pp 109-153. American Mathematical Society, Providence, RI, 1984.

[3] K. Conrad. The minimal polynomial and some applications. www.math.uconn.edu/~kconrad/blurbs/minpolyandappns.pdf.

[4] C. Curtis and I. Reiner. *Representation Theory of Finite Groups and Associative Algebras*. New York/London/Sydney: Interscience Publishers, 1962.

[5] O. Egecioglu and J. B. Remmel. A combinatorial interpretation of the inverse Kostka matrix. *Linear and Multilinear Algebra*, 26: pp 59-84, 1990.

[6] E. Egge, N. Loeh, and G. Warrington. From quasisymmetric expansions to Schur expansions via a modified inverse Kostka matrix. *European Journal of Combinatorics*, 31(8): pp 2014-2027, 2010.

[7] D. Eisenbud, J. Harris. 3264 & all that: intersection theory in algebraic geometry. isites.harvard.edu/fs/docs/icb.topic720403.files/book.pdf.

[8] W. Fulton. *Young Tableaux*. Cambridge UK: Cambridge University Press, Inc, 1997.

[9] W. Fulton and J. Harris. *Representation Theory: A First Course*. New York: Springer, 1991 (GTM 129).

[10] A. M. Garsia and J. Haglund. A proof of the q, t-Catalan positivity conjecture. *Discrete Mathematics*, 256(3): pp 677-717, 2002.

[11] K. Iijima. The first term of plethysms. arXiv: 1108.4815.

[12] D. Krammer. “The Symmetric Groups”. www.homepages.warwick.ac.uk/~masbal/MA5P2.0809/index.html.

[13] T. M. Langley and J. B. Remmel. The plethysm of $s_\lambda[s_\mu]$ at hook or near-hook shapes. *Electronic Journal of Combinatorics*, 11(1): Research Papers 11, 2004.

[14] D. E. Littlewood. Invariant theory, tensors and group characters. *Philosophical Transactions of the Royal Society of London. Series A*, 239: pp 305-365, 1944.

[15] D. E. Littlewood and A.R. Richardson. Group characters and algebra. *Philosophical Transactions of the Royal Society of London. Series A*. 233: pp 99141, 1934.

[16] N. A. Loehr and G. S. Warrington. Quasisymmetric expansion of Schur-function plethysms. *Proceedings of the American Mathematical Society*, 140: pp 1159-1171, 2012.

[17] I. G. MacDonald: *Symmetric Functions and Hall Polynomials*. New York: Oxford University Press Inc., 1995 (Second Edition).

[18] B. Sagan: *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*. New York: Springer, 2001 (GTM 203).

[19] I. Schur. Über eine Klasse von Matrizen, die sich einer gegebenen Matrix zuordnen lassen, Inaugural-Dissertation, Friedrich-Wilhelms-Universitat zu Berlin 1901; *Ges. Abhandlugen*, vol.1, pp 1-72.

[20] J-P. Serre. *Linear Representations of Finite Groups*. New York: Springer, 1977 (GTM 42).

[21] R. Stanley. *Enumerative Combinatorics, Vol. 2*. Cambridge UK: Cambridge University Press, Inc, 1999.

[22] J.R. Stembridge. A concise proof of the Littlewood-Richarson rule. *Electronic Journal of Combinatorics*, 9 (1): Note 5, 4 pp. (electronic), 2002.