Combinatorics of Symmetric Functions

Kevin Anderson
Supervised by Dr. Hadi Salmasian

August 28, 2019

Contents

Introduction

Symmetric polynomials are an important object of study in many fields of modern mathematics. The Schur polynomials (and their generalizations) are a particularly pervasive family of symmetric polynomials which arise naturally in such diverse settings as representation theory, algebraic geometry, and mathematical physics. In addition to their many applications, the Schur polynomials are notable for their fascinating combinatorial properties. The purpose of this report is to provide a thorough but streamlined survey of the basic theory of Schur polynomials, analyzed through the lens of algebraic combinatorics.

The report is divided into six sections and is organized as a synthesis of three main sources. In Section 1 we introduce terminology and notation related to integer partitions, which are the basic combinatorial object used to parametrize many important families of symmetric polynomials. Sections 2 and 3 are dedicated to developing the classical theory of symmetric polynomials and introducing several important families of symmetric polynomials which appear throughout the text, most notably the Schur polynomials. Our presentation of these topics is primarily adapted from that of Procesi in [7, Ch. 2]. Section 4 introduces some of the combinatorial theory of Young tableaux and is based on Fulton's treatment of the subject in [2]. Sections 5 and 6 are dedicated to studying some of the combinatorial properties of symmetric polynomials and are based primarily on the presentation offered by Macdonald in [6]. In particular, Section 5 focuses on the combinatorial theory of Schur polynomials, while Section 6 focuses on analogous properties of the Macdonald polynomials, which are natural generalizations of the Schur polynomials having many interesting applications, though they are seldom included in introductory texts on symmetric polynomials.

We intend for our treatment to be accessible to an undergraduate audience, with prerequisites kept to a minimum. Our treatment should be especially appealing to readers with little or no background in representation theory, and any reader with basic knowledge of algebra (*i.e.* group theory, ring theory, and linear algebra) at the undergraduate level should be able to take full advantage of these notes. We intend, as much as possible, for these notes to be self-contained, although a small number of results are delegated to external sources for the sake of brevity.

Acknowledgements. This report is the result of a summer research project at the University of Ottawa which was made financially possible due to the NSERC USRA program. I would like to thank to Dr. Hadi Salmasian for supervising this project; I am sincerely grateful for his mentorship, guidance, and illuminating mathematical discussions.

1 Preliminaries

1.1 Partitions

In this section, we introduce relevant notions and terminology related to *integer partitions*, which are used to parametrize many important families of symmetric polynomials, and present basic notation that we will use throughout. Our choice of notation and terminology will typically follow that of [6]. The current introduction will only introduce those concepts which will be used fairly often in later sections; a more exhaustive treatment of the basic theory of partitions can be found in [6, p. 1-17].

Definition 1.1. An *integer partition* is a finite or infinite sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots)$ of nonnegative integers which is weakly decreasing; that is, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq \cdots \geq 0$.

We call a nonzero term λ_i a *part* of the partition λ. The *length* (or *height*) of a partition, denoted $l(\lambda)$, is the number of nonzero terms. The sum $|\lambda| := \sum_i \lambda_i$ of all the terms is the *weight*; if $|\lambda| = n$ for some integer n, we say that λ is a partition of n and write $\lambda \vdash n$.

Notation. Let $\lambda := (\lambda_1, \ldots, \lambda_n)$ be an integer partition. We may write $\lambda = (1^{a_1} 2^{a_2} 3^{a_3} \cdots)$, where the *multiplicity* $a_i := \#\{j \mid \lambda_j = i\}$ counts the number of parts of the partition λ which are equal to the integer i.

There are several common partitions and weakly decreasing integer sequences which we will often use. For convenience we will often denote the *empty partition* $(0, 0, 0, \ldots)$ simply by 0. We denote by $\delta^{(n)}$ the following integer sequence of length $n - 1$:

$$\delta^{(n)} := (n-1, n-2, \ldots, 2, 1, 0)$$

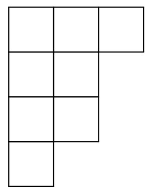
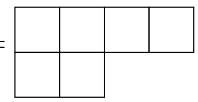
We will often omit the superscript when the length n is clear from context. It is often useful to define new sequences in terms of the sums of sequences of the same length; we do so componentwise. For instance, if $\lambda_1 := (m_1, m_2, \ldots, m_n)$ and $\lambda_2 := (p_1, p_2, \ldots, p_n)$, then we define

$$\lambda_1 + \lambda_2 := (m_1 + p_1, m_2 + p_2, \ldots, m_n + p_n)$$

We may also define $\mathbb{Z}$-scalar multiples and differences of integer partitions in an analogous way, provided that the resulting sequence is still weakly decreasing and each term is a nonnegative integer. For sequences of different length, we may apply a similar approach, first appending 0's to the end of the shorter partition so that the lengths are equal. In general it is customary to consider partitions which differ only by the number of trailing zeroes as the same.

A useful way to represent partitions is diagramatically in so-called *Young diagrams* (sometimes called *Ferrers* diagrams in the literature). This more geometric picture of partitions will be the starting point for many interesting combinatorial investigations beginning in Section 4. Given a partition $\lambda = (\lambda_1, \ldots, \lambda_n)$, the corresponding Young diagram will be a collection of boxes arranged in $l(\lambda)$ rows and λ_1 columns. In particular, we construct the Young diagram of λ by first forming a row consisting of λ_1 horizontally arranged empty boxes. Below this, we place a row of λ_2 boxes, aligned at the left, below this a row of λ_3 boxes, and so on until the bottommost row consisting of λ_n boxes.

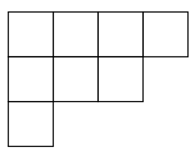
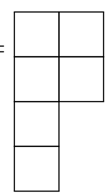
Example. If we have partitions $\lambda = (3, 2, 2, 1)$ and $\mu = (4, 2)$, then we have the respective diagrams

$\lambda =$  and $\mu =$ 

Remark. An alternate convention, reminiscent of coordinate geometry, places the rows of the Young diagram from bottom to top as the length of each row decreases. We will exclusively use the previous convention, in which rows of decreasing length are placed from top to bottom.

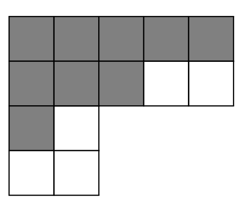
Given this geometric picture of integer partitions, we define the *conjugate* (or *dual*) of a partition λ to be the partition obtained by exchanging columns and rows. We denote the conjugate partition by λ', although it is worth noting that the notation $\tilde{\lambda}$ is also commonly used in the literature.

Example. If λ and μ are as in the previous example, then $\lambda' = (4, 3, 1)$ and $\mu' = (2, 2, 1, 1)$:

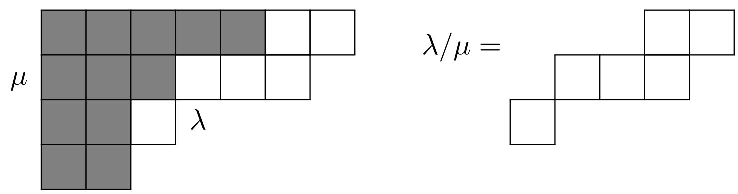
$\lambda' =$  $\mu' =$ 

Algebraically, we may express the parts of the conjugate partition by $\lambda'_i = \#\{j \mid \lambda_j \geq i\}$.

Another useful construction which is motivated by the geometric picture of diagrams is the notion of a *skew diagram*; these will be used extensively in Section 4. First we define the *containment* partial order on the set of partitions as follows. If λ and μ are partitions, we say that $\mu \subset \lambda$ if $\mu_i \leq \lambda_i$ for each i. As usual, we add as many trailing zeroes as necessary to the shorter of the two partitions while making this comparison. More geometrically, $\mu \subset \lambda$ means that if the diagram of μ is superimposed on the diagram of λ so that they are aligned at the top and left, then the diagram of μ is "contained" in λ. For example, one can easily see from the diagram below that $(5, 3, 1) = \mu \subset \lambda = (5, 5, 2, 2)$, where the shaded boxes correspond to the diagram of μ.



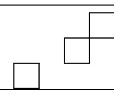
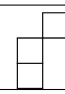
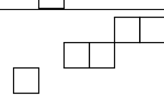
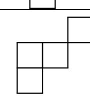
If λ and μ are partitions such that $\mu \subset \lambda$, then we may define the *skew diagram* μ/λ to be the difference of their diagrams. That is, λ/μ is the diagram comprising the boxes of λ which are not boxes of μ. This is generally not a Young diagram, and is not necessarily connected. In the example above, the unshaded boxes represent the skew diagram $(5,5,2,2)/(5,3,1)$. As another example, if $\lambda = (7,6,3,2)$ and $\mu = (5,3,2,2)$, then we have



A particularly useful type of skew diagram are *strips*.

Definition 1.2. Let $\lambda \supset \mu$ be partitions such that λ/μ is a skew diagram consisting of k boxes. We say that λ/μ is a *horizontal* (resp. *vertical*) *k-strip* if no two boxes of λ/μ lie in the same column (resp. row). Algebraically, λ/μ is a horizontal (resp. vertical) strip if and only if $\lambda_i' - \mu_i' \leq 1$ (resp. $\lambda_i - \mu_i \leq 1$) for each $i \geq 1$.

Example. The following table includes examples of skew diagrams which are horizontal or vertical strips, or neither.

	Horizontal Strip	Not a Horizontal Strip
Vertical Strip		
Not a Vertical Strip		

It is often useful to refer to the individual boxes of Young diagrams. In doing so, we follow the 1-based (row,column) indexing for entries of matrices. For example, the rightmost box in the Young diagram of λ above is in the (1,7) position, while the white box immediately above the λ label is in the (2,4) position. When referring to the boxes of a skew diagram λ/μ, we carry over the indices from λ. For instance, in the prior example $\lambda/\mu = (7,6,3,2)/(5,3,2,2)$, the rightmost box of λ/μ is still in the (1,7) position, *not* the (1,5) position.

2 Symmetric Polynomials

Symmetric polynomials, and their infinite variable generalizations, will be our primary algebraic object of study. The purpose of this section is to introduce some of the classical theory of symmetric polynomials, with a focus on introducing several important bases. In the final section 2.7 we outline the formal construction of symmetric "functions" in infinitely many variables. Sections 2.1-2.5 are adapted from Procesi's exposition of the topic; the relevant material can be found in [7, p. 7-9, 19-24]. Sections 2.6 and 2.7 are adapted from [6, p. 19-23] and [6, p. 17-19], respectively.

2.1 Elementary Symmetric Polynomials

Let $R = \mathbb{Z}[x_1, \ldots, x_n]$ be the ring of n-variable polynomials over $\mathbb{Z}$, and consider the action of the symmetric group S_n on R by permutation of the variables. We are particularly interested in the elements of R that are invariant under this action; we call these invariant elements *symmetric polynomials*. More explicitly, if $f(x_1, \ldots, x_n)$ is a polynomial in R, we define the action of S_n on R by

$$\sigma \cdot f(x_1, \ldots, x_n) := f(x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(n)})$$

where $\sigma \in S_n$ is a permutation. For expository purposes, we provide the following basic introductory definition, providing a slightly more formal treatment in section 2.5.

Definition 2.1 (Symmetric Polynomial)**.** Consider the action of S_n on $R = \mathbb{Z}[x_1, \ldots, x_n]$ by permutation of variables, as defined above. We call a polynomial $f \in R$ a *symmetric polynomial* if $\sigma \cdot f = f$ for every $\sigma \in S_n$.

Remark. The scope of the above definition need not be limited to polynomials over $\mathbb{Z}$. Indeed, the S_n-action defined above can be defined on any polynomial ring, and consequently symmetric polynomials can be defined (as the invariant elements under this action) over any ring of coefficients. For the vast majority of the results we will prove, it is necessary and sufficient that the ring of coefficients be commutative. For a select few results, it may also be necessary (and sufficient) for the ring of coefficients to be a characteristic 0 integral domain.

In what follows, we will typically consider symmetric polynomials to be symmetric polynomials having coefficients in $\mathbb{Z}$. The main exception to this tendency is in Section 2.2, where we consider coefficients in $\mathbb{Q}$, and Section 6, where our ring of coefficients will consist of multivariate rational functions.

We often denote the set of symmetric polynomials in n variables by Λ_n or, following a more general notation from group theory, $\mathbb{Z}[x_1, \ldots, x_n]^{S_n}$. The set of symmetric polynomials in n variables is a subring of the space of polynomials $\mathbb{Z}[x_1, \ldots, x_n]$.

There are several noteworthy classes of symmetric functions. One historically notable example are the *elementary symmetric polynomials*. If we consider polynomials over $\mathbb{Z}$ in variables x_i with an extra indeterminate t, we define the elementary symmetric functions implicitly by the generating function

$$E(t) := \prod_{i=1}^{\infty}(1+tx_i) := \sum_{i=0}^{\infty} e_i t^i \tag{2.1.1}$$

To obtain the $n+1$ elementary symmetric polynomials $e_i := e_i(x_1,\dots,x_n)$, $0 \le i \le n$, we simply set $x_i = 0$ (and consequently $e_i = 0$) for each $i > n$, which gives the following identity:

$$E(t) = \prod_{i=1}^{n}(1+tx_i) = \sum_{i=0}^{n} e_i t^i \tag{2.1.2}$$

By commutativity, we have $\prod_{i=1}^n(1+tx_i) = \prod_{i=1}^n(1+tx_{\sigma^{-1}(i)})$ for any permutation $\sigma \in S_n$. Hence the polynomials e_i are invariant under permutation of the variables x_i, and are therefore symmetric polynomials.

Example. It is clear that $e_0(x_1,\dots,x_n) = 1$ for any n. If $n = 2$, we have

$$E(t) = (1+tx_1)(1+tx_2) = 1 + (x_1+x_2)t + (x_1x_2)t^2$$

That is, $e_1(x_1,x_2) = x_1 + x_2$ and $e_2(x_1,x_2) = x_1x_2$.

The process of computing elementary symmetric functions using the implicit definition is quite tedious for larger values of n. Fortunately, the i-th elementary symmetric function in n variables admits a general combinatorial expression as the sum of all products of i distinct variables from $x_1,\dots,x_n$. More precisely, we have

Proposition 2.1. *For $1 \le i \le n$, we have*

$$e_i(x_1,\dots,x_n) = \sum_{1\le a_1<a_2<\cdots<a_i\le n} x_{a_1}x_{a_2}\cdots x_{a_i}$$

Proof. We proceed by induction on n; if $n = 1$ the conclusion is trivial. Now let $1 \le i \le n$ for some $n > 1$. Using the notation $e_i^{(n)} := e_i(x_1,\dots,x_n)$ for convenience, by (2.1.2) we have

$$\sum_{i=0}^{n} e_i^{(n)} t^i = \prod_{i=1}^{n}(1+tx_i) = (1+tx_n)\sum_{i=0}^{n-1} e_i^{(n-1)} t^i = \sum_{i=1}^{n}\left(e_{i-1}^{(n-1)}x_n + e_i^{(n-1)}\right)t^i$$

Comparing coefficients gives the useful recursive identity, valid for all $1 \le i \le n$:

$$e_i^{(n)} = e_{i-1}^{(n-1)}x_n + e_i^{(n-1)} \tag{2.1.3}$$

By our induction hypothesis, $e_i^{(n-1)}$ is the sum of all products $x_{a_1}\cdots x_{a_i}$ of i distinct variables in $\{x_1,\dots,x_{n-1}\}$. Evidently, this is precisely the sum of all products of i distinct variables in $\{x_1,\dots,x_n\}$ which do not contain x_n as a factor. Likewise, $e_{i-1}^{(n-1)}$ is the sum of all products of $i-1$ distinct variables in $\{x_1,\dots,x_{n-1}\}$. Multiplying by x_n consequently gives all products of i distinct variables in $\{x_1,\dots,x_n\}$ which contain x_n as a factor. Adding these sums together gives the sum of all products of i distinct variables in $\{x_1,\dots,x_n\}$; the conclusion follows by induction. $\square$

Example. Using the previous proposition, we may quickly verify the results of the previous example. We also readily see that the following hold:

$$\begin{aligned} e_2(x_1,x_2,x_3) &= x_1x_2 + x_1x_3 + x_2x_3 \\ e_3(x_1,x_2,x_3,x_4) &= x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 \\ e_2(x_1,x_2,x_3,x_4,x_5) &= x_1x_2 + x_1x_3 + x_1x_4 + x_1x_5 + x_2x_3 + x_2x_4 + x_2x_5 + x_3x_4 + x_3x_5 + x_4x_5 \end{aligned}$$

The elementary symmetric polynomials have received great interest historically in solving classical problems in algebra dealing with the roots of polynomials. In fact, expressions of the form $e_1(x_1,x_2) = x_1 + x_2$ and $e_2(x_1,x_2) = x_1x_2$ were first used by the ancient Mesopotamians approximately 4000 years ago to solve quadratic equations. Likewise, the more general study of symmetric functions arose within the past few centuries to study polynomial roots. Observe that

$$t^n E\left(-\tfrac{1}{t}\right) = t^n \prod_{i=1}^{n} \left(1 - \tfrac{x_i}{t}\right) = \prod_{i=1}^{n} t\left(1 - \tfrac{x_i}{t}\right) = \prod_{i=1}^{n} (t - x_i)$$

That is, the polynomial $t^n E\left(-\frac{1}{t}\right)$ has the variables $x_1, \ldots, x_n$ as its roots. However, by definition of E, we have

$$\prod_{i=1}^{n}(t - x_i) = t^n E\left(-\tfrac{1}{t}\right) = t^n \sum_{i=0}^{n} e_i \left(-\tfrac{1}{t}\right)^i = \sum_{i=0}^{n} (-1)^i e_i t^{n-i} \tag{2.1.4}$$

The equation (2.1.4) establishes an explicit correspondence between a polynomial and its roots. Let $f := \sum_{i=0}^{n} a_i t^{n-i}$ be a degree n polynomial over a field, say $\mathbb{R}$ or $\mathbb{C}$. Note that in our choice of indexing, a_0 is the leading coefficient of f as opposed to the constant term. If F is a splitting field for f, then we may write $f = a_0 \prod_{i=1}^{n}(t - \alpha_i)$, where $\alpha_1, \ldots, \alpha_n$ are the roots of f in the field F, counted with multiplicity. Equation (2.1.4) reveals that the coefficients a_i have an explicit form. In particular, $a_i = (-1)^i a_0 e_i(\alpha_1, \ldots, \alpha_n)$.

Consequently, many properties of the roots of polynomials can be deduced by studying expressions involving elementary symmetric functions in the roots (*i.e.* expressions involving the coefficients); the roots themselves need not be known explicitly. This is the premise of several important notions in the classical theory, such as the *discriminant* and *resultant* [7, p. 22-27]. We explore the former in greater depth in section 2.3.

2.2 Power Sums and Newton's Identities

Another important family of symmetric polynomials are the power sum symmetric polynomials. The k-th *power sum* p_k in n variables is defined to be the polynomial $p_k(x_1, \ldots, x_n) := \sum_{j=1}^{n} x_j^k$. Clearly each p_k is invariant under permutation of the variables.

The k-th power sum symmetric function is sometimes called the k-th *Newton function*. This is in reference to Isaac Newton, to whom we attribute a sequence of recursive relations, the so-called *Newton's identities*, which allow us to express any power sum symmetric function p_k as a polynomial in the elementary symmetric functions e_i, and vice versa. We will derive these identities by working in the ring of formal power series. Beginning with (2.1.2), we take the logarithm of both sides, which gives

$$\log\left(\sum_{i=0}^{n} e_i t^i\right) = \log\left(\prod_{i=1}^{n}(1+tx_i)\right) = \sum_{i=1}^{n}\log(1+tx_i)$$

Taking the formal derivative with respect to t yields

$$\frac{\sum_{i=1}^{n} i e_i t^{i-1}}{\sum_{i=0}^{n} e_i t^i} = \sum_{i=1}^{n}\frac{x_i}{1+tx_i}$$

The righthand side of the above equation can be expanded as follows:

$$\sum_{i=1}^{n}\frac{x_i}{1+tx_i} = \sum_{i=1}^{n}\left(x_i\sum_{h=0}^{\infty}(-tx_i)^h\right) = \sum_{i=1}^{n}\sum_{h=0}^{\infty}(-t)^h x_i^{h+1} = \sum_{h=0}^{\infty}(-t)^h p_{h+1}$$

Hence we obtain the following identity, setting $e_i = 0$ for each $i > n$:

$$\sum_{i=1}^{n} i e_i t^{i-1} = \left(\sum_{h=0}^{\infty}(-t)^h p_{h+1}\right)\left(\sum_{i=0}^{n} e_i t^i\right) = \left(\sum_{i=0}^{\infty}(-1)^i p_{i+1}t^i\right)\left(\sum_{i=0}^{\infty} e_i t^i\right) = \sum_{i=0}^{\infty} a_i t^i$$

In the rightmost expression, we have $a_i := \sum_{j=0}^{i}(-1)^j p_{j+1}e_{i-j}$. Comparing the m-th coefficients of the leftmost and rightmost sides of the above equation, we obtain the following identity:

$$(m+1)e_{m+1} = \sum_{j=0}^{m}(-1)^j p_{j+1}e_{m-j} = (-1)^m p_{m+1} + \sum_{i=1}^{m}(-1)^{i-1}p_i e_{m+1-i}$$

By rearranging the above equation, we obtain Newton's identities:

$$p_{m+1} = (-1)^m(m+1)e_{m+1} + \sum_{i=1}^{m}(-1)^{m+i}p_i e_{m+1-i} \tag{2.2.1}$$

The identity (2.2.1) is valid for any $m \geq 1$, provided that we follow the convention that $e_i = 0$ for each $i > n$, where n is the number of variables. If we opt to work in the field of quotients of our (characteristic 0) ring of coefficients, then we obtain a similar result which allows us to compute the elementary symmetric functions e_i in terms of the power sums p_k:

$$e_{m+1} = \frac{\sum_{i=0}^{m}(-1)^i p_{i+1}e_{m-i}}{m+1} \tag{2.2.2}$$

2.3 Classical Application: The Discriminant

As we alluded to in section 2.1, many properties of the roots of polynomials can be deduced without explicit knowledge of the roots, provided that the property admits some sort of expression in terms of elementary symmetric functions in the roots. To be concrete, we will consider the example of a polynomial's *discriminant*.

Let $f = \sum_{i=0}^{n} a_i t^{n-i}$ be a degree n polynomial over a field F. If K is a splitting field for f, then we may write $f = a_0 \prod_{i=1}^{n}(t - \alpha_i)$, where $\alpha_i \in K$ are the n roots of f, counted with multiplicity. One property of the roots that we may be interested in is whether or not f has any repeated roots. It is clear that f has a repeated root if and only if the polynomial $V(x) := \prod_{i<j}(x_i - x_j)$ is equal to 0 when evaluated at $x_i = \alpha_i$.

The polynomial V is called the *Vandermonde determinant*, and has applications in many areas of mathematics. It will be of particular interest to us in the theory of symmetric functions, in particular in section 3 when we introduce the Schur polynomials. As the name suggests, the polynomial $V(x)$ is readily obtained as the determinant of a matrix. Namely, $V(x) = \det(A)$, where we define the Vandermonde matrix A by setting $A_{i,j} := x_j^{n-i}$ as follows:

$$A := \begin{pmatrix} x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \\ x_1^{n-2} & x_2^{n-2} & \cdots & x_n^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} \tag{2.3.1}$$

Evidently, V is not a symmetric polynomial. In particular, if $(ij) \in S_n$ is a transposition, then $(ij)V = -V$. We call a polynomial with this property an *alternating* (or *antisymmetric*) polynomial. We have the following equivalent definition:

Definition 2.2 (Alternating Polynomial)**.** A polynomial $f(x_1, \ldots, x_n)$ is said to be an *alternating polynomial* if for any $\sigma \in S_n$,

$$\sigma f(x_1, \ldots, x_n) := f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \operatorname{sgn}(\sigma) f(x_1, \ldots, x_n)$$

The fact that V is alternating can be deduced from the fact that V is the determinant of the Vandermonde matrix, and a transposition of two variables x_i, x_j corresponds to a transposition of the i-th and j-th columns of A.

While V is not a symmetric polynomial, it is clear that $V^2 = \det(AA^T)$ is, and moreover V^2 maintains the same property that $V^2(\alpha_1, \ldots, \alpha_n) = 0$ iff f has a repeated root. Observe also that the entries of the matrix AA^T are all Newton functions. In particular, we have

$$AA^T_{i,j} = \sum_{k=1}^{n} A_{i,k} A_{j,k} = \sum_{k=1}^{n} x_k^{n-i} x_k^{n-j} = \sum_{k=1}^{n} x_k^{2n-(i+j)} = p_{2n-(i+j)}$$

By Newton's identities, we can express each entry of the matrix AA^T as a polynomial in the elementary symmetric functions e_i. Consequently, the same is true of the detmerinant V^2.

Definition 2.3 (Discriminant)**.** For $n \geq 1$, a polynomial D in the variables $e_1, \ldots, e_n$ such that $D(e_1, \ldots, e_n) = V^2(x_1, \ldots, x_n)$ is called the *discriminant*.

Remark. For the discriminant to be well-defined, it is important that this polynomial expression be unique. This is in fact the case, as we will establish in section 2.4.

Example. Let $n = 2$. Then $A = \left(\begin{smallmatrix} x_1 & x_2 \\ 1 & 1 \end{smallmatrix}\right)$, and $AA^T = \begin{pmatrix} x_1^2+x_2^2 & x_1+x_2 \\ x_1+x_2 & 2 \end{pmatrix} = \left(\begin{smallmatrix} p_2 & p_1 \\ p_1 & p_0 \end{smallmatrix}\right)$. Hence $V^2(x_1, x_2) = p_2p_0 - p_1^2 = 2p_2 - p_1^2$. By (2.2.1), we have $p_1 = e_1$ and $p_2 = e_1^2 - 2e_2$. Hence $V^2(x_1, x_2) = 2(e_1^2 - 2e_2) - e_1^2 = e_1^2 - 4e_2$. That is, $D(e_1, e_2) = e_1^2 - 4e_2$.

Recall that in section 2.1, we showed that the coefficients of our polynomial f are related to the roots by the explicit formula $a_i = (-1)^i a_0 e_i(\alpha_1, \ldots, \alpha_n)$. That is, the discriminant can be expressed in terms of the coefficients of f. Hence, given an arbitrary polynomial, we can determine whether or not it has any repeated roots simply by evaluating a polynomial expression of its coefficients.

Example. Consider the case of the previous example with $n = 2$. Suppose that we have a quadratic polynomial $ax^2 + bx + c$ $(a \neq 0)$ with roots α_1, α_2. By the previous remark, we have the identities $b = -ae_1(\alpha_1, \alpha_2)$ and $c = ae_2(\alpha_1, \alpha_2)$. Since the monic polynomial $x^2 + \frac{b}{a}x + \frac{c}{a}$ has the same set of roots, we see that $\alpha_1 = \alpha_2$ iff $D(e_1(\alpha_1, \alpha_2), e_2(\alpha_1, \alpha_2)) = 0$. But

$$D(e_1(\alpha_1, \alpha_2), e_2(\alpha_1, \alpha_2)) = D\left(\tfrac{-b}{a}, \tfrac{c}{a}\right) = \left(\tfrac{-b}{a}\right)^2 - 4\left(\tfrac{c}{a}\right)$$

Multiplying through by a^2, we see that $\alpha_1 = \alpha_2$ iff $b^2 - 4ac = 0$.

2.4 Fundamental Theorem of Symmetric Polynomials

As we saw in section 2.2, any power sum symmetric function can be expressed as a polynomial in the elementary symmetric functions. In fact, this result applies more generally to any symmetric polynomial. In particular, we have the following theorem:

Theorem 2.2 (Fundamental Theorem of Symmetric Polynomials)**.** *Any symmetric polynomial in n variables can be expressed uniquely as a polynomial in the elementary symmetric functions $e_1, \ldots, e_n$ with integer coefficients.*

Proof. The proof is by induction, simultaneously on the degree d and the number of variables n. If $d = 0$ or $n = 0$, then the conclusion is trivial, so we proceed directly to the inductive case. Let f be a symmetric polynomial of degree $d \geq 1$ in $n \geq 1$ variables $x_1, \ldots, x_n$, and suppose for our induction hypothesis that the theorem holds for all symmetric polynomials of degree $d' < d$ in at most n variables, as well as for all polynomials of degree at most d in $m < n$ variables.

Consider the homomorphism $\pi_{n,n-1} : \mathbb{Z}[x_1, \ldots, x_n] \to \mathbb{Z}[x_1, \ldots, x_{n-1}]$ given by evaluating f at $x_n = 0$. If $\pi_{n,n-1}(f) := f(x_1, \ldots, x_{n-1}, 0) = 0$, then $x_n \mid f$. Consequently, $x_i \mid f$ for each i by symmetry, so in particular $e_n^{(n)} \mid f$. Dividing by e_n, we obtain a polynomial of strictly smaller degree. Our induction hypothesis then implies the existence of a unique polynomial $g(e_1, \ldots, e_n)$ with coefficients in $\mathbb{Z}$ such that $f = e_n g$.

If instead $\pi_{n,n-1}(f) := f(x_1, \ldots, x_{n-1}, 0) \neq 0$, then $\pi_{n,n-1}(f)$ is a polynomial of degree at most d in $n-1$ variables. Our induction hypothesis implies that there exists a unique polynomial p in $n-1$ variables such that $p(e_1^{(n-1)}, \ldots, e_{n-1}^{(n-1)}) = f(x_1, \ldots, x_{n-1}, 0)$. By (2.1.3), we have $e_i^{(n)} = e_{i-1}^{(n-1)} x_n + e_i^{(n-1)}$ for all $1 \leq i \leq n$. In particular, when we evaluate at $x_n = 0$, we see that $e_i^{(n)} = e_i^{(n-1)}$. Hence $f - p(e_1^{(n)}, \ldots, e_{n-1}^{(n)})$ is a symmetric polynomial which vanishes at $x_n = 0$, which allows us to revert to the previous case. □

Example 2.1. For any $n \geq 0$, the symmetric polynomial $f(x_1, \ldots, x_n) := \sum_{i \neq j} x_i x_j^2$ can be uniquely written as $f(x_1, \ldots, x_n) = e_1^{(n)} e_2^{(n)} - 3e_3^{(n)}$.

Proof. By Lemma 2.8, it suffices to show that this identity holds for $n = 3$. We have

$$f(x_1, x_2, x_3) = x_1^2 x_2 + x_1 x_3^2 + x_2 x_1^2 + x_2 x_3^2 + x_3 x_1^2 + x_3 x_2^2$$

Evaluating at $x_3 = 0$ gives

$$\pi_3(f(x_1, x_2, x_3)) := f(x_1, x_2, 0) = x_1^2 x_2 + x_2 x_1^2 = (x_1 + x_2)(x_1 x_2) = e_1^{(2)} e_2^{(2)}$$

By the proof of Theorem 2.2, we see that $p(y_1, y_2) := y_1 y_2$ is the unique polynomial over $\mathbb{Z}$ satisfying $p(e_1, e_2) = \pi_3(f)$. Consequently, we have $\pi_3(f - e_1^{(3)} e_2^{(3)}) = 0$, and hence $e_3^{(3)} \mid f - e_1^{(3)} e_2^{(3)}$. In particular, we have

$$\begin{aligned}
f(x_1, x_2, x_3) - e_1^{(3)} e_2^{(3)} &= x_1^2 x_2 + x_1 x_3^2 + x_2 x_1^2 + x_2 x_3^2 + x_3 x_1^2 + x_3 x_2^2 \\
&\qquad - (x_1 + x_2 + x_3)(x_1 x_2 + x_1 x_3 + x_2 x_3) \\
&= x_1^2 x_2 + x_1 x_3^2 + x_2 x_1^2 + x_2 x_3^2 + x_3 x_1^2 + x_3 x_2^2 \\
&\qquad - (x_1^2 x_2 + x_1 x_3^2 + x_2 x_1^2 + x_2 x_3^2 + x_3 x_1^2 + x_3 x_2^2 + 3 x_1 x_2 x_3) \\
&= -3 x_1 x_2 x_3 = -3 e_3^{(3)}
\end{aligned}$$

Hence $f(x_1, x_2, x_3) = e_1^{(3)} e_2^{(3)} - 3e_3^{(3)}$, and consequently the identity $f(x_1, \ldots, x_n) = e_1^{(n)} e_2^{(n)} - 3e_3^{(n)}$ holds for each $n \geq 0$. □

Remark. The polynomial f used in the previous example is in fact an example of a *monomial symmetric polynomial* which we will define in the next section. In particular, we have $f(x_1, \ldots, x_n) = m_{(2,1)}(x_1, \ldots, x_n)$.

2.5 Monomial Symmetric Polynomials

Another basis of the symmetric polynomials are the *monomial symmetric polynomials.* In order to introduce these polynomials, and prove that they form a basis for the symmetric polynomials, we will briefly work in the more general setting of group actions.

Let G be a group acting on sets X and Y. Then we have an induced G-action on the set $Y^X = \{f \mid f : X \to Y\}$ of maps from X to Y by setting

$$(g \cdot f)(x) := g \cdot f(g^{-1}x)$$

For our purposes, it suffices to consider the special case where the G-action on Y is trivial. Using the notation ${}^g f := g \cdot f$, the induced G-action on Y^X then reduces to ${}^g f(x) = f(g^{-1}x)$. In this context, we say that a function $f : X \to Y$ is *invariant* if ${}^g f = f$. More explicitly, f is invariant iff $f(x) = f(g^{-1}x)$ for each $g \in G, x \in X$. Equivalently, f is invariant when $f(x) = f(gx)$ for all $g \in G, x \in X$ or, in other words, f is invariant iff it is constant on each G-orbit in X.

Example 2.2. Let X be a finite G-set, and let F be a field (in fact a ring suffices). Identify $x \in X$ with the characteristic function $\chi_{\{x\}} \in F^X$. It is easy to verify that in this way, X becomes a basis for F^X as a vector space with operations defined pointwise. The induced G-action on F^X is by linear maps which permute the basis elements. In this case, we call F^X a *permutation representation.*

Proposition 2.3. *Let $(F^X)^G$ denote the set of elements of a permutation representation F^X which are invariant under the action of G. Then $(F^X)^G$ is a subspace of F^X with a basis given by the characteristic functions of the G-orbits in X.*

Proof. Recall that $f \in F^X$ is invariant if and only if f is constant on the G-orbits in X. If f is invariant, so that $f(x) = a_i$ for each x in a given G-orbit $\mathcal{O}_i$, then we may uniquely write $f = \sum_i a_i \chi_{\mathcal{O}_i}$, where $\chi_{\mathcal{O}_i}$ is the characteristic function of the orbit $\mathcal{O}_i$. □

Remark. We may express the conclusion of Proposition 2.3 more concretely as follows. If $\mathcal{O}$ is a G-orbit, and we let $u_{\mathcal{O}} := \sum_{x \in \mathcal{O}} x$, then $\{u_{\mathcal{O}}\}$ is a basis for $(F^X)^G$.

We will now relate the preceding result to the theory of symmetric functions via the following discussion, in which we offer a formal treatment of monomials.

Definition 2.4. A *monomial* is a function $M : \{1, \dots, n\} \to \mathbb{N}$. The set of monomials form a commutative monoid under pointwise addition.

Notation. We indicate by x_k the monomial which is the characteristic function of the set $\{k\}$. Addition of monomials is written multiplicatively. Given a monomial M such that $M(k) = h_k$, we write $M = x_1^{h_1} x_2^{h_2} \cdots x_n^{h_n}$.

We define the *degree* of a monomial to be the quantity $\sum_i h_i$. Evidently monomials may be written unambiguously as n-tuples $(h_1, \dots, h_n)$. Moreover, any degree d monomial is equivalent under the induced action of S_n to a monomial $M := (k_1, \dots, k_n)$ with $k_1 \geq k_2 \geq \cdots \geq k_n$. In particular, we may associate M with a unique partition $\lambda(M) \vdash d$ with $l(\lambda(M)) \leq n$.

If we consider the set X of monomials under the induced action of S_n, and let F be a commutative ring, we may consider the corresponding permutation representation. The resulting permutation representation F^X is in fact isomorphic to the polynomial ring $F[x_1,\ldots,x_n]$. By our previous discussion, $F[x_1,\ldots,x_n]$ is an F-module with the monomials (identified with their respective characteristic functions) as a basis.

The symmetric group S_n acts on the set of monomials X by $\sigma M(k) := M(\sigma^{-1}(k))$. This action is compatible with addition of monomials. Moreover, we see that the S_n-action induced on the permutation representation $F[x_1,\ldots,x_n]$ is such that $\sigma f(x_1,\ldots,x_n) = f(x_{\sigma^{-1}(1)},\ldots,x_{\sigma^{-1}(n)})$. We may more formally define the set Λ_n of symmetric polynomials in n variables to be the set of invariant elements of this permutation representation (cf. 2.1). Proposition 2.3 gives an explicit basis for this space, which we call the *monomial symmetric polynomials*.

As we remarked earlier, the orbit of a monomial $x_1^{k_1}\cdots x_n^{k_n}$ can be uniquely associated with a partition $\lambda = (h_1 \geq h_2 \geq \ldots \geq h_n)$ of the degree $\sum_i k_i$. Moreover, the characteristic function of an orbit, given our identification in $F[x_1,\ldots,x_n]$ of the monomial M with the characteristic function of the singleton set $\{M\}$, is the sum of all monomials in the orbit. The following definition concretely describes the characteristic function of an S_n-orbit λ, interpreted as an element of $F[x_1,\ldots,x_n]$.

Definition 2.5 (Monomial Symmetric Polynomial)**.** Let $\lambda = (\lambda_1,\lambda_2,\ldots,\lambda_n)$ be an integer partition with $l(\lambda) \leq n$. Then we define the corresponding *monomial symmetric polynomial* to be the polynomial

$$m_\lambda(x_1,\ldots,x_n) := \sum_{\sigma\in S_n} x_{\sigma(1)}^{\lambda_1} x_{\sigma(2)}^{\lambda_2}\cdots x_{\sigma(n)}^{\lambda_n}$$

where the sum is taken over all *distinct* permutations σ.

Example. Note that although $|S_3| = 6$, the monomial symmetric polynomials below do not necessarily have 6 monomial terms as the summation in Definition 2.5 is taken over *distinct* permutations, not *all* permutations.

$$\begin{aligned}
m_{(1,1,0)}(x_1,x_2,x_3) &= x_1x_2 + x_1x_3 + x_2x_3\\
m_{(1,1,1)}(x_1,x_2,x_3) &= x_1x_2x_3\\
m_{(3,2,1)}(x_1,x_2,x_3) &= x_1^3x_2^2x_3 + x_1^2x_2^3x_3 + x_1^3x_2x_3^2 + x_1^2x_2x_3^3 + x_1x_2^3x_3^2 + x_1x_2^2x_3^3
\end{aligned}$$

For instance, in the first example $m_{(1,1,0)}$, the action of the permutations (13) and (132) on the monomial $x_1^1x_2^1x_3^0$ are the same, so we only include the term x_2x_3 in the sum once.

As we have previously remarked, Proposition 2.3 implies the following corollary.

Corollary 2.4. *The set $\{m_\lambda \mid l(\lambda) \leq n\}$ of monomial symmetric polynomials is an integral basis for the set Λ_n of symmetric polynomials in n variables.*

2.6 Complete Homogeneous Symmetric Polynomials

In this section we introduce a third basis for the symmetric polynomials, the *complete homogeneous symmetric polynomials* (or simply *complete symmetric polynomials*). These polynomials are in some sense dual to the elementary symmetric polynomials introduced in section 2.1. Like the elementary symmetric polynomials, we define the complete homogeneous symmetric polynomials implicitly with a generating function:

$$H(t) := \prod_{i=1}^{\infty} \frac{1}{1-tx_i} := \sum_{i=0}^{\infty} h_i t^i \tag{2.6.1}$$

As with the elementary symmetric polynomials, the complete symmetric polynomials admit an explicit combinatorial expression.

Proposition 2.5. *Let $n \geq 1$, $i \geq 0$. Then $h_i(x_1, \ldots, x_n)$ is the sum of all degree i monomials in n variables:*

$$h_i(x_1, \ldots, x_n) = \sum_{1 \leq a_1 \leq a_2 \leq \cdots \leq a_i \leq n} x_{a_1} x_{a_2} \cdots x_{a_i} \tag{2.6.2}$$

Equivalently, we have

$$h_i(x_1, \ldots, x_n) = \sum_{|\lambda|=i} m_\lambda(x_1, \ldots, x_n) \tag{2.6.3}$$

where the sum is taken over all partitions λ with $|\lambda| = i$ and $l(\lambda) \leq n$.

Proof. Expanding $(1-tx_i)^{-1}$ as a geometric series, we have $(1-tx_i)^{-1} = \sum_{k=0}^{\infty} x_i^k t^k$. Taking the product over all i, together with the identity (2.6.1), gives

$$\sum_{i=0}^{\infty} h_i t^i = \prod_{i=1}^{\infty} \frac{1}{1-tx_i} = \prod_{i=1}^{\infty} \sum_{k=0}^{\infty} x_i^k t^k$$

Expanding the righthand product, we wish to provide an expression for the coefficient of t^i, namely h_i. This will clearly be a sum of monomials in the x_j. Now suppose that one of the monomials appearing in this coefficient is $x_{j_1}^{k_1} \cdots x_{j_m}^{k_m}$. For each l, $1 \leq l \leq m$, the exponent k_l is the sum of the exponents of the x_{j_l}, summed over each occurence of x_{j_l} in the product which produced this monomial. Since the exponent of x_{j_l} in any given term of the product is equal to the exponent of t, it follows that the sum of the k_l is equal to the sum of the exponents of t which occured in the product. This is precisely the exponent of t in the lefthand series, namely i. That is, each monomial appearing in this sum has degree i. Since the product is taken over all combinations of x_i^k, then consequently every degree i monomial in the x_j appears in this sum.

To reiterate, the coefficient of t^i (that is, h_i) in the lefthand series is precisely the sum of all degree i monomials in the variables x_j. If we permute the terms of this sum to collect terms which lie in the same symmetric group orbit, we have the equivalent formulation

$$h_i = \sum_{|\lambda|=i} m_\lambda$$

where m_λ is the infinite-variable monomial symmetric function which corresponds to the partition λ. How to interpret such a symmetric function precisely will be handled in section 2.7.

Setting $x_i = 0$ for all $i > n$ gives the corresponding identities 2.6.2 and 2.6.3 for symmetric polynomials in n variables. $\square$

Example. Using the explicit expresison (2.6.2), we can easily compute

$$h_2(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$$

Using the expression (2.6.3), we may likewise compute

$$\begin{aligned} h_3(x_1, x_2, x_3) &= m_{(3,0,0)}(x_1, x_2, x_3) + m_{(2,1,0)}(x_1, x_2, x_3) + m_{(1,1,1)}(x_1, x_2, x_3) \\ &= (x_1^3 + x_2^3 + x_3^3) + (x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2) + (x_1x_2x_3) \end{aligned}$$

One of the interesting properties of the complete symmetric polynomials is that they act as a sort of dual to the elementary symmetric polynomials. In particular, consider the polynomials E and H which respectively define the elementary and complete symmetric polynomials (or equally well consider their extensions as formal power series). We have

$$H(t)E(-t) = \left(\prod_{k=1}^{n} \frac{1}{1 - tx_k}\right)\left(\prod_{k=1}^{n} 1 - tx_k\right) = 1 \tag{2.6.4}$$

Expanding the respective series expressions, we see that

$$1 = \left(\sum_{k=0}^{\infty} h_k t^k\right)\left(\sum_{k=0}^{\infty} e_k(-t)^k\right) = \sum_{k=0}^{\infty}\left(\sum_{i=0}^{k}(-1)^i e_i h_{k-i}\right) t^k$$

Comparing coefficients, we obtain the following identity, valid for each $k \geq 1$:

$$\sum_{i=0}^{k}(-1)^i e_i h_{k-i} = 0 \tag{2.6.5}$$

Remark. The above expression is in fact valid as an identity for symmetric functions (cf. section 2.7) and consequently (2.6.5) holds as an identity for symmetric polynomials in any number of variables n. To be explicit, the following symmetric polynomial identity holds for each $n, k \geq 1$:

$$\sum_{i=0}^{k}(-1)^i e_i^{(n)} h_{k-i}^{(n)} = 0$$

where we denote $h_i^{(n)} := h_i(x_1, \ldots, x_n)$, as with the elementary symmetric polynomials. A simple corollary of this identity is obtained by letting $k = 1$. We see then that $e_1^{(n)} = h_1^{(n)}$ for each $n \geq 1$. This can also be seen by comparing the combinatorial expressions in Propositions 2.1 and 2.5.

The expression (2.6.5) establishes an essential duality between the elementary and complete symmetric polynomials. This duality is exploited, and perhaps made more explicit, by the following result.

Theorem 2.6. *Any symmetric polynomial in n variables can be uniquely expressed as a polynomial, with coefficients in $\mathbb{Z}$, in the complete homogeneous symmetric polynomials $h_1, \ldots, h_n$ in n variables.*

Proof. By the Fundamental Theorem of Symmetric Polynomials 2.2, the same result is true for the elementary symmetric polynomials $e_1, \ldots, e_n$. As the e_i form a basis for the space Λ_n of n-variable symmetric polynomials, then a homomorphism on this space can be defined unambiguously by its action on the e_i. Making use of this fact, we define a homomorphism $\omega : \Lambda_n \to \Lambda_n$ by defining $\omega(e_k) = h_k$ for each k, $0 \leq k \leq n$. We in fact have that ω is an involution, namely $\omega^2 = \mathrm{Id}$.

Indeed, we may see this by induction on k, noting first that $h_0 = 1 = e_0$. Now consider $k > 0$ and suppose that $\omega^2(e_j) = \omega(h_j) = e_j$ for each $j < k$. By (2.6.5), we have the identity $h_k = \sum_{j=1}^{k} (-1)^{j+1} e_j h_{k-j}$. Applying ω, we obtain

$$\omega^2(e_k) = \omega(h_k) = \omega\left(\sum_{j=1}^{k} (-1)^{j+1} e_j h_{k-j}\right) = \sum_{j=1}^{k} (-1)^{j+1} \omega(e_j)\omega(h_{k-j}) = \sum_{j=1}^{k} (-1)^{j+1} h_j e_{k-j} = e_k$$

and so it follows by induction that ω is an involution. Consequently, ω is invertible and is therefore an automorphism of Λ_n. The fact that the h_i form a basis of Λ_n follows from Theorem 2.2 and the fact that an isomorphism maps bases to bases. □

2.7 Ring of Symmetric Functions

For many of the results we have proved regarding symmetric polynomials, the number of variables is immaterial. As our generating function definitions of the elementary and complete symmetric functions would suggest, it is often natural and useful to consider symmetric polynomials in infinitely many variables. We moreover approach these formal expressions with the idea in mind that an identity that holds for symmetric functions in infinitely many variables can be reduced to an analogous result for symmetric polynomials in n variables by setting all but finitely many variables x_i to 0. The purpose of this section is to construct this so-called *ring of symmetric functions* in infinitely many variables and to provide the useful lemma 2.8 which facilitates the proofs of many identities involving symmetric polynomials.

Consider the set Λ_n of symmetric polynomials in n variables. In light of the fact that the monomial symmetric polynomials m_λ with $l(\lambda) \leq n$ are a $\mathbb{Z}$-basis for Λ_n, we see that Λ_n is a *graded ring*. That is, we may decompose Λ_n as a direct sum

$$\Lambda_n = \bigoplus_{k \geq 0} \Lambda_n^k$$

where Λ_n^k is the set of symmetric polynomials in Λ_n which are homogeneous of degree k, together with the 0 polynomial. Equivalently, Λ_n^k is the $\mathbb{Z}$-module generated by the basis of all monomial symmetric polynomials m_λ such that $\lambda \vdash k$ and $l(\lambda) \leq n$. The above decomposition moreover satisfies the properties that $\Lambda_n^j \Lambda_n^k \subset \Lambda_n^{j+k}$ for each $j, k \geq 0$, and any element $f \in \Lambda_n$ can be written uniquely as $f = \sum_{k \geq 0} f_k$, where each f_k is an element of Λ_n^k.

Recall the map $\pi_{n,n-1}$ introduced in the proof of the Fundamental Theorem 2.2, defined by evaluating the polynomial $f(x_1, \ldots, x_n)$ at $x_n = 0$. We generalize this map as follows. Let $m \geq n$ and consider the homomorphism $\pi_{m,n} : \Lambda_m \to \Lambda_n$ defined by

$$\pi_{m,n}(x_i) := \begin{cases} x_i, & i \leq n \\ 0, & n < i \leq m \end{cases}$$

The effect of $\pi_{m,n}$ on the basis given by monomial symmetric polynomials is easily understood as follows. For all m_λ with $l(\lambda) > n$, we have $\pi_{m,n}(m_\lambda(x_1, \ldots, x_m)) = 0$. On the other hand, $\pi_{m,n}(m_\lambda(x_1, \ldots, x_m)) = m_\lambda(x_1, \ldots, x_n)$ when $l(\lambda) \leq n$. Consequently, $\pi_{m,n}$ is surjective. By restricting to Λ_m^k, we obtain homomorphisms $\pi_{m,n}^k : \Lambda_m^k \to \Lambda_n^k$ for all $k \geq 0$ and $m \geq n$. These maps are always surjective, and are bijective precisely when $m \geq n \geq k$.

We now form the inverse limit $\Lambda^k := \varprojlim_n \Lambda_n^k$ of the $\mathbb{Z}$-modules Λ_n^k relative to the homomorphisms $\pi_{m,n}^k$. Formally, an element of Λ^k is a sequence $\{f_n\}_{n \geq 0}$ with terms $f_n \in \Lambda_n^k$ such that $\pi_{m,n}^k(f_m) = f_n$ whenever $m \geq n$. Intuitively, we may think of an element $f \in \Lambda^k$ as an "infinite variable" extension of a symmetric polynomial in finitely many variables which is homogeneous of degree k. Now define a projection $\pi_n^k : \Lambda^k \to \Lambda_n^k$ which sends f to f_n.

Since $\pi_{m,n}^k$ is an isomorphism whenever $m \geq n \geq k$, it follows that the projection π_n^k is an isomorphism for all $n \geq k$. Consequently, an integral basis for Λ^k is given by all monomial symmetric functions $m_\lambda \in \Lambda^k$ such that $\lambda \vdash k$ and, for all $n \geq k$:

$$\pi_n^k(m_\lambda) = m_\lambda(x_1, \ldots, x_n)$$

Definition 2.6 (Ring of Symmetric Functions)**.** We define the *ring of symmetric functions* Λ in infinitely many variables to be the direct sum

$$\Lambda := \bigoplus_{k \geq 0} \Lambda^k$$

Remark. We see that Λ is the free $\mathbb{Z}$-module generated by the monomial symmetric functions m_λ for all partitions λ.

It is easy to see that Λ has a graded ring structure. By applying the projections π_n^k componentwise, we obtain surjective homomorphisms

$$\pi_n := \bigoplus_{k \geq 0} \pi_n^k : \Lambda^k \to \Lambda_n^k$$

for each $n \geq 0$. By the properties of the maps π_n^k, it follows that π_n is injective (and hence an isomorphism) when restricted to the space $\bigoplus_{k=0}^n \Lambda^k$ of all symmetric functions of degree $k \leq n$.

While it is seldom necessary to consider the foundational content underlying the central construction of this section, the homomorphisms defined therein are particularly useful with regards to the following lemmas, which we have already referred to in Example 2.1, and is useful in proving many identities involving symmetric polynomials.

Lemma 2.7. *Suppose that we have an identity* [1] *involving symmetric polynomials which is homogeneous of degree n and is valid in Λ_n. Then the identity is also valid in Λ, as well as in Λ_m for each $m \geq n$.*

Proof. By our prior remark, the map $\pi_n : \Lambda \to \Lambda_n$ is an isomorphism when restricted to the space of symmetric functions of degree $k \leq n$. That is, we have

$$\bigoplus_{k=0}^{n} \Lambda^k \cong \pi_n \left(\bigoplus_{k=0}^{n} \Lambda^k \right) = \bigoplus_{k=0}^{n} \Lambda_n^k$$

That is, the space of symmetric functions of degree at most n is isomorphic to the space of symmetric polynomials in n variables of degree at most n. Now suppose we have an identity that is homogeneous of degree n which holds in the space of symmetric polynomials in n variables. Then in particular, the identity will still hold in the space of symmetric polynomials in n variables of degree at most n. Then by the isomorphism, the identity will also hold in the space of symmetric functions of degree at most n. By applying projections π_m, which are isomorphisms for $m \geq n$, we see that the identity is also valid in Λ_m for each $m \geq n$. □

In many cases, an identity satisfying the hypotheses of Lemma 2.7 is in fact valid in Λ_m for each $m \geq 0$. A concrete and important class of examples are those of the form of Example 2.1. That is, identities which express a particular symmetric polynomial in terms of a basis which behaves well under the projections $\pi_{m,n}$, such as the elementary symmetric polynomials.

Lemma 2.8. *Let $p_1^{(n)}, \ldots, p_n^{(n)}$ be symmetric polynomials in n variables of degree at most n, and write $p_i^{(m)} := \left(\pi_m \circ \pi_n^{-1}\right)(p_i^{(n)})$. Let f be a symmetric polynomial in n variables which is homogeneous of degree n and let g be an arbitrary polynomial in n variables such that*

$$g(p_1^{(n)}, \ldots, p_n^{(n)}) = f(x_1, \ldots, x_n)$$

Then for any $m \geq 0$, the following identity holds:

$$g(p_1^{(m)}, \ldots, p_n^{(m)}) = \left(\pi_m \circ \pi_n^{-1}\right)(f) := f(x_1, \ldots, x_m)$$

Proof. The lemma essentially amounts to recalling the various properties of the projections π defined throughout this section. We note that the homomorphism $\pi_m \circ \pi_n^{-1}$ is well-defined on the space of

[1] To make the notion of "identity" precise, one may interpret an "identity" to be a first-order sentence in the (possibly enriched) language of rings.

symmetric functions of degree at most n. It is an isomorphism if $m \geq n$, and surjective (but not injective) when $m < n$. If we let $\pi_{n,m} := \pi_m \circ \pi_n^{-1}$, then we in fact recover our original definition for $\pi_{n,m}$ defined for $n \geq m$, and consequently extend this definition unambiguously for $n < m$. In particular, the map $\pi_{n,m}$ is well-defined for all polynomials under consideration in the statement of the lemma for all $m \geq 0$. The conclusion of the lemma follows by applying the homomorphism $\pi_{n,m}$ to both sides of the initial equation. $\square$

Remark. If $p_i^{(n)}$ is taken to be the elementary symmetric polynomial $e_i^{(n)}$ in the statement of the Lemma, then note that $\pi_{n,m}(e_i^{(n)}) = e_i^{(m)}$ for each $m \geq 0$. That is, the elementary symmetric polynomials "behave well" under the projections $\pi_{m,n}$. Hence we are justified in making the remark in Example 2.1 that it suffices to verify the identity (which homogeneous of degree 3) for $n = 3$, and conclude that it holds for each $n \geq 0$. More generally, if we have an identity involving elementary symmetric polynomials which is homogeneous of degree n and we wish to show that the identity holds in Λ_m for each $m \geq 0$, it suffices to show that the identity holds in Λ_n. The elementary symmetric polynomials are not unique in this respect; the same principle holds for the homogeneous symmetric polynomials by virtue of the automorphism ω, as well as for the Schur polynomials of Section 3.

We conclude this section with a brief discussion of bases for the ring of symmetric functions. As we remarked in Section 1.1, integer partitions are useful for parametrizing important families of symmetric functions. We have seen that this is the case for the monomial symmetric functions m_λ, which form a $\mathbb{Z}$-basis of Λ. In fact, many of the important $\mathbb{Z}$-bases of Λ can be paramatrized by partitions. In particular, this can be accomplished with the elementary and homogeneous symmetric functions, given a suitable construction.

As we have remarked, the elementary and homogeneous symmetric polynomials behave well with respect to the homomorphisms $\pi_{m,n}$ in the sense that $\pi_{n,m}(e_i^{(n)}) = e_i^{(m)}$ and $\pi_{n,m}(h_i^{(n)}) = h_i^{(m)}$ for all $m, n \geq 0$. Clearly we may just as well speak of elementary and homogeneous symmetric functions e_i and h_i in Λ by applying the appropriate inverse projections $\pi_i : \Lambda_i \to \Lambda$. Moreover, the symmetric functions we obtain by these homomorphisms agree completely with the original generating functions (2.1.1) and (2.6.1).

Given a partition $\lambda = (\lambda_1, \ldots, \lambda_n)$, we define symmetric functions $e_\lambda := e_{\lambda_1} \cdots e_{\lambda_n}$ and similarly $h_\lambda := h_{\lambda_1} \cdots h_{\lambda_n}$. Each e_λ can be written uniquely as a $\mathbb{Z}$-linear combination of monomial symmetric functions m_μ; a proof of this fact can be found in [6, p. 20]. Since the m_μ are a $\mathbb{Z}$-basis for Λ essentially by construction, the same is consequently true of the e_λ. Now recall the map $\omega : \Lambda_n \to \Lambda_n$ used in the proof of Theorem 2.6, which was defined by setting $\omega(e_k) = h_k$ for each $k \leq n$. We may just as well define the homomorphism $\omega : \Lambda \to \Lambda$ on symmetric functions by setting $\omega(e_k) = h_k$ for each $k \geq 0$. This is again an involution (and hence automorphism) by essentially the same proof as in the finite variable case. Since $\omega(e_\lambda) = h_\lambda$, then the h_λ also form a $\mathbb{Z}$-basis of Λ. This gives the following generalization of the Fundamental Theorem of Symmetric Polynomials and Theorem 2.6.

Proposition 2.9. *The set of symmetric functions* e_λ *(resp.* h_λ*), as* λ *runs over all partitions, forms a* $\mathbb{Z}$*-basis for the space* Λ *of symmetric functions. That is, any symmetric function can be expressed as a unique polynomial in the* e_λ *(resp.* h_λ*) with coefficients in* $\mathbb{Z}$*.*

Remark. Just as symmetric polynomials can be defined over any ring of coefficients, we may carry out the construction of the ring of symmetric functions over any (commutative) ring of coefficients. If we work in the ring $\Lambda_{\mathbb{Q}}$ of symmetric functions with rational coefficients, then we may recover Newton's identities from Section 2.2. If we define $p_\lambda := p_{\lambda_1} \cdots p_{\lambda_n}$, then Newton's identities together with Proposition 2.9 imply that the p_λ form a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}}$. See [6, Ch. I, §2] for a more comprehensive discussion.

3 Schur Polynomials

The purpose of the following sections is to introduce another classical symmetric polynomial basis, namely the Schur polynomials. The Schur polynomials in fact form an orthonormal basis with respect to a natural inner product, and admit many interesting combinatorial properties; these properties will be explored further in Section 5. In Sections 3.1, 3.2, and 3.4, we again follow Procesi's treatment of the topic; the relevant material can be found in [7, p. 28-33]. Section 3.3, which explores some relationships between Schur polynomials and some of the bases encountered in Section 2, is based primarily on results from [6, p. 41-42].

3.1 Alternating Polynomials

In section 2.3 we introduced the notion of an *alternating* (or *antisymmetric*) polynomial, which is a polynomial in n variables with the property that, for any permutation $\sigma \in S_n$, we have

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \operatorname{sgn}(\sigma) f(x_1, \ldots, x_n)$$

We have already seen that the Vandermonde determinant $V(x_1, \ldots, x_n) := \prod_{i<j}(x_i - x_j)$ is an example of an alternating polynomial (cf. Section 2.3). In fact, the following proposition indicates that the Vandermonde determinant is the prototypical alternating polynomial.

Proposition 3.1. *A polynomial $f(x_1, \ldots, x_n)$ is alternating if and only if*

$$f(x_1, \ldots, x_n) = V(x_1, \ldots, x_n) g(x_1, \ldots, x_n)$$

for some symmetric polynomial $g(x_1, \ldots, x_n)$.

Proof. ($\Leftarrow$) This follows from the easily verifiable general fact that if p_1 and p_2 are polynomials in n variables with p_1 alternating and p_2 symmetric, then $p_1 p_2$ is alternating.

($\Rightarrow$) Let $f(x_1, \ldots, x_n)$ be alternating. Consider two indices $i < j$, and substitute x_j for x_i to obtain a polynomial $f[x_j/x_i]$. Evidently, we will obtain the same polynomial if we apply the transposition (ij) before making the substitution. That is, $f[x_j/x_i] = ((ij) \cdot f)[x_j/x_i]$. Since f is alternating, we therefore have

$$f[x_j/x_i] = ((ij) \cdot f)[x_j/x_i] = (-f)[x_j/x_i] = -f[x_j/x_i]$$

Consequently, $f[x_j/x_i] = 0$. Since $f[x_j/x_i]$ is precisely f with x_i evaluated at x_j, we see that $x_i - x_j \mid f$. Since i and j were arbitrary, we see that $x_i - x_j \mid f$ for each $i < j$. That is, $V \mid f$. We may therefore write $f(x_1, \ldots, x_n) = V(x_1, \ldots, x_n) g(x_1, \ldots, x_n)$. If $\sigma \in S_n$, then

$$\operatorname{sgn}(\sigma) V g = \operatorname{sgn}(\sigma) f = \sigma \cdot f = \sigma \cdot (V g) = (\sigma \cdot V)(\sigma \cdot g) = (\operatorname{sgn}(\sigma) V)(\sigma \cdot g)$$

Dividing both sides by $\operatorname{sgn}(\sigma)V$ gives $\sigma \cdot g = g$, so g is symmetric. $\square$

Let A denote the $\mathbb{Z}$-algebra of alternating polynomials, S the $\mathbb{Z}$-algebra of symmetric polynomials. Then we may express Proposition 3.1 more formally in the following manner.

Proposition 3.2. *The algebra A of alternating polynomials is a free rank 1 module over the algebra S of symmetric polynomials generated by $V(x)$. That is, $A = V(x)S$.*

Consequently, if we have a basis for A over $\mathbb{Z}$, then we may obtain a basis for S over $\mathbb{Z}$ by dividing each basis element by $V(x)$. It is in this way that we will construct the *Schur polynomials* in Section 3.2. As such, we will begin here by constructing an integral basis for the space of antisymmetric polynomials.

Consider the polynomial ring $\mathbb{Z}[x_1, \ldots, x_n]$ with the basis given by the monomials, and consider the usual action of S_n by permutation of the variables. Recall from section 2.5 that the orbit of a given monomial corresponds to a partition λ of its degree. In particular, the weakly decreasing sequence of nonnegative integers given by $m_1 \geq m_2 \geq \cdots \geq m_n \geq 0$ corresponds to the orbit of the monomial $x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}$.

Let f be an alternating polynomial, (ij) a transposition. Since f is antisymmetric, then we have $(ij)f = -f$. However, applying the transposition (ij) fixes all monomials for which x_i and x_j have the same exponent. Consequently, each monomial with nonzero coefficient in f must have each exponent distinct. Moreover, if such a mononial lies in the orbit of $x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}$ with $m_1 > m_2 > \cdots > m_n \geq 0$, we have

$$x_{\sigma(1)}^{m_1} x_{\sigma(2)}^{m_2} \cdots x_{\sigma(n)}^{m_n} = \operatorname{sgn}(\sigma) x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}$$

for any $\sigma \in S_n$. From this we conclude the following theorem.

Theorem 3.3. *The polynomials*

$$A_{m_1 > m_2 > \cdots > m_n \geq 0}(x_1, \ldots, x_n) := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) x_{\sigma(1)}^{m_1} x_{\sigma(2)}^{m_2} \cdots x_{\sigma(n)}^{m_n} \tag{3.1.1}$$

form an integral basis for the space of alternating polynomials.

When performing computations with alternating polynomials, it is often convenient to reduce modulo S_n-orbits. More precisely, consider the space SM spanned by the *standard monomials*, which are monomials of the form $x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ with $k_1 > k_2 > \cdots > k_n \geq 0$. We then have a linear map L from the space of polynomials to SM which is the identity on SM and sends all nonstandard monomials to 0. Then $L(A_{m_1 > m_2 > \cdots > m_n \geq 0}) = x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$, and so L is a linear isomorphism between the space of alternating polynomials and SM. In particular, L maps the basis elements in (3.1.1) to the standard monomials.

3.2 Schur Polynomials

The basis of the antisymmetric polynomials defined in Theorem 3.3, together with Proposition 3.1, allows us to construct a new integral basis for the symmetric functions. Before we do so, we first introduce a useful result related to integer partitions.

Remark. The map

$$\begin{aligned}\lambda := (p_1, p_2, \ldots, p_{n-1}, p_n) &\mapsto (p_1 + n - 1, p_2 + n - 2, \ldots, p_{n-1} + 1, p_n)\\ &:= \lambda + \delta\end{aligned}$$

is a bijection between the sets of decreasing and strictly decreasing finite sequences. Hence we may unambiguously use the notation $A_{\lambda+\delta}$ to refer to the basis antisymmetric polynomials in (3.1.1).

The following proposition, which follows from the Leibniz determinant formula

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)}$$

gives a convenient identity which provides an alternate way of expressing the basis antisymmetric polynomials.

Proposition 3.4. *If $\lambda := (\lambda_1, \ldots, \lambda_n)$ is an integer partition, then*

$$A_{\lambda+\delta} = \det(\mathcal{A}_{\lambda+\delta})$$

where the $n \times n$ matrix $\mathcal{A}_{\lambda+\delta}$ is defined by $(\mathcal{A}_{\lambda+\delta})_{i,j} := x_j^{\lambda_i+n-i}$.

Remark. In the notation of the preceding proposition, we have:

$$V(x_1, \ldots, x_n) = A_{\delta^{(n)}} = \det(\mathcal{A}_{0+\delta^{(n)}})$$

Definition 3.1. Let λ be an integer partition with $l(\lambda) \leq n$. Then we define the *Schur polynomial* associated to λ by

$$S_\lambda(x_1, \ldots, x_n) := \frac{A_{\lambda+\delta^{(n)}}(x_1, \ldots, x_n)}{V(x_1, \ldots, x_n)} \tag{3.2.1}$$

Using the notation of Proposition 3.4, we may equivalently express Schur polynomials as a quotient of determinants:

$$S_\lambda := \frac{\det(\mathcal{A}_{\lambda+\delta})}{\det(\mathcal{A}_\delta)} \tag{3.2.2}$$

Example. As per the remark following Proposition 3.4, we have $V = A_\delta$ in any number of variables, and hence $S_0 = 1$.

If $n = 2$, and $\lambda = (3, 2)$, then $\mathcal{A}_{\lambda+\delta} = \left(\begin{smallmatrix} x_1^4 & x_2^4 \\ x_1^2 & x_2^2 \end{smallmatrix}\right)$. Computing the determinant, we obtain $A_{\lambda+\delta} = x_1^4 x_2^2 - x_1^2 x_2^4$. Dividing by the Vandermonde determinant gives $S_\lambda = x_1^3 x_2^2 + x_1^2 x_2^3$.

By Proposition 3.1, we see that each Schur function is in fact a symmetric polynomial. The same proposition, together with Theorem 3.3, moreover implies the following.

Theorem 3.5. *The polynomials S_λ, with $\lambda \vdash m$ and $l(\lambda) \leq n$, are an integral basis of the space of symmetric polynomials homogeneous of degree m in n variables.*

The preceding theorem in fact generalizes to symmetric functions in infinitely many variables. In order to work towards this result, summarized in Theorem 3.8, we must first develop some basic algebraic properties of the Schur polynomials. We begin with the following proposition.

Proposition 3.6. *Let $\lambda := (p_1, \ldots, p_n)$ be an integer partition, a a positive integer. Then*

$$A_{\lambda+\delta+(a^n)} = (x_1 \cdots x_n)^a A_{\lambda+\delta}, \quad S_{\lambda+(a^n)} = (x_1 \cdots x_n)^a S_\lambda \tag{3.2.3}$$

where (a^n), in multiplicity notation, is the partition of length n with each entry equal to a.

Proof. By Proposition 3.4, we have $A_{\lambda+\delta+(a^n)} = \det(\mathcal{A}_{\lambda+\delta+(a^n)})$. By definition of this matrix, we have

$$(\mathcal{A}_{\lambda+\delta+(a^n)})_{i,j} = x_j^{\lambda_i+a+n-i} = x_j^a x_j^{\lambda_i+n-i} = x_j^a (\mathcal{A}_{\lambda+\delta})_{i,j}$$

By multilinearity of the determinant, we factor out x_j^a from each column to obtain

$$A_{\lambda+\delta+(a^n)} = \det(\mathcal{A}_{\lambda+\delta+(a^n)}) = \left(\prod_{j=1}^n x_j^a\right) \det(\mathcal{A}_{\lambda+\delta}) = (x_1 \cdots x_n)^a A_{\lambda+\delta}$$

The analogous identity for Schur polynomials follows by dividing by the Vandermonde determinant. □

We now consider the effect of projecting Schur functions in n variables into the space of Schur functions in $n-1$ variables. This behaviour is summarized in the following lemma.

Lemma 3.7. *Consider the set of Schur polynomials in n variables $x_1, \ldots, x_n$. Under the map $\pi_{n,n-1}$ which evaluates x_n at 0, S_λ vanishes when $l(\lambda) = n$. Otherwise, S_λ maps to the corresponding Schur polynomial S_λ in $n-1$ variables.*

Proof. Fix a number of variables n and let $\lambda := (h_1, h_2, \ldots, h_n)$ with $h_1 \geq h_2 \geq \ldots \geq h_n \geq 0$. If $l(\lambda) = n$, then $h_n > 0$ and so by 3.6 we have

$$S_\lambda(x_1, \ldots, x_n) = (x_1 \cdots x_n) S_{\lambda-(1^n)}(x_1, \ldots, x_n)$$

Clearly, $\pi_{n,n-1}(S_\lambda) = 0$ in this case.

Suppose instead that $l(\lambda) < n$ (and so $h_n = 0$), and refer to the partition $(h_1, h_2, \ldots, h_{n-1})$ by the same symbol λ. Consider first the evaluation of the Vandermonde determinant at $x_n = 0$. We have

$$V(x_1, \ldots, x_{n-1}, 0) = \left(\prod_{i=1}^{n-1} x_i\right) \left(\prod_{i<j\leq n-1} (x_i - x_j)\right) = \left(\prod_{i=1}^{n-1} x_i\right) V(x_1, \ldots, x_{n-1})$$

Now consider the alternating polynomial $A_{\lambda+\delta}(x_1, \ldots, x_n)$. Denote $l_i := h_i + n - i$, so that $l_n = 0$ and

$$A_{\lambda+\delta}(x_1, \ldots, x_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) x_1^{l_{\sigma(1)}} \cdots x_n^{l_{\sigma(n)}}$$

Setting $x_n = 0$, we restrict the above sum to only those terms for which $\sigma(n) = n$ (namely, those terms for which $l_{\sigma(n)} = 0$). Since the subgroup of all such permutations is isomorphic to S_{n-1}, we have

$$A_{\lambda+\delta}(x_1, \ldots, x_{n-1}, 0) = \sum_{\sigma \in S_{n-1}} \operatorname{sgn}(\sigma) x_1^{l_{\sigma(1)}} \cdots x_{n-1}^{l_{\sigma(n-1)}} = A_{\lambda+(1^{n-1})+\delta}(x_1, \ldots, x_{n-1})$$

The rightmost identity follows since $l_i = h_i + n - i = (h_i + 1) + (n-1) - i$. Combining these results, we have

$$\begin{aligned} V(x_1, \ldots, x_{n-1}, 0) S_\lambda(x_1, \ldots, x_{n-1}, 0) &= A_{\lambda+\delta}(x_1, \ldots, x_{n-1}, 0) \\ &= A_{\lambda+(1^{n-1})+\delta}(x_1, \ldots, x_{n-1}) \\ &= \left(\prod_{i=1}^{n-1} x_i \right) A_{\lambda+\delta}(x_1, \ldots, x_{n-1}) \\ &= \left(\prod_{i=1}^{n-1} x_i \right) V(x_1, \ldots, x_{n-1}) S_\lambda(x_1, \ldots, x_{n-1}) \\ &= V(x_1, \ldots, x_{n-1}, 0) S_\lambda(x_1, \ldots, x_{n-1}) \end{aligned}$$

Dividing through by $V(x_1, \ldots, x_{n-1}, 0)$ implies that

$$\pi_{n,n-1}(S_\lambda(x_1, \ldots, x_n)) := S_\lambda(x_1, \ldots, x_{n-1}, 0) = S_\lambda(x_1, \ldots, x_{n-1})$$

whenever $l(\lambda) < n$, which completes the proof. □

Remark. Lemma 3.7 is particularly useful in light of the discussion surrounding Lemma 2.8 in section 2.7. In particular, we have shown that the Schur polynomials "behave well" with respect to the homomorphisms $\pi_{m,n}$ in the same way that the elementary symmetric polynomials do. That is, for any $m, n \geq 1$, we have $\pi_{m,n}(S_\lambda(x_1, \ldots, x_m)) = S_\lambda(x_1, \ldots, x_n)$.

Consequently, we have the same important conclusion regarding Schur polynomials as we did for the elementary symmetric polynomials. That is, if we wish to show an identity involving Schur polynomials which is homogeneous of degree n holds in any number of variables, it suffices to verify the identity holds in Λ_n.

We again make recourse to the homomorphisms π discussed in section 2.7 in order to prove the following generalization of Theorem 3.5.

Theorem 3.8. *The ring Λ of symmetric functions in infinitely many variables has as an integral basis all Schur functions S_λ. Restriction to symmetric functions in m variables amounts to setting all S_λ with $l(\lambda) > m$ to 0.*

Proof. Consider a fixed degree k, and let Λ_n^k denote the space of symmetric polynomials of degree k in n variables, as in section 2.7. By Theorem 3.5, an integral basis for this space is given by the

Schur polynomials S_λ in n variables with $\lambda \vdash k$ and $l(\lambda) \leq n$. Consequently, a $\mathbb{Z}$-basis for the space $\bigoplus_{k=0}^{n} \Lambda_n^k$ is given by all Schur polynomials S_λ with $|\lambda| \leq n$ and $l(\lambda) \leq n$. The latter condition is superfluous in this case, so the basis is precisely the set $B_n := \{S_\lambda \mid |\lambda| \leq n\}$.

In the proof of Lemma 2.7, we showed that the space $\bigoplus_{k=0}^{n} \Lambda_n^k$ is the isomorphic image of π_n applied to the space $\bigoplus_{k=0}^{n} \Lambda^k$. That is, a $\mathbb{Z}$-basis for the space $\bigoplus_{k=0}^{n} \Lambda^k$ of symmetric functions of degree at most n is given by applying π_n^{-1} to the basis B_n. In particular, the resulting basis is the set of Schur functions S_λ in infinitely many variables such that $|\lambda| \leq n$. Extending to the entire graded ring $\Lambda = \bigoplus_{k \geq 0} \Lambda^k$, we see that a $\mathbb{Z}$-basis for the space Λ of symmetric functions is given by all Schur functions S_λ.

Restriction to m variables amounts to applying π_m to this basis. The map π_m sends all Schur functions S_λ with $l(\lambda) \leq m$ to the corresponding m-variable Schur polynomial $S_\lambda(x_1, \ldots, x_m)$, and all Schur functions with $l(\lambda) > m$ to 0. □

3.3 Schur Polynomials and Other Symmetric Polynomial Bases

In this section we present some results relating the Schur functions to other families of symmetric functions we have introduced earlier. The first of these results establishes that the elementary symmetric functions are in fact a subset of the Schur functions, namely those corresponding to partitions containing a single column.

Proposition 3.9. *In the ring of symmetric functions, we have $e_k = S_{(1^k)}$ for each $k \geq 0$. For any $k, n \geq 0$, we have the symmetric polynomial identity*

$$e_k(x_1, \ldots, x_n) = S_{(1^k)}(x_1, \ldots, x_n) \tag{3.3.1}$$

Proof. Since e_k is homogeneous of degree k, then by Lemma 2.8 and the remark following Lemma 3.7, it suffices to prove the identity in Λ_k. In this case, we have

$$e_k(x_1, \ldots, x_k) = \prod_{i=1}^{k} x_i = S_0(x_1, \ldots, x_k) \prod_{i=1}^{k} x_i = S_{(1^k)}(x_1, \ldots, x_k)$$

which establishes the identity in the ring of symmetric functions, as well as the polynomial identity for all $k, n \geq 0$. □

Dually, the complete homogeneous symmetric polynomials can also be realized as a special case of the Schur polynomials, namely those corresponding to the conjugate partitions $(k) = (1^k)'$ containing a single row. Before we can prove this, we will require several intermediary results. Due to the Fundamental Theorem 2.2 and the corresponding Theorem 2.6 for the homogeneous symmetric polynomials, any given Schur polynomial can be expressed uniquely as a polynomial in terms of the elementary (or homogeneous) symmetric polynomials. In fact, these expressions are given by in a compact manner by the following *Jacobi-Trudi* identities.

Proposition 3.10 (Jacobi-Trudi Identities)**.** *Let λ be a partition such that $l(\lambda) \leq n$ and such that $\lambda_1 = l(\lambda') \leq m$. Then the following identity holds in the ring of symmetric functions.*

$$S_\lambda = \det(h_{\lambda_i - i + j})_{1 \leq i,j \leq n} = \det(e_{\lambda'_i - i + j})_{1 \leq i,j \leq m} \tag{3.3.2}$$

with the convention that $e_k = h_k = 0$ for $k < 0$.

Proof. A proof that the two determinants in (3.3.2) are equal can be found in [6, p. 22-23]. It therefore suffices to prove the leftmost equality that that $S_\lambda = \det(h_{\lambda_i - i + j})$. To do so, we will prove the identity in Λ_n. For $1 \leq k \leq n$, denote by $e_r^{(\neg k)}$ the elementary symmetric polynomial e_r in $n-1$ variables under the evaluation

$$e_r^{(\neg k)} := e_r(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$$

Let M denote the $n \times n$ matrix

$$M := \left((-1)^{n-i} e_{n-i}^{(\neg k)}\right)_{1 \leq i,k \leq n}$$

For any n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, define the $n \times n$ matrices $\mathcal{A}_\alpha$ and H_α by

$$\mathcal{A}_\alpha := \left(x_j^{\alpha_i}\right)_{1 \leq i,j \leq n} \quad \text{and} \quad H_\alpha := (h_{\alpha_i - n + j})_{1 \leq i,j \leq n}$$

We claim that $\mathcal{A}_\alpha = H_\alpha M$. Indeed, let $E^{(\neg k)}(t) := \sum_{r=0}^{n-1} e_r^{(\neg k)} t^r = \prod_{i \neq k}(1 + x_i t)$. Then if $H(t)$ is the generating function for the homogeneous symmetric functions as in (2.6.1), we have

$$H(t) E^{(\neg k)}(-t) = (1 - x_k t)^{-1} = \sum_{r=0}^{\infty} x_k^r t^r$$

Comparing the α_i-th coefficient of the leftmost and rightmost sides gives the identity

$$\sum_{j=1}^{n} h_{\alpha_i - n + j} (-1)^{n-j} e_{n-j}^{(\neg k)} = x_k^{\alpha_i}$$

The lefthand side is precisely the (i,k)-th entry of $H_\alpha M$, and the righthand side is the (i,k)-th entry of $\mathcal{A}_\alpha$, establishing the claim. By taking determinants, we obtain the following identity valid for any $\alpha \in \mathbb{N}^n$

$$A_\alpha := \det(\mathcal{A}_\alpha) = \det(H_\alpha) \det(M)$$

If we let $\alpha = \delta^{(n)}$, we see that the lefthand side becomes the Vandermonde determinant $A_\delta = V(x)$. The matrix H_α becomes upper triangular with 1's along the diagonal, so we have $\det(H_\delta) = 1$ and hence $\det(M) = V(x)$. We therefore have

$$A_\alpha = V(x) \det(H_\alpha)$$

for any $\alpha \in \mathbb{N}^n$. By letting $\alpha = \lambda + \delta$ and dividing both sides by the Vandermonde determinant $V(x)$, the lefthand side becomes S_λ and the righthand side becomes $\det(h_{\lambda_i - i + j})_{1 \leq i,j \leq n}$, which completes the proof. $\square$

Recall the involution $\omega : \Lambda \to \Lambda$ defined by setting $\omega(e_k) = h_k$ for each $k \geq 0$, which is an automorphism of Λ. By applying the map ω to the various expressions in (3.3.2), we deduce that

$$\omega(S_\lambda) = S_{\lambda'} \tag{3.3.3}$$

for any partition λ. In particular, this implies the following corollary, which follows from Proposition 3.9 at the beginning of this section.

Proposition 3.11. *In the ring of symmetric functions, we have $h_k = S_{(k)}$ for each $k \geq 0$. For any $k, n \geq 0$, we have the symmetric polynomial identity*

$$h_k(x_1, \dots, x_n) = S_{(k)}(x_1, \dots, x_n) \tag{3.3.4}$$

3.4 Cauchy Formulas

The Schur functions defined in the preceding sections admit some rather interesting identities in terms of generating functions. Remarkably, as we shall see in Section 5, these identities turn out to provide the foundation for a natural inner product on the ring of symmetric functions.

Proposition 3.12. *Let $\{x_i \mid i \in \mathbb{N}\}, \{y_j \mid j \in \mathbb{N}\}$ be sets of variables. Then*

$$\prod_{i,j=1}^{n} \frac{1}{1 - x_i y_j} = \sum_\lambda S_\lambda(x) S_\lambda(y) \tag{3.4.1}$$

where the righthand sum in (3.4.1) is taken over all partitions. If λ' denotes the conjugate partition of λ, in which columns and rows are exchanged, then for any m, n we have

$$\prod_{i=1,j=1}^{m,n} (1 + x_i y_j) = \sum_\lambda S_\lambda(x) S_{\lambda'}(y) \tag{3.4.2}$$

Proof. We will first prove (3.4.1). Consider the matrix A given by $A_{ij} := \frac{1}{1-x_i y_j}$. Our goal is to provide two expressions for the determinant of A, the equality of which will imply the desired identity. We will first show that

$$\det(A) = \frac{V(x)V(y)}{\prod_{i,j=1}^{n}(1 - x_i y_j)} \tag{3.4.3}$$

Beginning with the matrix A, subtract the first row from the i-th row for each $i > 1$. Doing so, we obtain a new matrix B with the same determinant. The entries of B are given by $B_{1j} = A_{1j}$ and, for each $i > 1$:

$$B_{ij} = \frac{1}{1 - x_i y_j} - \frac{1}{1 - x_1 y_j} = \frac{(x_i - x_1) y_j}{(1 - x_i y_j)(1 - x_1 y_j)}$$

Thus we may factor $x_i - x_1$ from the i-th row for each $i > 1$, and we may factor $\frac{1}{1-x_1y_j}$ from the j-th column. Hence

$$\det(B) = \left(\frac{1}{1-x_1y_1}\prod_{i=2}^{n}\frac{x_i-x_1}{1-x_1y_i}\right)\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \frac{y_1}{1-x_2y_1} & \frac{y_2}{1-x_2y_2} & \cdots & \frac{y_n}{1-x_2y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{y_1}{1-x_ny_1} & \frac{y_2}{1-x_ny_2} & \cdots & \frac{y_n}{1-x_ny_n} \end{pmatrix} \tag{3.4.4}$$

In the righthand matrix, we subtract the first column from the j-th for each $j > 1$. This yields a new matrix C with the same determinant as the righthand matrix above. The entries of C are given by $C_{11} = 1$, $C_{i1} = \frac{y_1}{1-x_iy_1}$ for each $i > 1$, $C_{1j} = 0$ for $j > 1$, and for each $i, j > 1$:

$$C_{ij} = \frac{y_j - y_1}{(1-x_iy_j)(1-x_iy_1)}$$

For each $i, j > 1$, we may extract a factor $\frac{1}{1-x_iy_1}$ from the i-th row and a factor $(y_j - y_1)$ from the j-th column. This alters the determinant by a factor of $\prod_{i=2}^{n}\frac{y_i-y_1}{1-x_iy_1}$. After doing so, we are left with

$$\det(A) = \left(\frac{1}{1-x_1y_1}\prod_{i=2}^{n}\frac{x_i-x_1}{1-x_1y_i}\prod_{i=2}^{n}\frac{y_i-y_1}{1-x_iy_1}\right)\det\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \frac{y_1}{1-x_2y_1} & \frac{1}{1-x_2y_2} & \cdots & \frac{1}{1-x_2y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{y_1}{1-x_ny_1} & \frac{1}{1-x_ny_2} & \cdots & \frac{1}{1-x_ny_n} \end{pmatrix}$$

Note that the $(1,1)$ submatrix of the rightmost matrix is precisely a square matrix of size $n-1$ with the same form as the original matrix A, albeit with the variables x_1 and y_1 removed. By expanding the above determinant along the first row, we see that (3.4.3) follows by induction.

We will now find an alternate expression for the determinant of A by considering each entry of A as a geometric series $\frac{1}{1-x_iy_j} = \sum_{k=0}^{\infty} x_i^k y_j^k$. In this way, we may realize each row (resp. column) of A as the formal sum of infinitely many rows (resp. columns). By multilinearity in the rows, we obtain

$$\det(A) = \sum_{k_1=0}^{\infty}\cdots\sum_{k_n=0}^{\infty}\det(A_{k_1,k_2,\ldots,k_n}), \quad \text{where } (A_{k_1,k_2,\ldots,k_n})_{ij} := (x_iy_j)^{k_i} \tag{3.4.5}$$

By factoring out $x_i^{k_i}$ from each row, we see that $\det(A_{k_1,k_2,\ldots,k_n}) = \left(\prod_i x_i^{k_i}\right)\det(B_{k_1,k_2,\ldots,k_n})$. Here we define $B_{k_1,k_2,\ldots,k_n}$ to be the matrix with entries $(B_{k_1,k_2,\ldots,k_n})_{ij} := y_j^{k_i}$. This determinant is 0 if the k_i are not all distinct. Otherwise, we may reorder the sequence of k_i to be strictly decreasing, introducing the sign of the permutation as we do so. We then collect all terms in which the k_i are a permutation of a given sequence $\lambda + \delta$. Doing so produces the term $A_{\lambda+\delta}(x)A_{\lambda+\delta}(y)$. Repeating this procedure for all partitions λ, together with (3.4.3), implies the following identity:

$$\frac{V(x)V(y)}{\prod_{i,j=1}^{n}(1-x_iy_j)} = \sum_{\lambda} A_{\lambda+\delta}(x)A_{\lambda+\delta}(y) \tag{3.4.6}$$

Dividing both sides by the Vandermonde determinants turns the alternating polynomials into the corresponding Schur polynomials, which implies (3.4.1).

The dual identity (3.4.2) follows from (3.4.1) by applying the automorphism ω. By the remark following the end of the proof, the identity (3.4.1) may just as well be written as

$$\prod_{i=1,j=1}^{m,n}(1-x_iy_j)^{-1}=\sum_{\lambda}S_\lambda(x)S_\lambda(y)$$

By applying ω to the symmetric functions in y, the lefthand side of the identity becomes

$$\omega\left(\prod_{i=1,j=1}^{m,n}(1-x_iy_j)^{-1}\right)=\omega\left(\prod_{j=1}^{n}H(y_j)\right)=\prod_{j=1}^{n}E(y_j)=\prod_{i=1,j=1}^{m,n}(1+x_iy_j)$$

which is the lefthand side of (3.4.2). By (3.3.3) we have seen that $\omega(S_\lambda)=S_{\lambda'}$, which turns the righthand side of (3.4.1) into the righthand side of (3.4.2). □

Remark. A variant of the identity (3.4.1) still holds if we have sets of variables $x_1,\dots,x_m$ and $y_1,\dots,y_n$ with $m>n$. Indeed, we may obtain $\prod_{i=1}^{m}\prod_{j=1}^{n}\frac{1}{1-x_iy_j}$ from $\prod_{i,j=1}^{n}\frac{1}{1-x_iy_j}$ by setting $y_j=0$ for all j with $n<j\le m$. Lemma 3.7 accounts for the corresponding change to the righthand sum, yielding

$$\prod_{i=1}^{m}\prod_{j=1}^{n}\frac{1}{1-x_iy_j}=\sum_{\lambda\vdash m,l(\lambda)\le n}S_\lambda(x_1,\dots,x_m)S_\lambda(y_1,\dots,y_n)\tag{3.4.1a}$$

If $m<n$, then by a symmetric argument we also obtain

$$\prod_{i=1}^{m}\prod_{j=1}^{n}\frac{1}{1-x_iy_j}=\sum_{\lambda\vdash n,l(\lambda)\le m}S_\lambda(x_1,\dots,x_m)S_\lambda(y_1,\dots,y_n)\tag{3.4.1b}$$

Of course, both of the above identities amount to taking the given product on the lefthand side and taking the righthand sum over all partitions λ. The way that we have written the identities above simply clarifies which terms in the sum are nonzero.

4 Combinatorics of Young Tableaux

In this section we move away from the algebraic theory of symmetric functions and focus instead on the combinatorial theory of Young tableaux. We will primarily follow the presentation of this subject offered by Fulton in Part I of [2]. We will refer to the same source for terminology and notation, most of which we have summarized in Section 4.1. Sections 4.2-4.4 correspond roughly to §1 and §2.1 of the same source, although the order of the material has been reworked in order to develop the theory more directly. Section 4.5 on the Robinson-Schensted-Knuth correspondence is adapted from [2, §4.1].

4.1 Young Tableaux

In Section 1.1 we introduced the notion of a Young diagram, which provided a means to reason geometrically with integer partitions. The choice of using boxes (as opposed to, say, dots) to populate the rows of the diagram is not entirely superfluous, as it enables us to fill the boxes. From a combinatorial perspective, it is natural to consider the number of ways that the boxes of a Young diagram can be filled, subject to certain constraints. Although this seems unrelated to our current discussion, the combinatorics of filling Young diagrams with positive integers turns out to have deep connections to the study of symmetric functions, as we shall see beginning in Section 5.3. The primary combinatorial object we will be concerned with in the following sections are *tableaux.* We provide in this section the following necessary definitions.

Definition 4.1. Any way of putting a positive integer in each box of a Young diagram is called a *filling* of the diagram. A filling in which each box contains a distinct entry is called a *numbering* of the diagram.

Definition 4.2. A *semistandard Young tableau*, or simply *tableau*, is a filling of a Young diagram such that the entries are

1. Weakly increasing across each row, and

2. Strictly increasing down each column.

We say that the tableau is a tableau *on* the diagram λ, or that λ is the *shape* of the tableau.

Definition 4.3. A *standard* tableau is a tableau in which the entries are the integers 1 to $n = |\lambda|$, each occurring once.

Example. The following are a tableau and standard tableau, respectively, on $\lambda = (5, 3, 2, 2)$:

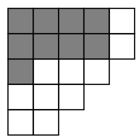
$$T_0 = \begin{array}{ccccc} 1 & 2 & 2 & 3 & 4 \\ 2 & 4 & 5 & & \\ 4 & 5 & & & \\ 5 & 6 & & & \end{array} \qquad \text{and} \qquad T_1 = \begin{array}{ccccc} 1 & 2 & 4 & 10 & 12 \\ 3 & 5 & 7 & & \\ 6 & 8 & & & \\ 9 & 11 & & & \end{array}$$

Remark. The entries of tableaux need not be positive integers. More generally, fillings of Young diagrams can use entries from any *alphabet* (*i.e.* totally ordered set).

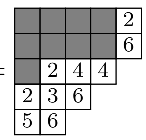
Recall that a skew diagram is the diagram obtained by removing a smaller Young diagram from a larger one that contains it. We may just as well consider fillings of skew diagrams, which prompts the following analogous definition.

Definition 4.4. A *skew tableau* is a filling of the boxes of a skew diagram with positive integers such that the entries weakly increase across each row and strictly increase down each column. We call the skew diagram of a skew tableau its *shape*.

Example. Let $\lambda = (5, 5, 4, 3, 2)$ and $\mu = (4, 4, 1)$. The the following are the skew diagram of λ/μ and a skew tableau S on λ/μ, respectively:



and $S =$



Remark. When working with tableaux or skew tableaux in the sections that follow, we must often compare which of two entries is lesser and which is greater. As a general convention, we consider the rightmost of two equal entries to be the greater of the two. For example, in the skew tableau S shown above, we would consider the 2 situated in the (4,1) box to be less than the 2 situated in the (3,2) box, which is in turn less than the 2 situated in the (1,5) box. Each of these 2's is still considered less than the 3 in the (4,2) box, of course, regardless of their position in the diagram.

Definition 4.5. Let λ, μ be Young diagrams such that $\mu \subset \lambda$ and consider the skew diagram λ/μ. A box (in λ/μ) such that the two boxes immediately below and immediately to the right are not contained in the diagram is called an *outside corner*. A box (in μ) such that the boxes immediately below and immedately to the right are not contained in μ is called an *inside corner*.

For instance, if λ and μ are as in the previous example, then the skew diagram λ/μ has two inside boxes, namely those in the (3,1) and (2,4) positions. There are four outside boxes in λ/μ, namely those in the (5,2), (4,3), (3,4), and (2,5) positions.

In the sections that follow, we will investigate various operations on tableaux and skew tableaux. When proving results related to these operations, it is helpful to eschew the geometric view of diagrams and instead consider how these operations act on lists of the entries of tableaux. In particular, we will be interested in examining how these operations act on the *row words* of skew tableaux, which we define as follows.

Definition 4.6. A *word* w is a finite sequence of letters taken from some alphabet, typically the positive integers. Given two words w, w', we denote their juxtaposition by $w \cdot w'$, or more often ww'.

Definition 4.7. Given a (possibly skew) tableau T with n rows, we define the *word* (or *row word*) of T, denoted $w(T)$ or $w_{row}(T)$, to be

$$w(T) = w_n w_{n-1} \cdots w_2 w_1$$

where w_i is the word consisting of the entries of the i-th row of T, written in increasing order. In other words, $w(T)$ is the word obtained by writing the entries of T from left to right within each row, starting with the bottom row and working up to the top row.

Example. Consider the tableaux T_0 and T_1, and the skew tableau S as in the previous examples of this section. We have

$$\begin{aligned} w(T_0) &= 5\ 6 \mid 4\ 5 \mid 2\ 4\ 5 \mid 1\ 2\ 2\ 3\ 4 \\ w(T_1) &= 9\ 11 \mid 6\ 8 \mid 3\ 5\ 7 \mid 1\ 2\ 4\ 10\ 12 \\ w(S) &= 5\ 6 \mid 2\ 3\ 6 \mid 2\ 4\ 4 \mid 6 \mid 2 \end{aligned}$$

The vertical strokes are included only to indicate row breaks; they are not part of the word.

Remark. A tableau T can be recovered from its word; simply introduce a row break each time an entry in the word is strictly larger than the one directly to its right, as in the first two examples above. We call each subsequence delimited by such a strict decrease a *piece* of the word.

Evidently not every word corresponds to a tableau. In particular, the pieces of the word must have weakly increasing length to ensure the shape of a Young diagram, and the usual order-theoretic properties of a tableau must be respected in each resulting column. On the other hand, every word corresponds to a skew tableau (in fact, many skew tableaux). For instance, if one partitions a word w into weakly increasing pieces, then these pieces can be put in rows. Then a skew tableau is obtained by placing each row above and entirely to the right of the preceding piece's row.

It is often useful to consider the number of times a given entry appears in a (skew) tableau, or equivalently the number of times that a given letter appears within a word.

Definition 4.8. We say that a skew tableau S with entries in $[n]$ has *content* (or *type* or *weight*) $\alpha = (\alpha_1, \ldots, \alpha_n)$ if the multiplicity of i in S is α_i for each i, $1 \leq i \leq n$.

Example. Let T_0, T_1, and S be as in the previous examples of this section. The skew tableau S has content $(0, 3, 1, 2, 1, 3)$. The tableau T_0 has content $(1, 3, 1, 3, 3, 1)$. The standard tableau T_1 has content (1^{12}). In general, a standard tableau with n boxes has content (1^n).

4.2 Schensted Row-Insertion Algorithm, Knuth-Equivalence

In this section we introduce the first of two useful operations which produce tableaux, namely the "row-insertion" algorithm. Along with the "sliding" procedure of Section 4.3, we will use the row-insertion algorithm to provide two equivalent definitions for an associative multiplication of tableaux; this will be the content of Section 4.4.

We begin by describing how the Schensted *row-insertion* (or *row-bumping*) algorithm acts on tableaux, later investigating how the algorithm acts on the words of tableaux. This algorithm takes a tableau T and positive integer x, and constructs a new tableau $T \leftarrow x$ which contains one more box than T. The procedure runs as follows.

- If x is at least as large as all the entries in the first row of T, append a new box to the end of the first row containing x and halt.
- If not, find the least (*i.e.* leftmost) entry, say x_1, of the first row that is strictly greater than x. Remove ("bump") x_1 from its box and place x in the now empty box.
- Take the entry x_1 that was bumped from the first row and repeat the above procedure, row-inserting x_1 into the second row of T.
- Repeat this process on each subsequent row until one of the following conditions is met:
  - The bumped entry is maximal in the row below, and can be appended to the end of the row it is bumped into, or
  - The entry is bumped from the bottom row, in which case it forms a new bottom row containing one box.

Example. Row-inserting $x = 2$ into the tableau $T = \begin{array}{|c|c|c|c|}\hline 1&2&2&3\\\hline 2&3&5&5\\\hline 4&4&6\\\cline{1-3} 5&6\\\cline{1-2}\end{array}$ yields:

$$\begin{array}{|c|c|c|c|l}\hline 1&2&2&3&\leftarrow\ 2\\\hline 2&3&5&5\\\hline 4&4&6\\\cline{1-3} 5&6\\\cline{1-2}\end{array} \longrightarrow \begin{array}{|c|c|c|c|l}\hline 1&2&2&2\\\hline 2&3&5&5&\leftarrow\ 3\\\hline 4&4&6\\\cline{1-3} 5&6\\\cline{1-2}\end{array} \longrightarrow \begin{array}{|c|c|c|c|}\hline 1&2&2&2\\\hline 2&3&3&5\\\hline 4&4&6&\multicolumn{1}{l}{\leftarrow\ 5}\\\cline{1-3} 5&6\\\cline{1-2}\end{array} \longrightarrow \begin{array}{|c|c|c|c|}\hline 1&2&2&2\\\hline 2&3&3&5\\\hline 4&4&5\\\cline{1-3} 5&6&6\\\cline{1-3}\end{array}$$

It is not difficult to verify that the Schensted algorithm will always produce a tableau; it suffices to show that the order-theoretic properties of the tableau rows and columns are preserved at each step. Indeed, by construction we have that each row remains weakly increasing. Now suppose that an entry y bumps an entry z from a given row. The entry directly below z (if one exists) is strictly greater than z by definition of a tableau. Hence z either stays in the same column or moves to the left as it is bumped into the next row. The entry directly above z's new position is at most y, and hence strictly less than z.

The Schensted algorithm is in fact reversible, provided that one identifies the new box that has been added. Running the algorithm in reverse will produce the original tableau as well as the entry that was row-inserted into the first row:

- If z is the entry in the new box, locate the position of z in the row above. In particular, find the greatest (*i.e.* rightmost) entry in the row above which is strictly less than z.

- Place z in the box from the previous step and bump the original lesser entry into the next row above.

- Repeat the above two steps until an entry has been bumped out of the top row.

In fact, this reverse row-insertion algorithm can be run on any tableau and any box in the tableau which is an outside box.

Definition 4.9. A row-insertion $T \leftarrow x$ determines a collection R of boxes, namely those where an element is bumped from a row, together with the box where the last bumped element is placed. We call R the *bumping route* of the row-insertion. We call the bottommost box of the bumping route, namely the box which was added to the diagram of T, the *new box* of the row-insertion.

Example. If we consider the row-insertion of the previous example, then the bumping route R is indicated by the shaded boxes below:

1	2	2	2
2	3	3	5
4	4	5	
5	6	6	

The box in the (4,3) position is the new box of the row-insertion.

We say that a route R is *strictly* (respectively *weakly*) *left* of a bumping route R' if in each row containing a box of R', the route R has a box in the same row which is left of (respectively left of or equal to) the box in R'. We give similar strict and weak definitions when speaking of bumping routes or boxes being above, below, or to the right of one another. The preceding definitions provide us with the necessary terminology for the following lemma, which will be of use in proving the Pieri formulas for Schur polynomials in Section 5.3.

Lemma 4.1 (Row Bumping Lemma)**.** *Consider two successive row-insertions, first row-inserting a positive integer x into a tableau T, then row-inserting an element x' into the tableau $T \leftarrow x$. Suppose that these two row-insertions produce respective bumping routes R and R' and respective new boxes B and B'.*

1. *If $x \leq x'$, then R is strictly left of R', and B is strictly left of and weakly below B'.*

2. *If $x > x'$, then R' is weakly left of R and B' is weakly left of and strictly below B.*

Proof. Let $x \leq x'$, and suppose that x bumps an element y from the first row. The element y' bumped by x' from the first row must lie strictly to the right of the box where x bumped y since all boxes weakly left of x have entries at most x (and hence at most x'). Since the box formerly occupied by y is strictly left of the box formerly occupied by y', we see that $y \leq y'$, and so the same argument applies to each subsequent row. Hence R is strictly left of R'. Note that R cannot stop in a row strictly above the row in which R' stops; moreover, if R and R' stop in the same row, then B must lie to the left of B' by the previous argument. Hence B is strictly left of and weakly below B'.

Now let $x > x'$ and suppose that x and x' bump elements y and y' from the first row, respectively. The box where x' bumps y' must be weakly left of the box where x bumps y. This follows from the fact that x' bumps the least entry strictly greater than it; since $x > x'$, then any such entry is at most x, so $y' \leq x$ and the box of y' cannot be to the right of the box of x. Since $y' \leq x$ and $x < y$, then $y' < y$ and so the same argument applies to each subsequent row. Hence R' is weakly left of R. Moreover, y' will never be maximal in a row occupied by y, so R' must extend at least one row further down than R. Hence B' is weakly left of and strictly below B. $\square$

The row-bumping lemma can be extended as follows.

Proposition 4.2. *Let T be a tableau of shape μ, and let $U := (\cdots((T \leftarrow x_1) \leftarrow x_2) \leftarrow \cdots) \leftarrow x_p$ for some positive integers $x_1, \ldots, x_p$. Let λ be the shape of U. If $x_1 \leq x_2 \leq \cdots \leq x_p$ (resp. $x_1 > x_2 > \cdots > x_p$), then no two boxes in λ/μ lie in the same column (resp. row).*

Conversely, suppose U is a tableau on $\lambda \supset \mu$, with p boxes in λ/μ. If no two boxes in λ/μ lie in the same column (resp. row), then there is a unique tableau T of shape μ and unique positive integers $x_1 \leq x_2 \leq \cdots \leq x_p$ (resp. $x_1 > x_2 > \cdots > x_p$) such that

$$U = (\cdots((T \leftarrow x_1) \leftarrow x_2) \leftarrow \cdots) \leftarrow x_p$$

Proof. Note that the boxes of λ/μ are precisely the new boxes of the subsequent row-insertions of the x_i. The first half of the proposition follows from $p-1$ applications of the row-bumping lemma.

For the converse, first consider the case where no two of the boxes lie in the same column. Perform p consecutive reverse row-bumpings on U using the boxes in λ/μ, proceeding from the rightmost to the leftmost. This gives a tableau T on μ and elements $x_1 \leq x_2 \leq \cdots \leq x_p$ whose uniqueness is guaranteed by the row-bumping lemma. The case where no two boxes lie in the same row is handled similarly, albeit with the reverse row-bumpings proceeding from the bottommost to the topmost box in λ/μ. $\square$

We will now move away from the diagrammatic treatment of tableaux and consider how the Schensted algorithm operates on the words of tableaux. Suppose that a letter (*i.e.* positive integer) x is inserted into a row. Then the Schensted algorithm asks us to factor the word of this row as $u \cdot x' \cdot v$, where u and v are (possibly empty) words and x' is a letter such that $u \leq x$ and $x < x'$. Here $u \leq x$ means that $u_i \leq x$ for each letter u_i in u. We perform the row-insertion by replacing x' with x and bumping x' to the next row down; this yields the tableau with word $x' \cdot u \cdot x \cdot v$. In sum, the Schensted algorithm acts on the words of rows by

$$(u \cdot x' \cdot v) \cdot x \rightsquigarrow x' \cdot (u \cdot x \cdot v), \quad u \leq x < x' \leq v \tag{4.2.1}$$

where the letters of u and v are weakly increasing, and $u \leq v$ means $u_i \leq v_j$ for any two letters u_i in u, v_j in v.

Example. Consider the tableau T from the first example of this section. The row word of T is given by $w(T) = 5644623551223$. Row-inserting $x = 2$ into this word gives

$$\begin{aligned}(56)(446)(2355)(1223)\cdot 2 &\rightsquigarrow (56)(446)(2355)\cdot 3\cdot (1222)\\ &\rightsquigarrow (56)(446)\cdot 5\cdot (2335)(1222)\\ &\rightsquigarrow (56)\cdot 6\cdot (445)(2335)(1222)\\ &\rightsquigarrow (566)(445)(2335)(1222)\end{aligned}$$

which is the word of the tableau $T \leftarrow x$ as obtained earlier.

Although the transformation (4.2.1) does correctly describe the action of the Schensted algorithm on the word of a tableau, the description is incomplete in the sense that it does not describe how to find such a factorization $u \cdot x' \cdot v$. In order to address this issue, we will look more closely at how the Schensted algorithm acts on triplets of neighbouring letters in a word. Maintaining the same notation as in (4.2.1), then within a given row, the Schensted algorithm amounts to performing two sequences of transformations acting on adjacent letters. The first sequence of transformations runs as follows:

$$\begin{aligned}ux'v_1\cdots v_{q-1}v_qx &\longmapsto ux'v_1\cdots v_{q-1}xv_q && (x < v_{q-1} \le v_q)\\ &\longmapsto ux'v_1\cdots v_{q-2}xv_{q-1}v_q && (x < v_{q-2} \le v_{q-1})\\ \cdots &\longmapsto ux'v_1xv_2\cdots v_{q-1}v_q && (x < v_1 \le v_2)\\ &\longmapsto ux'xv_1\cdots v_{q-1}v_q && (u \le x < x' \le v_1)\end{aligned}$$

For this stage, the reiterated transformation is

$$xyz \mapsto xzy \qquad (z < x \le y) \tag{K'}$$

At this point, x bumps x' and x' moves successively to the left, yielding a second sequence of transformations:

$$\begin{aligned}u_1\cdots u_{p-1}u_px'xv &\longmapsto u_1\cdots u_{p-1}x'u_pxv && (u_p \le x < x')\\ &\longmapsto u_1\cdots x'u_{p-1}u_pxv && (u_{p-1} \le u_p < x')\\ \cdots &\longmapsto u_1x'u_2u_3\cdots u_pxv && (u_2 \le u_3 < x')\\ &\longmapsto x'u_1u_2\cdots u_pxv && (u_1 \le u_2 < x')\end{aligned}$$

For this stage, the reiterated transformation is

$$xyz \mapsto yxz \qquad (x \le z < y) \tag{K''}$$

Definition 4.10. The transformations K', K'', or their inverses, are called *elementary Knuth transformations* on a word when applied to three consecutive letters in the word.

Another way to realize the elementary Knuth transformations is as follows. Suppose that xyz are three consecutive letters in a word. If one of y and z is smaller than x and the other greater than x, then the transformations K' and K'^{-1} allow us to interchange y and z. This still applies if one of the neighbours y or z is equal to x, following the usual convention that the rightmost of two equal letters is greater. Likewise, if one of x and y is smaller than z and the other greater than z, then the transformations K'' and K''^{-1} allow us to interchange x and y.

An important observation is that the elementary Knuth transformations may be used to define the following equivalence relation.

Definition 4.11. Two words w, w' are *Knuth-equivalent* if they can be changed to the other by a sequence of elementary Knuth transformations. In this case we write $w \equiv w'$.

We have the following proposition, which follows immediately from the initial discussion motivating the elementary Knuth transformations.

Proposition 4.3. *If T is a tableau and x is a letter, then $w(T \leftarrow x) \equiv w(T) \cdot x$.*

Although the notion of Knuth-equivalence may seem unmotivated at present, its significance will become apparent in the sections that follow. In particular, Knuth-equivalence provides a crucial link between the Schensted algorithm and the seemingly unrelated "sliding" algorithm we will explore in the following section. The important connection is established by the following central theorem of Knuth-equivalence on words.

Theorem 4.4. *Every word is Knuth-equivalent to the word of a unique tableau.*

The existence of such a tableau is easy to establish, and there is a so-called *canonical procedure* to construct it. In particular, if $w = x_1 \cdots x_r$ is a word, then the word

$$P(w) := \Big(\Big(\cdots\Big(\Big(\boxed{x_1} \leftarrow x_2\Big) \leftarrow x_3\Big) \leftarrow \cdots\Big) \leftarrow x_{r-1}\Big) \leftarrow x_r$$

constructed by the canonical procedure is the word of a tableau, and is Knuth-equivalent to w by Proposition 4.3. The uniqueness claim is far less trivial to prove, and requires a number of new ideas and results. As these results are not of tremendous importance for what follows, we have omitted the uniqueness proof from the present discussion. The interested reader can find the proof in Chapter 3 of [2].

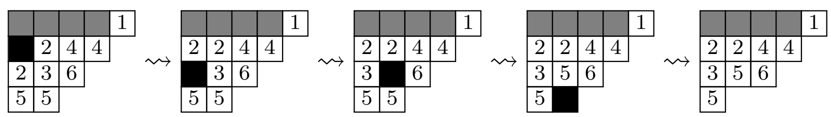
4.3 Sliding, Rectification of Skew Tableaux

In this section we concern ourselves with a second tableau operation, namely the "sliding" algorithm due to Schützenberger. While the Schensted algorithm of the previous section operated on a tableau and letter to create a new tableau including that letter, Schützenberger's algorithm suitably rearranges the box of a skew tableau so as to form a tableau after sufficiently many iterations.

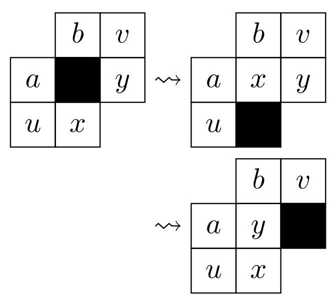
Given a skew tableau S, the basic sliding algorithm runs as follows.

- Select an inside corner ("empty box") of the skew tableau S, and consider the two neighbouring entries immediately below and immediately to the right of the empty box.
- Identify the smaller of these two neighbouring entries, taking the leftmost of equal entries to be smaller, as usual. Slide this smaller entry into the empty box, creating a new empty box.
- Repeat the previous step with the new empty box until the empty box becomes an outside corner. At this point, remove the empty box from the diagram.

Example. Starting from the inside corner in the (2,1) position, the sliding algorithm acts on the following skew tableau as follows:



The fact that the sliding algorithm produces a skew diagram is clear, as the algorithm amounts to adding a box to an inside corner, and removing the box from an outside corner. The fact that the sliding algorithm produces a skew *tableau* is also simple to verify. At each stage, there are two cases to consider:



$(x \leq y)$

$(x > y)$

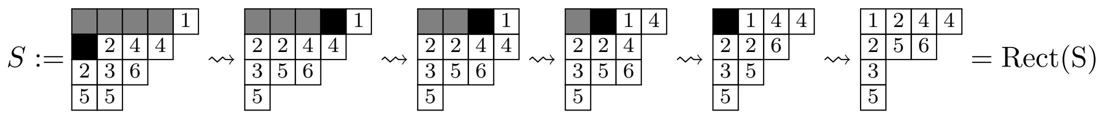
Note that any of the boxes in the basic configuration above may be empty. In each of the above cases, it is easy to verify that the rows remain weakly increasing and columns remain strictly increasing. Consequently, the tableau ordering is preserved throughout the entire procedure.

As with the Schensted algorithm, the sliding algorithm is reversible subject to additional information. In particular, given a skew tableau and the (outside corner) box that was removed, we can perform the *reverse slide* to recover the original skew tableau. The reverse slide is conducted by running the sliding algorithm in reverse, as follows. Starting with the removed box, slide the greater of its two neighbours immediately above and immediately to the left into the empty box. Repeat this procedure until the empty box becomes an inside corner.

Observe that applying the sliding algorithm to any inside corner of a skew tableau S of shape λ_0/μ_0 produces a new skew tableau S_1 of shape λ_1/μ_1 with $|\lambda_1| = |\lambda_0|$ and $|\mu_1| = |\mu_0| - 1$. Moreover, the entries of S_1 are simply a rearrangement the entries of S. Consequently, the sliding algorithm can be performed on each resulting skew tableau until one obtains a skew tableau with no inside corners; namely, a tableau. The entries of this tableau will moreover be a rearrangement of the entries of S.

Definition 4.12. The resulting tableau is called the *rectification* of S, which we denote by $\mathrm{Rect}(S)$. The process of applying subsequent slidings to obtain $\mathrm{Rect}(S)$ is called the *jeu de taquin.*

Example. Continuing from the previous example, we obtain



Although it is not obvious at this point, the rectification of a skew tableau is in fact well-defined. That is, the tableau $\mathrm{Rect}(S)$ is independent of the choice of inside corner at each stage of the jeu de taquin. In order to justify this claim, we will consider how the sliding algorithm affects the words of skew tableaux. In particular, we have the following proposition.

Proposition 4.5. *If one skew tableau can be obtained from another by a sequence of slides, then their words are Knuth-equivalent.*

Proof. We first note that the result of a slide need not be a skew tableau; quite often it will be a skew tableau with a hole. In any case, the word is defined in the same way by listing the entries from left to right in each row, from the bottom row to the top. Thus it suffices to show that the Knuth-equivalence class of the word is unchanged by a single horizontal slide or a single vertical slide. In the case of a horizontal slide, the Knuth class of the word is clearly preserved as the word itself is unchanged. The case of a vertical slide requires a more involved analysis. As the details are ultimately unimportant for our purposes, we will omit the proof of this case. The interested reader can find a complete proof in [2, p. 20-22]. $\square$

This proposition, together with Theorem 4.4, implies the following corollary, which establishes that the rectification is well-defined.

Corollary 4.6. *The rectification of a skew tableau S is the unique tableau whose word is Knuth-equivalent to the word of S. If S, S' are skew tableaux, then $Rect(S) = Rect(S')$ if and only if $S \equiv S'$.*

4.4 Products of Tableaux

At this point we are now equipped with the concepts required to define a multiplication of tableaux. In light of the results we have collected on Knuth-equivalence, we will define our product in terms of the words of tableaux, and use the algorithms of Schensted and Schützenberger to facilitate explicit computations. The utility of this tableau product will become apparent beginning in Section 5.3 when we establish an explicit correspondence between Schur polynomials and tableaux. We will then be able to translate questions about products and factorizations of Schur polynomials (or more specifically their monomial components) into corresponding questions about products and factorizations of tableaux. With this in mind, we define the product of two tableaux as follows.

Definition 4.13. Let T and U be tableaux with words $w(T) = t_1 \cdots t_r$ and $w(U) = u_1 \cdots u_s$, respectively. We define the *product tableau* $T \cdot U$ to be the unique tableau whose word is Knuth-equivalent to the word $w(T) \cdot w(U)$ given by juxtaposition.

By Theorem 4.4, we see that this multiplication is well-defined. It is moreover easy to derive several basic properties, which can be seen by investigating the corresponding properties of the juxtaposition product of words, and are summarized in the following proposition.

Proposition 4.7. *Consider the set of tableaux with the product $\cdot$ defined as above. The product $\cdot$ is associative, and an identity element is given by the empty tableau $\varnothing$, the unique tableau on the empty partition* (0)*. In other words, the set of tableaux with the product $\cdot$ and identity $\varnothing$ forms a monoid.*

Remark. If T and U are tableaux consisting of r and s boxes, respectively, then it is clear from the definition that the product tableau $T \cdot U$ consists of $r + s$ boxes. As any nonempty tableau consists of a strictly positive number of boxes, it can easily be seen that $\varnothing$ is the only invertible element of this monoid.

By repeated applications of Proposition 4.3, we deduce the following method to explicitly compute the product of tableaux using the row-insertion algorithm. Let T, U be tableaux and suppose that U has row word $w(U) = u_1 \cdots u_s$. Then the product tableau $T \cdot U$ is given by

$$T \cdot U = ((\cdots((T \leftarrow u_1) \leftarrow u_2) \leftarrow \cdots) \leftarrow u_{s-1}) \leftarrow u_s \tag{4.4.1}$$

Example.

$$\begin{ytableau} 1 & 2 & 3 \\ 3 & 3 \\ 4 & 6 \end{ytableau} \cdot \begin{ytableau} 1 & 3 \\ 2 \end{ytableau} = \begin{ytableau} 1 & 2 & 2 \\ 3 & 3 & 3 \\ 4 & 6 \end{ytableau} \cdot \begin{ytableau} 1 & 3 \end{ytableau} = \begin{ytableau} 1 & 1 & 2 \\ 2 & 3 & 3 \\ 3 & 6 \\ 4 \end{ytableau} \cdot \begin{ytableau} 1 \end{ytableau} = \begin{ytableau} 1 & 1 & 2 & 3 \\ 2 & 3 & 3 \\ 3 & 6 \\ 4 \end{ytableau}$$

Given this simple means to perform explicit computations, we can also readily observe that the product $\cdot$ defined on tableaux is *not* commutative in general, even in simple cases.

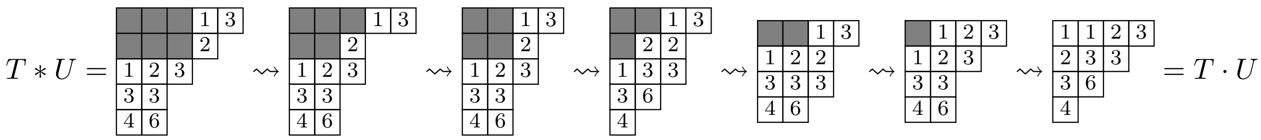
Example.

$$\begin{ytableau} 1 & 2 \end{ytableau} \cdot \begin{ytableau} 1 \end{ytableau} = \begin{ytableau} 1 & 1 \\ 2 \end{ytableau} \neq \begin{ytableau} 1 & 1 & 2 \end{ytableau} = \begin{ytableau} 1 \end{ytableau} \cdot \begin{ytableau} 1 & 2 \end{ytableau}$$

A second explicit method for computing tableau products is given by the sliding algorithm. Let T and U be tableaux, and define the skew tableau $T * U$ diagramatically by placing U immediately above and to the right of T. It is clear that $w(T * U) = w(T) \cdot w(U)$. By Corollary 4.6, we see that

$$T \cdot U = \mathrm{Rect}(T * U) \tag{4.4.2}$$

Example. If $T = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 3 & 3 \\ \cline{1-2} 4 & 6 \\ \cline{1-2}\end{array}$ and $U = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}$ as in our earlier example, then we have



Having now defined a multiplication of tableaux, we are ready to establish an initial correspondence between tableaux and polynomials. To do so, first recall our earlier remark that the set of tableaux form a monoid. As with any monoid, we may form a corresponding monoid ring by treating the monoid elements as indeterminates over an additive group, such as $\mathbb{Z}$. In particular, we will consider this construction with the monoid M of tableaux with entries in the alphabet $[m] := \{1, \ldots, m\}$, where m is a positive integer, and the additive group $\mathbb{Z}$. We define the *tableau ring* $R_{[m]}$ to be the free $\mathbb{Z}$-module with basis given by the tableaux with entries in $[m]$. Multiplication in $R_{[m]}$ is of course defined by our usual multiplication of tableaux, and $R_{[m]}$ forms an associative, noncommutative ring.

There is a canonical ring homomorphism from $R_{[m]}$ into the polynomial ring $\mathbb{Z}[x_1, \ldots, x_m]$. In particular, if T is a tableau with entries in $[m]$, then we define the monomial x^T to be the product

$$x^T := x^\alpha := \prod_{i=1}^{m} x_i^{\alpha_i} \tag{4.4.3}$$

where $\alpha = (\alpha_1, \ldots, \alpha_m)$ is the content of T (cf. Definition 4.8). In other words, α_i denotes the multiplicity of i in the tableau T; that is, the number of boxes of T which contain i as an entry. We then have the canonical ring homomorphism $R_{[m]} \to \mathbb{Z}[x_1, \ldots, x_m]$ given by mapping a tableau to its monomial, *i.e.* $T \mapsto x^T$. The key observation to see that this is a homomorphism is the fact that the entries of a product of two tableaux are precisely the entries of the two tableau factors. Hence the multiplicity behaves additively with respect to tableau multiplication.

Remark. A more general variation of the above construction can be carried out by extending the alphabet of our tableaux to be the entire set of positive integers. In this construction, we allow elements of our tableau ring R to be formal power series of tableaux with coefficients in $\mathbb{Z}$. In this case, we define the monomial of a tableau and the canonical homomorphism $T \mapsto X^T$ in the same way; the degree of each monomial will remain finite since any given tableau contains only finitely many entries. The difference in this case is that the canonical homomorphism maps elements of the tableau ring to elements of the ring of formal power series in infinitely many variables with coefficients in $\mathbb{Z}$.

We conclude this section with a special case of the central result of Section 5.3, namely that the Schur polynomials admit a remarkably simple expression as a sum of tableau monomials. In particular, we will concern ourselves for the time being with tableau monomial expressions for the elementary and complete homogeneous symmetric polynomials, which are provided in the following proposition.

Proposition 4.8. *Let λ be a partition and denote by $\mathcal{T}(\lambda)$ the set of tableaux with shape λ and positive integer entries. Then if $k \geq 0$, we have*

$$e_k = \sum_{T \in \mathcal{T}(1^k)} x^T \qquad \text{and} \qquad h_k = \sum_{T \in \mathcal{T}(k)} x^T \tag{4.4.4}$$

Restriction to tableaux with entries in $[m]$ yields the corresponding expressions for symmetric polynomials in m variables, valid for all $k \leq m$:

$$e_k(x_1, \ldots, x_m) = \sum_{T \in \mathcal{T}_{[m]}(1^k)} x^T \qquad \text{and} \qquad h_k(x_1, \ldots, x_m) = \sum_{T \in \mathcal{T}_{[m]}(k)} x^T \tag{4.4.5}$$

where $\mathcal{T}_{[m]}(\lambda)$ denotes the set of tableaux with shape λ and entries in $[m]$.

Proof. We will first prove (4.4.5). Recall that the partitions (1^k) and (k) correspond to a column and row of length k, respectively. Tableaux on (k) with entries in $[m]$ correspond (by their words) to weakly increasing sequences in $[m]$ of length k. Likewise, tableaux on (1^k) with entries in $[m]$ correspond (by reversing their words) to strictly increasing sequences in $[m]$ of length k. The identities in (4.4.5) follow by comparing with the combinatorial expressions in Propositions 2.1 and 2.5 for elementary and complete symmetric polynomials, respectively.

The proof of (4.4.4) is a simple generalization of this argument, noting that the combinatorial expressions in Propositions 2.1 and 2.5 generalize naturally for symmetric functions in infinitely many variables. □

Remark. Recall that the elementary and homogeneous symmetric functions are special cases of the Schur functions by Propositions 3.9 and 3.11, with $e_k = S_{(1^k)}$ and $h_k = S_{(k)}$ for each $k \geq 0$. In fact, the expressions of the previous proposition generalize directly to arbitrary Schur polynomials, although the proof is more involved in the general case; this will be the content of Sections 5.2 and 5.3.

4.5 Robinson-Schensted-Knuth Correspondence

In this section, we introduce the Robinson-Schensted-Knuth (RSK) correspondence, a remarkable combinatorial bijection between pairs of tableaux and matrices with nonnegative integer entries.

Recall that for each word w, there exists a unique tableau $P(w)$ whose word is Knuth-equivalent to w. In particular, if $w = x_1 \cdots x_r$, then $P(w)$ can be constructed by the canonical procedure:

$$P(w) = \left(\left(\cdots\left(\left(\boxed{x_1} \leftarrow x_2\right) \leftarrow x_3\right) \leftarrow \cdots\right) \leftarrow x_{r-1}\right) \leftarrow x_r$$

Given a tableau constructed by the canonical procedure, we may be interested in determining which word was used as the blueprint. The tableau alone is insufficient to determine this word uniquely; it will only provide a Knuth-equivalence class of possible words. Recall however that the Schensted algorithm is completely reversible, provided that one has the additional information of which box has been added to the tableau at each stage. Consequently, one can recover the word w from the tableau $P(w)$ together with the numbering of the boxes that arise in the canonical procedure. To preserve this information, we use a *recording* (or *insertion*) *tableau*, which we construct as follows.

Definition 4.14. Let $w = x_1 \cdots x_r$ be a word of length r. The *recording tableau* $Q(w)$ is a standard tableau with the same shape as $P(w)$. The entries of $Q(w)$ are the integers $1, \dots, r$ with the integer k placed in the box added at the k-th step of the canonical construction of $P(w)$. If we let

$$P_k := \Big(\Big(\cdots\Big(\Big(\boxed{x_1} \leftarrow x_2\Big) \leftarrow x_3\Big) \leftarrow \cdots\Big) \leftarrow x_{k-1}\Big) \leftarrow x_k$$

and denote by Q_k the corresponding insertion tableau obtained at the k-th step of the canonical procedure, then we construct $Q(w) = Q_r$ recursively as follows. We begin by letting $Q_1 = \boxed{x_1}$. We then construct Q_k from Q_{k-1} by adding a new box in the same position as the new box added to P_k from P_{k-1}; we fill this new box with the integer k.

Note that the new box added to P_k (and hence Q_k) at each stage of the canonical procedure is always an outside corner. Hence the entry k which is placed in this box will be larger than the entries above or to its left, so the result will be a tableau.

Example. Let $w = 5\,4\,8\,2\,3\,4$. Then for each $k \geq 1$, the pair (P_k, Q_k) is given by

$$k=1: \quad \begin{ytableau} 5 \end{ytableau} \quad \begin{ytableau} 1 \end{ytableau} \qquad k=2: \quad \begin{ytableau} 4 \\ 5 \end{ytableau} \quad \begin{ytableau} 1 \\ 2 \end{ytableau} \qquad k=3: \quad \begin{ytableau} 4 & 8 \\ 5 \end{ytableau} \quad \begin{ytableau} 1 & 3 \\ 2 \end{ytableau} \qquad k=4: \quad \begin{ytableau} 2 & 8 \\ 4 \\ 5 \end{ytableau} \quad \begin{ytableau} 1 & 3 \\ 2 \\ 4 \end{ytableau}$$

$$k=5: \quad \begin{ytableau} 2 & 3 \\ 4 & 8 \\ 5 \end{ytableau} \quad \begin{ytableau} 1 & 3 \\ 2 & 5 \\ 4 \end{ytableau} \qquad k=6: \quad P(w) = \begin{ytableau} 2 & 3 & 4 \\ 4 & 8 \\ 5 \end{ytableau} \quad Q(w) = \begin{ytableau} 1 & 3 & 6 \\ 2 & 5 \\ 4 \end{ytableau}$$

To recover the original word w from the pair (P, Q), we perform the following procedure to reduce from (P_k, Q_k) to (P_{k-1}, Q_{k-1}). Begin by locating the box in Q_k with the largest element (namely k). We obtain P_{k-1} by reverse row-inserting the element in the corresponding box in P_k. The entry that is bumped out of the top row is the k-th letter x_k in the word w. We obtain Q_{k-1} from Q_k by removing the box containing the entry k.

By virtue of the fact that the canonical procedure is reversible, we see that any pair of tableaux (P, Q), with Q a standard tableau, arises in this way. Indeed, if Q is a standard tableau, then we may perform the word recovery procedure to obtain a word w such that $P = P(w)$ and $Q = Q(w)$. We therefore have the following proposition.

Proposition 4.9. *The canonical procedure together with the construction of a recording tableau, and the word recovery inverse procedure, establish a bijective correspondence between the following sets:*

- *The set of words w of length r in the alphabet $[n]$, where $n \geq r$, and*

- *The set of ordered pairs (P, Q) of tableaux on the same shape $\lambda \vdash r$, with entries of P taken from $[n]$ and Q a standard tableau.*

We call this bijection the Robinson-Schensted correspondence.

The Robinson-Schensted correspondence is in fact a generalization of an earlier result due to Robinson, which is interesting in itself. Suppose that $n = r$, so that the letters of the word w are the letters $1, \ldots, n$, each occurring once. Then w corresponds to a permutation of the set $[n]$, namely that which sends i to the i-th letter of w. In this case, P is a standard tableau. Conversely, if P is a standard tableau, then $n = r$ and w corresponds to a permutation $\sigma \in S_n$. We call this bijective correspondence between permutations in S_n and pairs of standard tableaux on the same shape $\lambda \vdash n$ the *Robinson correspondence.*

Knuth in turn generalized the Robinson-Schensted correspondence to arbitrary ordered pairs (P, Q) of tableau on the same shape. The generalized procedure and its inverse require only minor modifications from the Robinson-Schensted correspondence. In particular, suppose that (P, Q) are tableaux on the same shape, with P having entries in $[m]$ and Q having entries in $[n]$. Then one can perform the reverse process described above to obtain a sequence of pairs of tableaux on the same shape:

$$(P, Q) = (P_r, P_r), (P_{r-1}, P_{r-1}), \ldots, (P_2, P_2), (P_1, P_1)$$

each having one fewer box than the preceding pair. Explicitly, we construct (P_{k-1}, Q_{k-1}) from (P_k, Q_k) as follows. Begin by locating the box in Q_k containing the greatest element, as usual treating the rightmost of equal elements as the greatest. We obtain Q_{k-1} by removing this box from Q_k. We obtain P_{k-1} by reverse row-inserting the entry from the box located in the corresponding position in P_k.

If we let u_k be the element removed from Q_k, and let v_k be the element bumped out of the top row of P_k, then one obtains a two-rowed array $w = \left(\begin{smallmatrix} u_1 & \cdots & u_r \\ v_1 & \cdots & v_r \end{smallmatrix} \right)$. By construction, the $u_i \in [m]$ appearing in the top row of w are weakly increasing, so $u_1 \leq u_2 \leq \cdots \leq u_r$. By the row-bumping lemma 4.1, the $v_i \in [n]$ appearing in the bottom row of w satisfy $v_{k-1} \leq v_k$ if $u_{k-1} = u_k$. We say that a two-rowed array w satisfying these two properties is *in lexicographic order.* Evidently, given a two-rowed array w in lexicographic order, we can construct a pair (P, Q) of tableaux on the same shape using essentially the same procedure as before. In particular, we construct P by the canonical procedure using the entries of the bottom row of w, from left to right. The entries of the recording tableau Q are given in this case by the corresponding entries of the top row of w.

This construction of course reduces to the earlier Robinson-Schensted correspondence. In particular, if Q is a standard tableau, then we have an array $w = \left(\begin{smallmatrix} 1 & \cdots & r \\ v_1 & \cdots & v_r \end{smallmatrix} \right)$ which we may interpret as the word $v_1 \cdots v_r$. In particular, we call a two-rowed array a *word* if $u_i = i$ for each index $1 \leq i \leq r$. When the v_i are the r distinct elements of $[r]$, we call the array a *permutation.* We summarize this in the following theorem.

Theorem 4.10 (RSK Theorem)**.** *The above procedures establish a bijective correspondence between two-rowed lexicographic arrays w and ordered pairs (P, Q) of tableaux on the same shape. In particular,*

1. *w has r entries in each row $\iff$ P and Q each have r boxes. The entries of P are the entries of the bottom row of w, and the entries of Q are the entries of the top row of w.*

2. *w is a word $\iff$ Q is a standard tableau.*

3. *w is a permutation $\iff$ P and Q are both standard tableaux.*

We call this bijection the Robinson-Schensted-Knuth (RSK) correspondence*.*

The RSK correspondence can be extended in turn to a correspondence between ordered pairs (P, Q) of tableaux and matrices with nonnegative integer entries. We first note that an arbitrary two-rowed array determines a unique lexicographic array by rearranging its columns into lexicographic order. Thus, for the purposes of the RSK correspondence, we may work with two-rowed arrays modulo permutation of the columns. There is a natural way to express two-rowed arrays in terms of matrices as follows.

Consider a two-rowed array such that the entries of the top row are elements of $[m]$, and the bottom row entries are elements of $[n]$. We can identify this array by a collection of columns, which we represent as pairs (i, j) with $i \in [m]$ and $j \in [m]$, with each pair having some nonnegative multiplicity; namely, the number of times that $\binom{i}{j}$ appears as a column in the array. Consequently, we may associate an equivalence class of arrays with an $m \times n$ matrix A, where A_{ij} is the multiplicity of (i, j) in the array.

Example. We associate the array $w = \left(\begin{smallmatrix} 1 & 1 & 1 & 2 & 2 & 3 & 3 & 3 & 3 \\ 1 & 2 & 2 & 1 & 2 & 1 & 1 & 1 & 2 \end{smallmatrix}\right)$ with the matrix $A = \left(\begin{smallmatrix} 1 & 2 \\ 1 & 1 \\ 3 & 1 \end{smallmatrix}\right)$.

Consequently, the RSK correspondence provides a bijective correspondence between matrices with nonnegative integer entries and ordered pairs (P, Q) of tableaux on the same shape. There are a number of simple properties that follow from the correspondence. For instance, if a matrix A corresponding to a pair of tableaux (P, Q) has size $m \times n$, then P will have entries in $[n]$ and Q will have entries in $[m]$.

A more interesting observation is that the i-th row sum of A is precisely the number of times that i occurs in the top row of the corresponding matrix, or equivalently the number of times that i occurs in the tableau Q. Similarly, the j-th column sum is the number of times that j occurs in P. Consequently, A is the matrix of a word if and only if each row of A consists of one 1 with all other entries 0. Likewise, A is the matrix corresponding to a permutation w if and only if the same property holds for each column. In this case, A is in fact the permutation matrix of w, following the convention that the permutation matrix of w has a 1 in the $w(i)$-th column of the i-th row, with all other entries 0.

In this section, we have defined the RSK correspondence between pairs of tableaux and matrices by factoring into two distinct correspondences with two-rowed arrays. In fact, there are combinatorial algorithms which establish the RSK correspondence between pairs of tableaux and matrices directly, with no intermediate reference to an array. This so-called "matrix-ball" construction will not be of particular use for our purposes, so we will not provide any further discussion. A thorough description of the algorithms comprises Section 4.2 of [2]. One property of the RSK correspondence that follows quite clearly from the matrix-ball construction which we will make use of is the following theorem.

Theorem 4.11 (Symmetry Theorem)**.** *If an array* $\left(\begin{smallmatrix} u_1 & \cdots & u_r \\ v_1 & \cdots & v_r \end{smallmatrix}\right)$ *corresponds to the pair of tableaux* (P, Q)*, then the array* $\left(\begin{smallmatrix} v_1 & \cdots & v_r \\ u_1 & \cdots & u_r \end{smallmatrix}\right)$ *corresponds to the pair of tableaux* (Q, P)*. In particular, if* w *is a permutation, then* $P(w^{-1}) = Q(w)$ *and* $Q(w^{-1}) = P(w)$*.*

We conclude with some interesting corollaries of the Symmetry Theorem.

Corollary 4.12. *Symmetric matrices with nonnegative integer entries correspond bijectively with tableaux.*

Proof. By the definition of our array-matrix correspondence, turning an array upside down amounts to taking the transpose of the corresponding matrix. In this context, the Symmetry Theorem states that a matrix A corresponds to a pair of tableaux (P, Q) iff its transpose A^T corresponds to the pair (Q, P). In particular, symmetric matrices of size n correspond to pairs of the form (P, P) where P is a tableau with entries in $[n]$, hence the conclusion follows. $\square$

Corollary 4.13. *Standard tableaux with* n *boxes correspond bijectively with involutions* $w \in S_n$*.*

Proof. Consider the special case of Corollary 4.12 where the tableau P in question is a standard tableau. By the RSK Theorem, P corresponds to a permutation $w \in S_n$, where n is the number of boxes in P. By the previous corollary the corresponding permutation matrix is symmetric, which forces w to be an involution. $\square$

5 Combinatorics of Schur Polynomials

Having introduced some core aspects of the theory of Young tableaux, we are now ready to connect some of these combinatorial results to our study of symmetric polynomials. Our primary goal in the first two sections will be to arrive at an alternate definition for the Schur polynomials in terms of tableaux and prove that this agrees with the original determinantal definition from Section 3.2. This will enable us in the following sections to prove a number of algebraic results regarding Schur polynomials, most notably the Littlewood-Richardson Rule, which provides a relatively simple combinatorial method to compute the product of two Schur polynomials.

There are many ways that one may prove the equivalence of the determinantal and combinatorial definitions of the Schur polynomials. We will follow the primarily algebraic approach of Macdonald; this comprises the content of Sections 5.1, 5.2, and the first half of Section 5.3, which are adapted from [6, Ch. I, §4, 5]. Once the equivalence of definitions for the Schur polynomials is established, we return to Fulton's treatment of the combinatorial theory. In particular, the proof of the Pieri formulas in Section 5.3 is adapted from [2, p. 24-25] and Section 5.4 on the Littlewood-Richardson Rule is based on [2, p. 58-66].

5.1 Hall Inner Product

As we have seen in Section 3.2, the Schur functions S_λ form a $\mathbb{Z}$-basis for the ring of symmetric functions. Remarkably, there is a natural inner product on the ring of symmetric functions such that the Schur functions form an orthonormal basis. This utility of this inner product will become apparent in the next section when we use it to define the *skew Schur polynomials*, which in turn will help establish a link between the algebraic and combinatorial aspects of Schur polynomials. In this section we will restrict our focus to the definition and basic properties of the Hall inner product.

The Hall inner product is in some sense defined by series expansions of the Cauchy product $\prod_{i,j}(1-x_iy_j)^{-1}$ from Section 3.4. We have already seen one expansion of the Cauchy product in terms of Schur polynomials, namely the expansion given by (3.4.1). Recall that for any partition λ, we define $h_\lambda := \prod_{i=1}^{l(\lambda)} h_{\lambda_i}$, and that the h_λ form a $\mathbb{Z}$-basis of the ring of symmetric functions. The Cauchy product in fact admits a second series expansion in terms of the h_λ and m_λ.

Proposition 5.1. *Let $x = \{x_1, x_2, \ldots\}$ and $y = \{y_1, y_2, \ldots\}$ be sets of independent variables. Then*

$$\prod_{i,j}(1-x_iy_j)^{-1} = \sum_\lambda h_\lambda(x)m_\lambda(y) = \sum_\lambda m_\lambda(x)h_\lambda(y) \tag{5.1.1}$$

where the sums are taken over all partitions λ.

Proof. We have

$$\prod_{i,j}(1-x_iy_j)^{-1} = \prod_j H(y_j) = \prod_j \sum_{k=0}^{\infty} h_k(x)y_j^r = \sum_\alpha h_\alpha(x)y^\alpha = \sum_\lambda h_\lambda(x)m_\lambda(y)$$

where the second to last sum is taken over all sequences $\alpha \in \bigoplus_{i=1}^{\infty} \mathbb{N}$ of nonnegative integers with finitely many nonzero terms. The rightmost expression above follows by rearranging the terms of the series into their symmetric group orbits. The equality of the rightmost expression in (5.1.1) follows by a symmetric argument, beginning with the identity $\prod_{i,j}(1 - x_i y_j)^{-1} = \prod_i H(x_i)$. □

With the identity (5.1.1) established, we are ready to define the Hall inner product and list some basic properties and related results, the most significant of which is Theorem 5.2 which states that the Schur functions are an orthonormal basis of Λ.

Definition 5.1 (Hall Inner Product)**.** We define the *Hall inner product* on Λ to be the $\mathbb{Z}$-valued bilinear form $\langle u, v\rangle$ such that the $\mathbb{Z}$-bases given by the h_λ and m_λ are dual to one another. In particular, we define the Hall inner product by requiring

$$\langle h_\lambda, m_\mu \rangle := \delta_{\lambda\mu} := \begin{cases} 1 & \lambda = \mu \\ 0 & \lambda \neq \mu \end{cases}$$

Theorem 5.2. *For all partitions λ, μ, we have*

$$\langle S_\lambda, S_\mu \rangle = \delta_{\lambda\mu} \tag{5.1.2}$$

That is, the Schur functions are an orthonormal basis for the symmetric functions with the Hall inner product. If Λ^n denotes the space of symmetric functions homogeneous of degree n, then the Schur functions S_λ with $\lambda \vdash n$ form an orthonormal basis of Λ^n.

Proof. Since the h_λ and m_λ form $\mathbb{Z}$-bases, we may write $S_\lambda = \sum_\rho a_{\lambda\rho} h_\rho$, $S_\mu = \sum_\rho b_{\mu\rho} m_\rho$. Then by computing the inner product $\langle S_\lambda, S_\mu\rangle$, we see that

$$\langle S_\lambda, S_\mu \rangle = \sum_\rho a_{\lambda\rho} b_{\mu\rho}$$

By virtue of the Cauchy identities (3.4.1) and (5.1.1), we see that

$$\sum_\rho h_\rho(x) m_\rho(y) = \sum_\rho S_\rho(x) S_\rho(y) = \sum_\rho \left(\sum_\lambda a_{\rho\lambda} h_\lambda(x)\right)\left(\sum_\mu b_{\rho\mu} m_\mu(y)\right)$$

By comparing coefficients in the leftmost and rightmost expressions, we obtain

$$\sum_\rho a_{\rho\lambda} b_{\rho\mu} = \delta_{\lambda\mu}$$

Now consider square matrices A and B whose entries are indexed by partitions so that $A = (a_{\rho\sigma})$ and $B = (b_{\rho\sigma})$. In this context, the previous equation can be interpreted as saying that $A^T B = I$, namely $A^T = B^{-1}$. This is of course the case if and only if $A = (A^T)^T = (B^{-1})^T = (B^T)^{-1}$. Hence we have the identity $AB^T = I$, which translates componentwise into the identity

$$\sum_\rho a_{\lambda\rho} b_{\mu\rho} = \delta_{\lambda\mu}$$

Since the lefthand expression is equal to $\langle S_\lambda, S_\mu \rangle$, this completes the proof. □

Remark. The proof of Theorem 5.2 can be generalized to show that the power sum symmetric functions $p_\lambda := \prod_{i=1}^{l(\lambda)} p_{\lambda_i}$ form an orthogonal basis of the space $\Lambda_\mathbb{Q}$ of symmetric functions with rational coefficients. The relevant discussion can be found in [6, p. 62-64]. The same argument can also be found later in this document in the proof of Proposition 6.2.

Proposition 5.3. *The Hall inner product is symmetric and positive definite. That is, for all $f, g \in \Lambda$ we have $\langle f, g \rangle = \langle g, f \rangle$ and $\langle f, f \rangle \geq 0$, with equality iff $f = 0$.*

Proof. Let $f, g \in \Lambda$. If we write $f = \sum_\lambda a_\lambda S_\lambda$, $g = \sum_\mu b_\mu S_\mu$, then by bilinearity we may expand the inner product as $\langle f, g \rangle = \sum_\lambda a_\lambda \sum_\mu b_\mu \langle S_\lambda, S_\mu \rangle$. Since $\delta_{\lambda\mu} = \delta_{\mu\lambda}$, then $\langle S_\lambda, S_\mu \rangle = \langle S_\mu, S_\lambda \rangle$. By applying the symmetry of the Hall inner product on Schur functions and moving the linear combinations back into the inner product, we obtain the desired identity $\langle f, g \rangle = \langle g, f \rangle$.

Since the Schur functions are an orthonormal basis, then if $f = \sum_\lambda a_\lambda S_\lambda$ as before, we have $\langle f, f \rangle = \sum_\lambda a_\lambda^2$. As the a_λ are integers then clearly this is strictly positive if $f \neq 0$, and 0 if $f = 0$. □

Proposition 5.4. *The involution $\omega : \Lambda \to \Lambda$ defined by $e_k \mapsto h_k$ is an isometry. That is, $\langle \omega(f), \omega(g) \rangle = \langle f, g \rangle$ for all $f, g \in \Lambda$.*

Proof. By bilinearity, it suffices to verify that $\langle \omega(f), \omega(g) \rangle = \langle f, g \rangle$ for all f, g in a particular basis, say the basis given by Schur functions. By (3.3.3), we have $\omega(S_\lambda) = S_{\lambda'}$, and so for any two Schur functions S_λ, S_μ we have

$$\langle \omega(S_\lambda), \omega(S_\mu) \rangle = \langle S_{\lambda'}, S_{\mu'} \rangle = \delta_{\lambda'\mu'} = \delta_{\lambda\mu} = \langle S_\lambda, S_\mu \rangle$$

□

5.2 Skew Schur Polynomials

By Theorem 5.2, we have that the Schur polynomials S_λ form an orthonormal $\mathbb{Z}$-basis for the space Λ of symmetric functions. Consequently, any symmetric function f is uniquely defined by the values $\langle f, S_\lambda \rangle$, namely its Hall inner product with any given Schur polynomial. This allows us to define the following generalization of the Schur functions as follows.

Definition 5.2. Let λ, ν be partitions. We define the *skew Schur function* $S_{\nu/\lambda}$ to be the unique symmetric polynomial such that

$$\langle S_{\nu/\lambda}, S_\mu \rangle = \langle S_\nu, S_\lambda S_\mu \rangle \tag{5.2.1}$$

for each partition μ.

The following definition allows us to give an equivalent formulation of the skew Schur functions.

Definition 5.3. Let λ, μ, ν be partitions. We define the *Littlewood-Richardson coefficient* $c^{\nu}_{\lambda\mu}$ to be the coefficient of S_ν in the linear combination expressing the product $S_\lambda S_\mu$ in terms of Schur functions. That is, we define the Littlewood-Richardson coefficients by

$$S_\lambda S_\mu = \sum_\nu c^{\nu}_{\lambda\mu} S_\nu \tag{5.2.2}$$

Remark. Surprisingly, the Littlewood-Richardson coefficients are nonnegative for any choice of λ, μ and ν. In fact, there are a number of combinatorial algorithms which can be used to compute these coefficients. An investigation of some of these algorithms will be the focus of Section 5.4.

With the Littlewood-Richardson coefficients defined as above, the skew Schur function $S_{\nu/\lambda}$ can be equivalently expressed as

$$S_{\nu/\lambda} = \sum_\mu c^{\nu}_{\lambda\mu} S_\mu \tag{5.2.3}$$

From this formulation, it is clear that $S_{\nu/0} = S_\nu$ for any partition ν. Moreover, since S_λ is homogeneous of degree $|\lambda|$, we see that $c^{\nu}_{\lambda\mu} \neq 0$ only if $|\nu| = |\lambda| + |\mu|$. Consequently, $S_{\nu/\lambda}$ is homogeneous of degree $|\nu| - |\lambda|$, and is 0 if $|\nu| < |\lambda|$.

At present, the formula (5.2.3) is not particularly useful due to the *a priori* nonconstructive nature of the Littlewood-Richardson coefficients $c^{\nu}_{\lambda\mu}$. This will be remedied in Section 5.4, in which we describe combinatorial algorithms for determining these coefficients. A more immediately useful expression is the following generalization of the Jacobi-Trudi identity.

Proposition 5.5. *Let λ be a partition with $l(\lambda) \leq n$ and $\lambda_1 = l(\lambda') \leq m$. Then*

$$S_{\lambda/\mu} = \det(h_{\lambda_i - \mu_j - i + j})_{1 \leq i,j \leq n} = \det(e_{\lambda'_i - \mu'_j - i + j})_{1 \leq i,j \leq m} \tag{5.2.4}$$

with the usual convention that $e_k = h_k = 0$ if $k < 0$.

Proof. As with the proof of the original Jacobi-Trudi identity (3.3.2) for Schur functions, a proof of the right equality can be found in [6, p. 22-23]. We will now prove the left equality. Let $x = \{x_1, x_2, \ldots\}, y = \{y_1, y_2, \ldots\}$ be sets of independent variables. By the identities (5.2.3) and (5.2.2) we have

$$\sum_\lambda S_{\lambda/\mu}(x) S_\lambda(y) = \sum_{\lambda,\nu} c^{\lambda}_{\mu\nu} S_\nu(x) S_\lambda(y) = \sum_\nu S_\nu(x) S_\mu(y) S_\lambda(y)$$

By applying the Cauchy identities (3.4.1) and (5.1.1) to the rightmost expression we obtain

$$\sum_\lambda S_{\lambda/\mu}(x) S_\lambda(y) = S_\mu(y) \sum_\nu h_\nu(x) m_\nu(y)$$

Now suppose that the set of variables $y = \{y_1, \ldots, y_n\}$ is finite, so that the above sums are restricted to partitions λ, ν of length at most n. After multiplying by the Vandermonde determinant, we may rewrite the previous equation as

$$\sum_{\lambda} S_{\lambda/\mu}(x) A_{\lambda+\delta}(y) = \sum_{\nu} h_\nu(x) m_\nu(y) A_{\mu+\delta}(y) = \sum_{\alpha} h_\alpha(x) \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) y^{\alpha+\sigma(\mu+\delta)}$$

with the rightmost sum taken over all n-tuples $\alpha \in \mathbb{N}^n$ of nonnegative integers. Consequently, $S_{\lambda/\mu}(x)$ is equal to the coefficient of $y^{\lambda+\delta}$ in the rightmost expression. In particular, we have the identity

$$S_{\lambda/\mu} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) h_{\lambda+\delta-\sigma(\mu+\delta)} = \det(h_{\lambda_i - \mu_j - i + j})_{1 \le i,j \le n}$$

following the convention that $h_\alpha = 0$ if at least one of the components α_i is negative. □

As with Schur functions, we have the simple corollary that the involution ω acts on the skew Schur functions by conjugation of the corresponding skew diagram. That is,

$$\omega(S_{\lambda/\mu}) = S_{\lambda'/\mu'} \tag{5.2.5}$$

The Jacobi-Trudi identity (5.2.4) also provides some necessary conditions for a skew Schur polynomial to be nonzero, as well as a useful factorization of skew Schur functions.

Proposition 5.6. *If λ, μ are partitions, then $S_{\lambda/\mu} \neq 0$ only if $\mu \subset \lambda$.*

Proof. Suppose $\mu \not\subset \lambda$, so that $\lambda_r < \mu_r$ for some r. Then $\lambda_i \le \lambda_r < \mu_r \le \mu_j$ for $1 \le j \le r \le i \le n$. Thus $\lambda_i - \mu_j - i + j < 0$ for this range of (i, j), so $(h_{\lambda_i - \mu_j - i + j})_{1 \le i,j \le n}$ has an $(n - r + 1) \times r$ block of zeroes in the bottom left corner. Therefore its determinant vanishes, so $S_{\lambda/\mu} = 0$ by (5.2.4). □

Proposition 5.7. *The skew Schur polynomial $S_{\lambda/\mu}(x_1, \ldots, x_n)$ is nonzero only if $0 \le \lambda'_i - \mu'_i \le n$ for all $i \ge 1$.*

Proof. Suppose that $\lambda'_r - \mu'_r > n$ for some $r \ge 1$. Then since $e_k = 0$ for all $k > n$, the matrix $(e_{\lambda'_i - \mu'_j - i + j})_{1 \le i,j \le m}$ has a rectangular block of 0's in the top right corner with one vertex of the rectangle on the main diagonal. Consequently, the determinant is zero. □

Remark. The converse of Proposition 5.7 also holds by virtue of Theorem 5.12, for if each column of λ/μ has height at most n, then there exists a skew tableau on λ/μ with entries in $[n]$, and hence $S_{\lambda/\mu}(x_1, \ldots, x_n)$ will have a nonzero monomial term. For the present discussion (indeed, for the purposes of proving Theorem 5.12), only the previously stated implication is required.

For the next result, we first introduce a general fact about skew diagrams. Any skew diagram can be decomposed as a collection of skew diagrams which have no "gaps". More precisely, we have the following definition.

Definition 5.4. Let λ/μ be a skew diagram, and let ϕ be a subset of λ/μ. A *path* is a sequence of boxes $x_0, \ldots, x_n \in \phi$ such that x_{i-1} and x_i share a common side for each i, $1 \leq i \leq m$. We say that ϕ is *connected* if any two boxes in ϕ are joined by a path in ϕ. If ϕ is a maximally connected subset of λ/μ, we call ϕ a *connected component* of λ/μ.

Evidently, any skew diagram can be partitioned into components, and of course each component is itself a skew diagram. This leads to the following factorization result for skew Schur functions.

Proposition 5.8. *Suppose that λ/μ is a skew diagram with components θ_i (cf. Definition 5.4). Then the skew Schur function $S_{\lambda/\mu}$ is the product of the skew Schur functions of the components. That is, we have $S_{\lambda/\mu} = \prod_i S_{\theta_i}$.*

Proof. Suppose that $\lambda_{r+1} \leq \mu_r$ for some r so that the skew diagram λ/μ consists of two disjoint skew diagrams θ and ϕ. In this case, the matrix $(h_{\lambda_i-\mu_j-i+j})_{1\leq i,j\leq n}$ has the form $\left(\begin{smallmatrix} A & B \\ 0 & C \end{smallmatrix}\right)$, where A has size $r \times r$ and C has size $(n-r) \times (n-r)$. The determinant in this case is $\det(A)\det(C)$, and A and C correspond to the respective Jacobi-Trudi matrices for the components θ and ϕ. The conclusion follows by induction. $\square$

We will now concern ourselves with expressions for skew Schur polynomials in multiple sets of independent variables, as reduction of these results to important special cases will provide us with highly useful combinatorial expressions for skew Schur functions. Before proceeding to the general case, we first consider expressions for Schur functions in two sets of independent variables, given by the following proposition.

Proposition 5.9. *Let $x = \{x_1, x_2, \ldots\}, y = \{y_1, y_2, \ldots\}$ be sets of independent variables, and denote by $S_\lambda(x, y)$ the Schur function in variables $(x_1, x_2, \ldots, y_1, y_2, \ldots)$ corresponding to the partition λ. Then*

$$S_\lambda(x,y) = \sum_\mu S_{\lambda/\mu}(x)S_\mu(y) = \sum_{\mu,\nu} c^\lambda_{\mu\nu} S_\nu(x)S_\mu(y) \tag{5.2.6}$$

Proof. The right equality follows from (5.2.3). To prove the left equality, first introduce a third set

$z = \{z_1, z_2, \ldots\}$ of independent variables. We have

$$\begin{aligned}
\sum_{\lambda,\mu} S_{\lambda/\mu}(x)S_\lambda(z)S_\mu(y) &= \sum_{\lambda,\mu,\nu} c^\lambda_{\mu\nu} S_\nu(x)S_\lambda(z)S_\mu(y) \\
&= \sum_{\mu,\nu} S_\nu(x)S_\mu(y)S_\mu(z)S_\nu(z) \\
&= \left(\sum_\mu S_\mu(y)S_\mu(z)\right)\left(\sum_\nu S_\nu(x)S_\nu(z)\right) \\
&= \left(\prod_{i,k}(1-x_iz_k)^{-1}\right)\left(\prod_{j,k}(1-y_jz_k)^{-1}\right) \\
&= \sum_\lambda S_\lambda(x,y)S_\lambda(z)
\end{aligned}$$

The left equality in (5.2.6) is obtained by comparing the coefficients of $S_\lambda(z)$ in the initial and final expressions. ☐

We may generalize the previous result quite easily to obtain an expression for an arbitrary skew Schur function in two sets of independent variables.

Proposition 5.10. *Let $x = \{x_1, x_2, \ldots\}, y = \{y_1, y_2, \ldots\}$ be sets of independent variables, and denote by $S_{\lambda/\mu}(x, y)$ the skew Schur function in variables $(x_1, x_2, \ldots, y_1, y_2, \ldots)$ corresponding to the skew diagram λ/μ. Then*

$$S_{\lambda/\mu}(x,y) = \sum_\nu S_{\lambda/\nu}(x)S_{\nu/\mu}(y) \tag{5.2.7}$$

where the sum is taken over all partitions ν such that $\mu \subset \nu \subset \lambda$.

Proof. Introduce a third set $z = \{z_1, z_2, \ldots\}$ of independent variables. By the previous proposition we have

$$\sum_\mu S_{\lambda/\mu}(x,y)S_\mu(z) = S_\lambda(x,y,z) = \sum_\nu S_{\lambda/\nu}(x)S_\nu(y,z) = \sum_{\mu,\nu} S_{\lambda/\nu}(x)S_{\nu/\mu}(y)S_\mu(z)$$

The conclusion follows by comparing coefficients of $S_\mu(z)$. ☐

If $x^{(1)}, \ldots, x^{(n)}$ are n sets of independent variables, then we may generalize Proposition 5.10 inductively as follows.

Proposition 5.11. *Let $x^{(1)}, \ldots, x^{(n)}$ be n sets of independent variables. Then*

$$S_{\lambda/\mu}(x^{(1)}, \ldots, x^{(n)}) = \sum_{(\nu)} \prod_{i=1}^n S_{\nu^{(i)}/\nu^{(i-1)}}(x^{(i)}) \tag{5.2.8}$$

where the sum is taken over all sequences $(\nu) = (\nu^{(0)}, \nu^{(1)}, \ldots, \nu^{(n)})$ of partitions such that

$$\mu = \nu^{(0)} \subset \nu^{(1)} \subset \cdots \subset \nu^{(n)} = \lambda$$

5.3 Schur Polynomials as Sums of Tableau Monomials

An important special case of Proposition 5.11 is obtained by letting each set of variables $x^{(i)}$ contain a single variable x_i. In doing so, we reduce to the skew Schur polynomial $S_{\lambda/\mu}(x_1, \dots, x_n)$ and obtain a novel expression for this polynomial without any reference to Littlewood-Richardson coefficients. This is summarized in the following theorem, which establishes a fundamental connection between (skew) Schur functions and (skew) tableaux.

Theorem 5.12. *Let λ and μ be partitions and let $\mathcal{T}(\lambda/\mu)$ denote the set of skew tableaux on λ/μ with positive integer entries. Then the skew Schur function λ/μ admits the combinatorial expression*

$$S_{\lambda/\mu} = \sum_{S \in \mathcal{T}(\lambda/\mu)} x^S \tag{5.3.1}$$

where x^S denotes the skew tableau monomial x^α, where α is the content of the skew tableau S. If we let $\mathcal{T}_{[n]}(\lambda/\mu)$ denote the set of skew tableaux on λ/μ with entries in $[n]$, then we have the following identity for skew Schur polynomials in n variables:

$$S_{\lambda/\mu}(x_1, \dots, x_n) = \sum_{S \in \mathcal{T}_{[n]}(\lambda/\mu)} x^S \tag{5.3.2}$$

Proof. We will first prove (5.3.2). As we remarked at the beginning of this section, the proof of this result follows by reducing Proposition 5.11 in the special case of n sets of variables $x^{(i)}$ each containing a single variable x_i. In this way the lefthand side of (5.2.8) becomes the skew Schur polynomial $S_{\lambda/\mu}(x_1, \dots, x_n)$.

Consider a single term $S_{\nu^{(i)}/\nu^{(i-1)}}(x_i)$ appearing as a factor in one of the summands of the righthand side of (5.2.8). By Proposition 5.7 with $n = 1$, we see that $S_{\nu^{(i)}/\nu^{(i-1)}}(x_i) \neq 0$ only if the skew diagram $\nu^{(i)}/\nu^{(i-1)}$ is a horizontal strip. If this is the case, then the connected components of $\nu^{(i)}/\nu^{(i-1)}$ are horizontal rows (k_j) of length k_j, whose individual skew Schur functions are given by $S_{(k_j)}(x_i) = h_{k_j}(x_i) = x_i^{k_j}$. By Proposition 5.8 the resulting skew Schur polynomial $S_{\nu^{(i)}/\nu^{(i-1)}}(x_i)$ is simply $x_i^{|\nu^{(i)}/\nu^{(i-1)}|}$. In this way, the righthand side of (5.2.8) becomes a sum of monomials $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, where $\alpha_i = |\nu^{(i)}/\nu^{(i-1)}|$. The sum in this case is taken over all sequences (ν) of nested partitions $\mu = \nu^{(0)} \subset \nu^{(1)} \subset \cdots \subset \nu^{(n)} = \lambda$ such that each diagram $\nu^{(i)}/\nu^{(i-1)}$ is a horizontal strip.

An equivalent way of comprehending the sequences (ν) that arise in this expression is by interpreting each sequence as a skew tableau on λ/μ with entries in $[n]$. Indeed, given a skew tableau S on λ/μ with entries in n, we may recursively define a suitable sequence of partitions (ν) by letting $\nu^{(0)} = \mu$ and defining $\nu^{(i)}$ to be the diagram $\nu^{(i-1)}$ together with all the boxes of S containing the entry i. Conversely, given such a sequence of $n+1$ nested partitions such that each skew diagram $\nu^{(i)}/\nu^{(i-1)}$ is a horizontal strip, we may recover the corresponding skew tableau. We do so by simply filling the boxes of $\nu^{(i)}/\nu^{(i-1)}$ with the entry i. By this correspondence, the products being summed in (5.2.8) become the monomials x^α, where α is the content of the corresponding skew tableau on λ/μ. In this way, we obtain (5.3.2).

The symmetric polynomial identity (5.3.2) in fact implies the symmetric function identity (5.3.1). This follows because the identity holds in Λ_n for each $n \geq 1$. In particular, the identity, which is homogeneous of degree $|\lambda/\mu|$, holds if $n = |\lambda/\mu|$, so we may apply Lemma 2.7. □

By letting $\mu = 0$, we specialize to Schur functions and obtain the immediate corollary.

Corollary 5.13. *Let λ be a partition and let $\mathcal{T}(\lambda)$ denote the set of tableaux on λ. Then we have the identity*

$$S_\lambda = \sum_{T \in \mathcal{T}(\lambda)} x^T \tag{5.3.3}$$

where x^T is the tableau monomial whose exponents are given by the content of T. If $\mathcal{T}_{[n]}(\lambda)$ denotes the set of tableaux on λ with entries in $[n]$, then the following identity holds for Schur functions in n variables:

$$S_\lambda(x_1, \ldots, x_n) = \sum_{T \in \mathcal{T}_{[n]}(\lambda)} x^T \tag{5.3.4}$$

Remark. By considering diagrams λ consisting of either a single column or a single row, we specialize further to elementary and complete homogeneous symmetric functions and retrieve Proposition 4.8.

This correspondence between Schur functions and tableaux is one of the fundamental results of algebraic combinatorics; in fact, many authors use the identity (5.3.3) as the definition of the Schur functions. With this crucial result, we are now equipped to begin a more in depth study of the combinatorial properties of Schur functions.

In order to establish combinatorial results related to Schur polynomials, we will typically work in the tableau ring $R_{[n]}$ (cf. Section 4.4). Indeed, in the tableau ring $R_{[n]}$, we may define $s_\lambda[n]$ to be the sum of all tableaux $T \in \mathcal{T}_{[n]}(\lambda)$. By virtue of (5.3.4) the homomorphism $T \mapsto x^T$ acts on the $s_\lambda[n] \in R_{[n]}$ by mapping it to the corresponding Schur polynomial in n variables, namely $s_\lambda[n] \mapsto S_\lambda(x_1, \ldots, x_n)$.

With this in mind, we will now consider the problem of expressing the product of two Schur functions as a $\mathbb{Z}$-linear combination of Schur functions. The central result in this vein is the Littlewood-Richardson Rule, which provides combinatorial expressions for the elusive Littlewood-Richardson coefficients $c^\nu_{\lambda\mu}$; this will be our focus in the following section. For the time being, we will consider an important special case. In particular, we will concern ourselves with expressions for a product of two Schur functions, where one of these Schur functions is an elementary or complete homogeneous symmetric function. This expression is given by the *Pieri Formulas*, which are in fact little more than an elaborate restatement of the row-bumping lemma.

Proposition 5.14 (Pieri Formulas)**.** *The following identity holds in the ring of symmetric functions for all $k \geq 0$:*

$$S_\lambda h_k = \sum_\mu S_\mu \tag{5.3.5}$$

where the sum is taken over all diagrams $\mu \supset \lambda$ such that μ/λ is a horizontal k-strip. Likewise, for each $k \geq 0$ we have

$$S_\lambda e_k = \sum_\mu S_\mu \tag{5.3.6}$$

where the sum is taken over all diagrams $\mu \supset \lambda$ such that μ/λ is a vertical k-strip.

Proof. We will begin by proving both formulas in the tableau ring $R_{[n]}$ for arbitrary $n \geq 1$. In this setting, the homogeneous and elementary symmetric polynomials can be realized as special cases of the $s_\lambda[n]$ defined above. In particular, by Proposition 4.8 we have $s_{(k)}[n] \mapsto h_k(x_1, \dots, x_n)$ and $s_{(1^k)}[n] \mapsto e_k(x_1, \dots, x_n)$ under the canonical homomorphism $T \mapsto x^T$. With this in mind, we will prove that the following identities hold in $R_{[n]}$ for each $n \geq 1$:

$$s_\lambda s_{(k)} = \sum_\mu s_\mu \tag{5.3.5'}$$

summed over all diagrams $\mu \supset \lambda$ such that μ/λ is a horizontal k-strip, and

$$s_\lambda s_{(1^k)} = \sum_\mu s_\mu \tag{5.3.6'}$$

summed over all diagrams $\mu \supset \lambda$ such that μ/λ is a vertical k-strip.

By definition, s_λ is the sum of all tableaux on λ with entries in $[n]$, and likewise $s_{(k)}$ is the sum of all tableaux on (k) with entries in $[n]$. As the diagram (k) is a single row of length k, all such tableaux on (k) correspond to weakly increasing sequences in $[n]$ of length k. By definition of the tableau ring, the product $s_\lambda s_{(k)}$ is simply the sum of all product tableaux $T_\lambda T_{(k)}$ where $T_\lambda \in \mathcal{T}_{[n]}(\lambda)$ and $T_{(k)} \in \mathcal{T}_{[n]}(k)$.

If we consider the tableau product via row-insertion, then we may apply Proposition 4.2. This proposition asserts that all tableaux $T_\lambda T_{(k)}$ appearing in the sum have shape $\mu \sup \lambda$ such that $|\mu| = |\lambda| + k$ and no two boxes in μ/λ lie in the same column. In other words, all such tableaux have shape μ such that μ/λ is a horizontal k-strip. Since the converse of the proposition also holds, then we see that all such products $T_\lambda T_{(k)}$ appear if we sum over all such μ. Rearranging the sum modulo shape yields (5.3.5').

The proof of (5.3.6') is similar. In particular, the key observation is that a tableau on (1^k) with entries in $[n]$ corresponds to a strictly increasing sequence in $[n]$ of length k. Since the entries of a tableau on (1^k) are row-inserted into T_λ in reverse (*i.e.* strictly decreasing) order, then Proposition 4.2 may be applied to give the desired conclusion.

Applying the canonical homomorphism $T \mapsto x^T$ establishes the corresponding results for symmetric polynomials in n variables for each $n \geq 1$:

$$S_\lambda(x_1, \ldots, x_n) h_k(x_1, \ldots, x_n) = \sum_\mu S_\mu(x_1, \ldots, x_n) \tag{5.3.5''}$$

$$S_\lambda(x_1, \ldots, x_n) e_k(x_1, \ldots, x_n) = \sum_\mu S_\mu(x_1, \ldots, x_n) \tag{5.3.6''}$$

where the sums are taken over all diagrams μ of the forms previously specified. Since these results are valid for arbitrarily many variables n, then we may apply Lemma 2.7 to obtain the corresponding identities for symmetric functions. □

Let $\lambda = (\lambda_1 \geq \ldots \geq \lambda_n \geq 0)$. Then a diagram μ appears in the sum in (5.3.5) if and only if it satisfies the following algebraic conditions:

1. $|\mu| = |\lambda| + k$, $l(\mu) \leq n + 1$
2. $\mu_1 \geq \lambda_1 \geq \mu_2 \geq \cdots \geq \mu_n \geq \lambda_n \geq \mu_{n+1} \geq \mu_{n+2} = 0$

A diagram μ appears in (5.3.6) if and only if it satisfies

1. $|\mu| = |\lambda| + k$, $l(\mu) \leq n + k$
2. $\lambda_i \leq \mu_i \leq \lambda_i + 1$ for each i, $1 \leq i \leq n + k$

Remark. As we remarked earlier, there are many different ways to prove the equivalence of the determinantal and combinatorial definitions of the Schur functions. We have chosen to follow Macdonald's approach here because the associated discussion of the Hall inner product and skew Schur functions motivates similar discussions in Section 6. An elementary proof of the equivalence was first published by Proctor in [8]; a more detailed presentation of the same proof can be found in [1, p. 161-169]. Yet another alternate proof, which can be found in [9, §4.5, 4.6], proves the Jacobi-Trudi identity directly from each of the two definitions and uses the fact that the two definitions agree for the homogeneous symmetric polynomials.

5.4 Littlewood-Richardson Rule

The purpose of the current section will be to establish several explicit combinatorial expressions for the implicitly defined Littlewood-Richardson coefficients $c^\nu_{\lambda\mu}$ introduced in Definition 5.3. This will provide us with a relatively simple means to express a product of any two Schur functions as a $\mathbb{Z}$-linear combination of Schur functions.

Recall that we define the Littlewood-Richardson coefficient $c^\nu_{\lambda\mu}$ by setting $S_\lambda S_\mu := \sum_\nu c^\nu_{\lambda\mu} S_\nu$. The identity (5.3.3) provides us with a more nuanced way of looking at this equation. In particular, we have

$$S_\lambda S_\mu = \left(\sum_{T_\lambda \in \mathcal{T}(\lambda)} x^{T_\lambda} \right) \left(\sum_{T_\mu \in \mathcal{T}(\mu)} x^{T_\mu} \right) = \sum_{\substack{T_\lambda \in \mathcal{T}(\lambda) \\ T_\mu \in \mathcal{T}(\mu)}} x^{T_\lambda \cdot T_\mu}$$

In order to turn the rightmost expression into a sum over Schur functions S_ν, we simply rearrange the sum so that all monomials $x^{T_\lambda \cdot T_\mu}$ with $T_\lambda \cdot T_\mu \in \mathcal{T}(\nu)$ are grouped together. Consequently, determining the value of the Littlewood-Richardson coefficient $c^\nu_{\lambda\mu}$ amounts to counting the number of ways a tableau T with shape ν factors as a product $T = T_\lambda \cdot T_\mu$ of tableaux on λ and μ, respectively. Central to this investigation is the following lemma.

Lemma 5.15. *Suppose $w = \left(\begin{smallmatrix} u_1 & \cdots & u_m \\ v_1 & \cdots & v_m \end{smallmatrix}\right)$ is a lexicographic array, corresponding by the RSK correspondence to the pair (P, Q) of tableaux. Let T be any tableau and perform the row-insertions*

$$(\cdots((T \leftarrow v_1) \leftarrow v_2) \leftarrow \cdots) \leftarrow v_m$$

and place $u_1, \ldots, u_m$ successively in the new boxes. Then the entries $u_1, \ldots, u_m$ form a skew tableau S whose rectification is Q.

Proof. Fix a tableau T and take a tableau T_0 with the same shape as T, using an alphabet whose letters are strictly smaller than the letters u_i in S (*e.g.* using negative integers). The pair (T, T_0) corresponds to some lexicographic array $\left(\begin{smallmatrix} s_1 & \cdots & s_n \\ t_1 & \cdots & t_n \end{smallmatrix}\right)$. The lexicographic array $\left(\begin{smallmatrix} s_1 & \cdots & s_n & u_1 & \cdots & u_m \\ t_1 & \cdots & t_n & v_1 & \cdots & v_m \end{smallmatrix}\right)$ corresponds in turn to a pair $(T \cdot P, V)$ of tableaux. Here $T \cdot P$ is the result of the successive row insertions of $v_1, \ldots, v_m$ into T, and V is a tableau whose entries $s_1, \ldots, s_n$ make up T_0 and whose entries $u_1, \ldots, u_m$ make up S.

Now invert this concatenated array and put the result in lexicographic order. The terms $\left(\begin{smallmatrix} v_i \\ u_i \end{smallmatrix}\right)$ will occur in this array in lexicographic order, with terms $\left(\begin{smallmatrix} t_j \\ s_j \end{smallmatrix}\right)$ interspersed (but also in lexicographic order). By the Symmetry Theorem 4.11, this inverted array corresponds to the pair $(V, T \cdot P)$ and the resulting array with the columns $\left(\begin{smallmatrix} t_j \\ s_j \end{smallmatrix}\right)$ removed corresponds to the pair (Q, P). Consequently, the word on the bottom row of this array is Knuth-equivalent to the row word $w(V)$ and removing the s_j's from this word yields a word which is Knuth-equivalent to $w(Q)$. Of course, removing the s_j's from $w(V)$ leaves precisely $w(S)$. Consequently $w(S)$ and $w(Q)$ are Knuth-equivalent, so $\mathrm{Rect}(S) = Q$. □

Let λ, μ, and ν be partitions. In light of the discussion preceding the lemma, we would like to consider the number of ways that a given tableau V with shape ν can be written as a product $T \cdot U$ of a tableau T on λ and a tableau U on μ. Clearly this will be 0 unless $|\nu| = |\lambda| + |\mu|$ and $\lambda \subset \nu$. One equivalent way of realizing this quantity is by considering the tableau product via rectification (4.4.2). Evidently, the number of ways to factor $V = T \cdot U$ as above is equal to the number of skew tableaux on the shape $\lambda * \mu$ whose rectification is V. Remarkably, this also happens to be the number of skew tableaux with shape ν/λ whose rectification is a given tableau on μ. To prove this fact, we first introduce the following notation.

Notation. For any tableau U_0 with shape μ, define $\mathcal{S}(\nu/\lambda, U_0)$ to be the set of all skew tableaux on ν/λ whose rectification is U_0. For any tableau V_0 on ν, define $\mathcal{T}(\lambda, \mu, V_0)$ to be the set of all pairs of tableaux $(T, U) \in \mathcal{T}(\lambda) \times \mathcal{T}(\mu)$ such that $T \cdot U = V_0$.

Proposition 5.16. *Let U_0 be a tableau on μ, V_0 a tableau on ν. Then there is a canonical bijective correspondence*

$$\mathcal{S}(\nu/\lambda, U_0) \longleftrightarrow \mathcal{T}(\lambda, \mu, V_0)$$

Proof. Fix $(T, U) \in \mathcal{T}(\lambda, \mu, V_0)$ and fix a tableau T_0 on λ whose entries are all strictly less than the entries of U_0 (using negative entries if necessary). Consider the unique lexicographic arrays $\left(\begin{smallmatrix} t_1 & \cdots & t_n \\ x_1 & \cdots & x_n \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} u_1 & \cdots & u_m \\ v_1 & \cdots & v_m \end{smallmatrix}\right)$ corresponding to the pairs (T, T_0) and (U, U_0), respectively, by the RSK correspondence. By concatenating these arrays, we obtain the unique lexicographic array $\left(\begin{smallmatrix} t_1 & \cdots & t_n & u_1 & \cdots & u_m \\ x_1 & \cdots & x_n & v_1 & \cdots & v_m \end{smallmatrix}\right)$. It follows from the tableau multiplication procedure via row-insertion that this concatenated array corresponds to a unique pair (V_0, V'). In particular, V' is a tableau on ν obtained from T_0 by placing the entries $u_1, \ldots, u_m$ into the new boxes of the successive row-insertions $T \leftarrow v_1 \leftarrow \cdots \leftarrow v_m$. We may uniquely obtain a skew tableau S on ν/λ by removing the boxes of T_0 from the tableau V'. Since S was constructed according to the hypotheses of Lemma 5.15, we have $\mathrm{Rect}(S) = U_0$ and so $S \in \mathcal{S}(\nu/\lambda, U_0)$. Moreover, the skew tableau S we obtain is independent of the original choice of T_0.

Evidently this procedure is invertible. Indeed, given the skew tableau S on ν/λ that we obtain, fix an arbitrary tableau T' on λ whose entries are all strictly less than the entries of S. Define a tableau $(T')_S$ on ν to be equal to T' on λ and equal to S on ν/λ. Then the pair $(V_0, (T')_S)$ will correspond to a unique lexicographic array $\left(\begin{smallmatrix} t'_1 & \cdots & t'_n & u_1 & \cdots & u_m \\ x_1 & \cdots & x_n & v_1 & \cdots & v_m \end{smallmatrix}\right)$, where the t'_i are the entries of t' and the u_i are the entries of S. We may split this into two smaller arrays $\left(\begin{smallmatrix} t'_1 & \cdots & t'_n \\ x_1 & \cdots & x_n \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} u_1 & \cdots & u_m \\ v_1 & \cdots & v_m \end{smallmatrix}\right)$. Since the bottom row of the concatenated array is precisely the bottom row of the concatenated array used to obtain S, it is clear from the row-insertion involved in the RSK correspondence that these arrays correspond to the pairs (T, T') and (U, U_0), where T and U are the original tableaux with $(T, U) \in \mathcal{T}(\lambda, \mu, V_0)$. $\square$

An important consequence of this correspondence is the following. Suppose that U_0 and U_1 are tableaux on μ, and let V_0 be a tableau on ν. By the proposition, $\#\mathcal{S}(\nu/\lambda, U_0) = \#\mathcal{T}(\lambda, \mu, V_0)$ and $\#\mathcal{S}(\nu/\lambda, U_1) = \#\mathcal{T}(\lambda, \mu, V_0)$. Consequently, $\#\mathcal{S}(\nu/\lambda, U_0) = \#\mathcal{S}(\nu/\lambda, U_1)$ for any two tableaux U_0, U_1 on μ. Likewise, $\#\mathcal{T}(\lambda, \mu, V_0) = \#\mathcal{T}(\lambda, \mu, V_1)$ for any choice of tableaux V_0, V_1 on ν. By the discussion preceding Lemma 5.15, the number of times a monomial x^T with T a tableau on ν appears in the product $S_\lambda S_\mu$ is precisely $\#\mathcal{T}(\lambda, \mu, T)$ times. Since this value is independent of the tableau T on ν, we may regroup the monomial terms of the sum into Schur functions S_ν, and the coefficient of S_ν in this sum will be $\#\mathcal{T}(\lambda, \mu, T)$. By comparing with the definition of the Littlewood-Richardson coefficients, this proves the following proposition.

Proposition 5.17. *Let λ, μ, and ν be partitions and let U_0 and V_0 be tableaux on μ and ν, respectively. Then*

$$c^{\nu}_{\lambda\mu} = \#\mathcal{S}(\nu/\lambda, U_0) = \#\mathcal{T}(\lambda, \mu, V_0) \tag{5.4.1}$$

That is, the Littlewood-Richardson coefficient $c^{\nu}_{\lambda\mu}$ is equal to the number of ways that V_0 can be written as a product $V_0 = T \cdot U$ with $T \in \mathcal{T}(\lambda)$ and $U \in \mathcal{T}(\mu)$, which is in turn equal to the number of skew tableaux on ν/λ whose rectification is U_0.

Without much additional work, we may extend the proposition as follows.

Proposition 5.18. *The following sets all have cardinality $c^{\nu}_{\lambda\mu}$:*

1. *$\mathcal{S}(\nu/\mu, T_0)$ for any tableau T_0 on λ,*
2. *$\mathcal{T}(\mu, \lambda, V_0)$ for any tableau V_0 on ν,*
3. *$\mathcal{S}(\nu'/\lambda', U'_0)$ for any tableau U'_0 on the conjugate diagram μ',*
4. *$\mathcal{T}(\lambda', \mu', V'_0)$ for any tableau V'_0 on the conjugate diagram ν', and*
5. *$\mathcal{S}(\lambda * \mu, V_0)$ for any tableau V_0 on ν.*

Proof. The cardinality of (5) is equal to $c^{\nu}_{\lambda\mu}$ by the discussion preceding Proposition 5.17. Taking U_0 to be a standard tableau on μ, there is a natural bijection between $\mathcal{S}(\nu/\lambda, U_0)$ and $\mathcal{S}(\nu'/\lambda', U'_0)$ by taking transposes. This establishes that the cardinality of (3) is $c^{\nu}_{\lambda\mu}$, and consequently (4) has the same cardinality by Proposition 5.17.

If we let V_0 be a standard tableau, then transposition is again a natural bijection between $\mathcal{S}(\lambda * \mu, V_0)$ and $\mathcal{S}((\lambda * \mu)', V'_0)$. Since $(\lambda * \mu)' = \mu' * \lambda'$, then $\mathcal{S}(\mu' * \lambda', V'_0)$ has cardinality $c^{\nu}_{\lambda\mu}$ for any tableau V'_0 on ν'. Using the fact that (4) and (5) have equal cardinalities, we have

$$c^{\nu}_{\lambda\mu} = \#\mathcal{S}(\mu' * \lambda', V'_0) = \#\mathcal{T}(\mu'', \lambda'', V''_0) = \#\mathcal{T}(\mu, \lambda, V_0)$$

which establishes the cardinality of (2). Finally, (1) and (2) have the same cardinality by Proposition 5.17. □

Remark. In particular, the proposition implies that $c^{\nu}_{\lambda\mu} = c^{\nu}_{\mu\lambda}$ and $c^{\nu}_{\lambda\mu} = c^{\nu'}_{\lambda'\mu'}$. From the former identity we obtain the additional constraint that $c^{\nu}_{\lambda\mu} = 0$ unless $\lambda \subset \nu$ and $\mu \subset \nu$.

There is a great diversity of combinatorial expressions for the Littlewood-Richardson coefficients beyond those listed in Propositions 5.17 and 5.18. The next such expression we will look at, which is often called the "Littlewood-Richardson Rule", requires a new notion, namely the idea of a *reverse lattice word*.

Definition 5.5. Let $w = x_1 \cdots x_r$ be a word of positive integers. We call w a *reverse lattice* (or *Yamanouchi*) *word* if for all integers $s < r$, $k \geq 1$, the integer k appears at least as many times as the integer $k+1$ in the sequence $x_r, x_{r-1}, \ldots . x_{r-s}$.

Example. 2 1 3 2 1 2 1 is a reverse lattice word, but 1 2 3 2 1 2 1 is not since the last 6 letters contain three 2's but only two 1's.

One of the important consequences of this definition is that the property of being a reverse lattice word is preserved under Knuth-equivalence.

Lemma 5.19. *If $w \equiv w'$ are Knuth-equivalent words, then w is a reverse lattice word if and only if w' is a reverse lattice word.*

Proof. It suffices to check that this property is preserved by the elementary Knuth transformations. First consider the elementary Knuth transformation

$$w = uxzyv \mapsto uzxyv = w' \qquad (x \leq y < z)$$

We are concerned with possible changes in the number of consecutive integers k and $k+1$ reading the words from right to left. If $x < y < z$ there is no change, so it suffices to check the case where $x = y = k$ and $z = k+1$. For w or w' to be a reverse lattice word, the number of k's appearing in v must be at least as large as the number of $(k+1)$'s appearing in v. In this case, both $xzyv$ and $zxyv$ are reverse lattice words. Now consider the other elementary Knuth transformation

$$w = uyxzv \mapsto uyzxv = w' \qquad (x < y \leq z)$$

The only nontrivial case to consider is when $x = k$ and $y = z = k+1$. In this case, w or w' can only be reverse lattice words if the number of k's in v is strictly greater than the number of $(k+1)$'s. If this is the case, then $yxzv$ and $yzxv$ will each have at least as many k's as $(k+1)$'s, which completes the proof. $\square$

Definition 5.6. Let S be a skew tableau. We call S a *Littlewood-Richardson skew tableau* if its row word $w(S)$ is a reverse lattice word.

The definitions for reverse lattice words and Littlewood-Richardson skew tableaux can be motivated in part by the following lemma. In particular, the rectification of a Littlewood-Richardson skew tableau has a rather simple form.

Notation. Let μ be a partition. We define the tableau $U(\mu)$ to be the tableau on μ such that all entries of the i-th row are equal to i. For example, if $\mu = (4, 4, 3, 2)$, we have

$$U(\mu) = \begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 1 \\ \hline 2 & 2 & 2 & 2 \\ \hline 3 & 3 & 3 \\ \cline{1-3} 4 & 4 \\ \cline{1-2} \end{array}$$

Lemma 5.20. *A skew tableau S is a Littlewood-Richardson skew tableau with content μ if and only if $Rect(S) = U(\mu)$.*

Proof. We will first prove that $U(\mu)$ is the unique Littlewood-Richardson tableau with content μ. Indeed, suppose that U is a Littlewood-Richardson tableau of content μ. Since $w(U)$ is a reverse lattice word, then its final letter must be a 1. Since the final letter of the row word corresponds to the rightmost entry of the first row, we see that this entry must be a 1. Since U is a tableau, then consequently each entry of the first row must be a 1. Since U is a tableau, then each entry of the second row must be at least 2. If we again consider the rightmost entry of the second row, we see that this entry corresponds to the rightmost letter in $w(U)$ which is not necessarily a 1. Since $w(U)$ is a reverse lattice word, then this letter is at most 2. Thus the rightmost entry of the second row of U is a 2, so each entry of the second row is a 2. The same argument applies for each subsequent row.

Now that we have proven that $U(\mu)$ is the unique Littlewood-Richardson tableau with content μ, it suffices to show that a skew tableau is a Littlewood-Richardson skew tableau if and only if its rectification is a Littlewood-Richardson tableau. Indeed, this is the case by virtue of Lemma 5.19 and the fact that sliding preserves Knuth-equivalence (Proposition 4.5). $\square$

With this result, we may easily derive the Littlewood-Richardson Rule.

Proposition 5.21 (Littlewood-Richardson Rule)**.** *The Littlewood-Richardson coefficient $c^{\nu}_{\lambda\mu}$ is equal to the number of Littlewood-Richardson skew tableaux with shape ν/λ and content μ.*

Proof. Since $U(\mu)$ is a tableau on μ, then by Proposition 5.17, we have $c^{\nu}_{\lambda\mu} = \#\mathcal{S}(\nu/\lambda, U(\mu))$. Unpacking this equation, it states that the Littlewood-Richardson coefficient is equal to the number of skew tableaux on ν/λ whose rectification is $U(\mu)$. By the previous lemma, all such skew tableau on ν/λ are Littlewood-Richardson skew tableau on ν/λ with content μ, and conversely. Hence the conclusion follows. $\square$

6 Generalizations of Schur Polynomials

Although the symmetric polynomials that we have been considering up to this point have typically had coefficients in $\mathbb{Z}$ or $\mathbb{Q}$, recall that we may very well consider symmetric polynomials over any ring of coefficients. In the following sections, we will study generalizations of the Schur polynomials whose coefficients are rational functions in one or two additional parameters. These are the Jack and Macdonald polynomials, respectively, and we will devote most of our attention to the latter. These polynomials arise naturally in a diverse range of contexts. For instance, the Macdonald polynomials appear in algebraic geometry [3], as well as mathematical physics where they are eigenfunctions of certain differential operators related to the quantum n-body problem [4]. As in our treatment of the Schur polynomials, we will be primarily concerned with the algebraic properties of these polynomials, as well as related combinatorial expressions.

We will typically follow the development of this theory presented by Macdonald in Ch. VI of [6]. Section 6.1 is adapted from [6, Ch. VI, §1,2]. The construction of the Macdonald polynomials provided in Section 6.2 follows the more condensed treatment of [5] while the basic algebraic and duality results that follow can be found in [6, Ch. VI, §4,5]. Sections 6.3 and 6.4, which will focus on deriving combinatorial expressions for the Macdonald polnomials, are adapted from §6 and §7 of [6, Ch. VI], respectively.

6.1 Generalized Hall Inner Product

In this section we will revisit the Hall inner product of Section 5.1, which we defined by requiring the symmetric function bases given by the h_λ and the m_λ to be dual to one another. As per the remark following Theorem 5.2, with this inner product the power sum symmetric functions p_λ form an orthogonal $\mathbb{Q}$-basis of the space $\Lambda_{\mathbb{Q}}$ of symmetric functions with rational coefficients. In fact, by virtue of (4.6) and (4.7) in [6, Ch. I], the Hall inner product can equivalently be defined by requiring that

$$\langle p_\lambda, p_\mu \rangle = \delta_{\lambda\mu} z_\lambda := \delta_{\lambda\mu} \prod_{i \geq 1} i^{m_i} \cdot m_i! \tag{6.1.1}$$

In the definition of z_λ[2], m_i denotes the multiplicity of i in $\lambda = (1^{m_1} 2^{m_2} 3^{m_3} \cdots)$. The Schur functions are in fact characterized uniquely by the following two properties, as per the first remark in [5, p. 139].

(a) The Schur functions S_λ are pairwise orthogonal relative to the Hall inner product defined by (6.1.1), namely $\langle S_\lambda, S_\mu \rangle \neq 0$ iff $\lambda = \mu$, and

[2] The value z_λ admits a combinatorial interpretation. In particular, if $\lambda \vdash n$, then z_λ is the number of permutations in S_n which commute with a given permutation $\sigma_\lambda \in S_n$ of cycle type λ [10, p. 299]. For our purposes, however, it just suffices to recognize that z_λ is a scalar parametrized by λ.

(b) Each Schur function S_λ can be written (uniquely) in the form

$$S_\lambda = m_\lambda + \sum_{\mu<\lambda} K_{\lambda\mu} m_\mu \tag{6.1.2}$$

for a suitable choice of coefficients $K_{\lambda\mu}$. Here (and in what follows, unless otherwise stated) $<$ refers to the *dominance partial order* on partitions, where we say that $\mu \leq \lambda$ if and only if $\mu_1 + \mu_2 + \cdots + \mu_i \leq \lambda_1 + \lambda_2 + \cdots + \lambda_i$ for each $i \geq 1$.

Remark. The coefficients $K_{\lambda\mu}$ appearing in (6.1.2) are called the *Kostka numbers*. Given the combinatorial formula (5.3.3) for the Schur functions, it is easy to see that the Kostka number $K_{\lambda\mu}$ is equal to the number of tableaux on λ with content μ. A brief discussion of the Kostka numbers can be found in [2, p. 25-26]. Among the results included is the fact that when $|\lambda| = |\mu|$, $K_{\lambda\mu} \neq 0$ only if $\mu \leq \lambda$ in the dominance order and hence the indexing of the sum in (6.1.2) is valid.

The Jack and Macdonald polynomials generalize the Schur polynomials in the following sense. The Jack polynomials are also characterized by the conditions (a) and (b), albeit with the scalar product defined instead by

$$\langle p_\lambda, p_\mu \rangle = \alpha^{l(\lambda)} \delta_{\lambda\mu} z_\lambda \tag{6.1.3}$$

where α is a positive real number. Likewise, the Macdonald polynomials are characterized by the same conditions, with the inner product defined by

$$\langle p_\lambda, p_\mu \rangle = \langle p_\lambda, p_\mu \rangle_{q,t} = \delta_{\lambda\mu} z_\lambda(q,t) := \delta_{\lambda\mu} z_\lambda \prod_{i=1}^{l(\lambda)} \frac{1-q^{\lambda_i}}{1-t^{\lambda_i}} \tag{6.1.4}$$

By taking the appropriate limits, we see that the Jack and Macdonald polynomials reduce to the Schur polynomials. In particular, the Jack polynomials reduce to the Schur polynomials by letting $\alpha = 1$ in (6.1.3). Likewise, the Macdonald polynomials specialize to the Schur polynomials by letting $q = t$ so that $z_\lambda(q,t) = z_\lambda$. The Macdonald polynomials in fact generalize the Jack polynomials. The latter can be retrieved by letting $q = t^\alpha$ and taking the limit $t \to 1$. In this way each factor $\frac{1-q^{\lambda_i}}{1-t^{\lambda_i}}$ tends to α so that in the limit we have $z_\lambda(t^\alpha, t) \to z_\lambda \alpha^{l(\lambda)}$. By virtue of this specialization, we will not study the Jack polynomials in any more depth, and will devote what remains to studying the Macdonald polynomials.

Now that we have provided some motivation for its definition, we will devote the remainder of this section to investigating the inner product $\langle \cdot, \cdot \rangle_{q,t}$ defined by (6.1.4). Let $F = \mathbb{Q}(q,t)$ be the field of rational functions in indeterminates q and t with rational coefficients. In what follows, we will work in the ring Λ_F of symmetric functions with coefficients in F, the part Λ_F^k which is homogeneous of degree k, or the n-variable projections $\Lambda_{n,F}$. It is in this setting that we will make sense of the variables q and t appearing in (6.1.4).

It can also be useful to think of q and t as real variables. In fact, we need only consider $q, t \in (0, 1]$ by virtue of the following proposition.

Proposition 6.1. *On each homogeneous component Λ_F^n of Λ_F, the scalar product $\langle\cdot,\cdot\rangle_{q,t}$ differs from the inner product defined with q^{-1} and t^{-1} only by a scalar. In particular,*

$$\langle f,g\rangle_{q^{-1},t^{-1}} = (q^{-1}t)^n\langle f,g\rangle_{q,t}$$

for $f,g\in\Lambda_F^n$.

Proof. This follows from the fact that

$$(q^{-1}t)^{|\lambda|}\left(\prod_{i=1}^{l(\lambda)}\frac{1-q^{\lambda_i}}{1-t^{\lambda_i}}\right) = \left(\prod_{i=1}^{l(\lambda)}q^{-\lambda_i}t^{\lambda_i}\right)\left(\prod_{i=1}^{l(\lambda)}\frac{1-q^{\lambda_i}}{1-t^{\lambda_i}}\right) = \prod_{i=1}^{l(\lambda)}\frac{q^{-\lambda_i}(1-q^{\lambda_i})}{t^{-\lambda_i}(1-t^{\lambda_i})} = \prod_{i=1}^{l(\lambda)}\frac{1-(q^{-1})^{\lambda_i}}{1-(t^{-1})^{\lambda_i}}$$

and hence $z_\lambda(q^{-1},t^{-1}) = (q^{-1}t)^{|\lambda|}z_{\lambda(q,t)}$, from which the conclusion follows. □

As with the original Hall inner product in Λ, orthogonality with respect to the (q,t)-Hall inner product is closely related to (generalizations of) the Cauchy identities. In order to state the appropriate generalizations, we first introduce the following notation.

Notation. Let a,q be indeterminates. We denote by $(a;q)_\infty$ the infinite product

$$(a;q)_\infty := \prod_{r=0}^{\infty}(1-aq^r)$$

which we regard as a formal power series in a and q.

Definition 6.1. Let $x=(x_1,x_2,\ldots)$, $y=(y_1,y_2,\ldots)$ be independent sets of indeterminates. We define the (q,t)-*Cauchy product* to be the formal power series expression

$$\Pi(x,y;q,t) := \prod_{i,j}\frac{(tx_iy_j;q)_\infty}{(x_iy_j;q)_\infty} \tag{6.1.5}$$

By setting $q=t$, we see that

$$\Pi(x,y;q,q) = \prod_{i,j}\frac{(qx_iy_j;q)_\infty}{(x_iy_j;q)_\infty} = \prod_{i,j}\frac{\prod_{r\geq 0}(1-q^{r+1}x_iy_j)}{\prod_{r\geq 0}(1-q^rx_iy_j)} = \prod_{i,j}\frac{\prod_{r\geq 1}(1-q^rx_iy_j)}{\prod_{r\geq 0}(1-q^rx_iy_j)} = \prod_{i,j}(1-x_iy_j)^{-1}$$

and hence we obtain the classic Cauchy product appearing on the lefthand side of (3.4.1). As with the classical Cauchy product, the (q,t)-Cauchy product admits several expressions as sums of symmetric functions. For instance, we have the identity [6, Ch. VI, (2.6)]:

$$\Pi(x,y;q,t) = \sum_\lambda z_\lambda(q,t)^{-1}p_\lambda(x)p_\lambda(y) \tag{6.1.6}$$

As we previously remarked, expansions of the (q,t)-Cauchy product are closely related to orthogonality with respect to the (q,t)-Hall inner product. In particular, we have the following generalization of Theorem 5.2. Note that in what follows we will often omit the subscripts on the inner product bracket if the inner product in question is clear from context.

Proposition 6.2. *For each $n \geq 0$, let $\{u_\lambda\}, \{v_\lambda\}$ be F-bases of Λ_F^n which are parametrized by partitions of n. Then the following are equivalent:*

(a) $\langle u_\lambda, v_\mu \rangle = \delta_{\lambda\mu}$ *for all* λ, μ, *and*

(b) $\sum_\lambda u_\lambda(x) v_\lambda(y) = \Pi(x, y; q, t)$.

Proof. The proof is similar to that of Theorem 5.2. Let $p_\lambda^* := z_\lambda(q,t)^{-1} p_\lambda$ so that $\langle p_\lambda^*, p_\mu \rangle = \delta_{\lambda\mu}$. If we write $u_\lambda = \sum_\rho a_{\lambda\rho} p_\rho^*$ and $v_\mu = \sum_\sigma b_{\mu\sigma} p_\sigma$, then evaluating the inner product $\langle u_\lambda, v_\mu \rangle$ gives

$$\langle u_\lambda, v_\mu \rangle = \sum_\rho a_{\lambda\rho} b_{\mu\rho}$$

and hence (a) is equivalent to

$$\sum_\rho a_{\lambda\rho} b_{\mu\rho} = \delta_{\lambda\mu} \tag{a'}$$

By virtue of (6.1.6), we see that (b) is equivalent to the identity

$$\sum_\lambda p_\lambda^*(x) p_\lambda(y) = \sum_\lambda u_\lambda(x) v_\lambda(y) = \sum_\lambda \left(\sum_\rho a_{\lambda\rho} p_\rho^*(x) \right) \left(\sum_\sigma b_{\lambda\sigma} p_\sigma(y) \right) = \sum_{\lambda,\rho,\sigma} a_{\lambda\rho} b_{\lambda\sigma} p_\rho^*(x) p_\sigma(y)$$

By comparing the coefficients appearing in the leftmost and rightmost expressions, we see that (b) is equivalent to

$$\sum_\lambda a_{\lambda\rho} b_{\lambda\sigma} = \delta_{\rho\sigma} \tag{b'}$$

If we interpret (a') and (b') as statements about the components of products of matrices, then we see that they are equivalent statements by the argument presented in the proof of Theorem 5.2. $\square$

As we shall see in the following section, Proposition 6.2 will provide us with a (q,t)-generalization of the Cauchy identities for Schur functions involving the Macdonald polynomials. In the meantime, we will work towards a (q,t)-generalization of the Cauchy identity (5.1.1) involving the homogeneous and monomial symmetric functions. This will first require defining a suitable analogue of the homogeneous symmetric functions in Λ_F. We do so as follows.

Definition 6.2. Let y be a single independent variable. We denote by $g_n(x; q, t)$ the coefficient of y^n in the power series expansion of the (q,t)-Cauchy product

$$\Pi(x, y; q, t) := \sum_{n \geq 0} g_n(x; q, t) y^n \tag{6.1.7}$$

For any partition λ, we define $g_\lambda(x; q, t) := \prod_{i \geq 1} g_{\lambda_i}(x; q, t)$.

The polynomial $g_n(x; q, t)$ admits a simple closed form expression, and the g_λ are in fact a basis of Λ_F.

Proposition 6.3. *Let $n \geq 0$. Then*

$$g_n(x;q,t) = \sum_{\lambda \vdash n} z_\lambda(q,t)^{-1} p_\lambda(x) \tag{6.1.8}$$

and the g_λ form an F-basis of Λ_F which is dual to the basis given by the m_λ.

Proof. Beginning with the Cauchy identity (6.1.6) we set $y_2 = y_3 = \cdots = 0$ to obtain

$$\sum_{n\geq 0} g_n(x;q,t)y^n = \sum_\lambda z_\lambda(q,t)^{-1} p_\lambda(x)p_\lambda(y) = \sum_\lambda z_\lambda(q,t)^{-1}p_\lambda(x)y^{|\lambda|}$$

The identity (6.1.8) follows by comparing the coefficients of $y^{|\lambda|}$. The duality between the g_λ and m_λ can be seen as follows. Again letting $y = (y_1, y_2, \ldots)$ be a set of independent variables, we see that

$$\Pi(x,y;q,t) = \prod_j \left(\sum_{n\geq 0} g_n(x;q,t)y_j^n \right) = \sum_\lambda g_\lambda(x;q,t)m_\lambda(y)$$

By Proposition 6.2 we therefore have $\langle g_\lambda, m_\mu \rangle = \delta_{\lambda\mu}$ for all partitions λ and μ. □

We will conclude this section by defining a generalization of the involution $\omega : \Lambda \to \Lambda$ which acts on the Schur functions by transposition (cf. (3.3.3)). Let $u, v \in F$ with $u, v \neq 1$. We define the F-algebra endomorphism $\omega_{u,v}$ of Λ_F by

$$\omega_{u,v}(p_k) := (-1)^{k-1}\frac{1-u^k}{1-v^k}p_k \tag{6.1.9}$$

for each $r \geq 1$. For each partition λ, we likewise have

$$\omega_{u,v}(p_\lambda) = \varepsilon_\lambda \prod_{i=1}^{l(\lambda)} \frac{1-u^{\lambda_i}}{1-v^{\lambda_i}} p_\lambda \tag{6.1.9'}$$

where $\varepsilon(\lambda) := (-1)^{|\lambda|-l(\lambda)}$. Unlike the map ω, the endomorphism $\omega_{u,v}$ is not an involution in general, although it is invertible (and hence an automorphism). In particular, it is clear from the definition that $w_{v,u} = \omega_{u,v}^{-1}$. Moreover, $\omega_{u,u} = \omega$, the original involution exchanging h_k and e_k. This can be seen by (2.13) in [6, Ch. I], which states that $\omega(p_k) = (-1)^{k-1}p_k$. The following proposition provides another property that the g_n share with the homogeneous symmetric functions in Λ; namely, the way in which the map $\omega_{q,t}$ acts on them.

Proposition 6.4. *For all $k \geq 0$ we have $\omega_{q,t}(g_k(x;q,t)) = e_k(x)$.*

Proof. By (6.1.8) we have

$$\begin{aligned}
\omega_{q,t}(g_k(x;q,t)) &= \sum_{\lambda \vdash k} z_\lambda(q,t)^{-1}\omega_{q,t}(p_\lambda(x)) \\
&= \sum_{\lambda \vdash k} z_\lambda(q,t)^{-1}\varepsilon_\lambda \prod_{i=1}^{l(\lambda)} \frac{1-q^{\lambda_i}}{1-t^{\lambda_i}} p_\lambda(x) \\
&= \sum_{\lambda \vdash k} z_\lambda^{-1}\varepsilon_\lambda p_\lambda(x) \\
&= e_k
\end{aligned}$$

The final equality is (2.14') in [6, Ch. I]. □

6.2 Macdonald Polynomials

In this section we will prove the existence of the Macdonald polynomials alluded to at the beginning of the previous section by defining them to be the simultaneous eigenfunctions of a family of linear operators. We offer here a condensed summary of the construction as outlined in [5, §2]. For more details, one may refer to [6, Ch. VI, §3,4]. The central result of this section is the following existence theorem for the Macdonald polynomials.

Theorem 6.5. *For each partition λ there is a unique symmetric function $P_\lambda = P_\lambda(x;q,t) \in \Lambda_F$ such that*

(*a*) $P_\lambda = m_\lambda + \sum_{\mu<\lambda} u_{\lambda\mu} m_\mu$ *with coefficients* $u_{\lambda\mu} \in F$*, and*

(*b*) $\langle P_\lambda, P_\mu\rangle_{q,t} = 0$ *if* $\lambda \neq \mu$.

We call the symmetric functions P_λ the Macdonald *symmetric functions.*

Proof. We will first prove the existence of the Macdonald polynomials in finitely many variables $x = (x_1, \dots, x_n)$. We will do so by constructing an F-linear operator $D = D_{q,t} : \Lambda_{n,F} \to \Lambda_{n,F}$ with the following properties:

(i) The matrix of D relative to the basis given by the m_λ is triangular. Namely, $Dm_\lambda = \sum_{\mu \leq \lambda} c_{\lambda\mu} m_\mu$ for each partition λ with $l(\lambda) \leq n$;

(ii) D is self-adjoint. That is, $\langle Df, g\rangle_{q,t} = \langle f, Dg\rangle_{q,t}$ for all $f, g \in \Lambda_{n,F}$;

(iii) The eigenvalues of D are distinct. Namely, $\lambda \neq \mu$ implies $c_{\lambda\lambda} \neq c_{\mu\mu}$.

We will explicitly construct such a linear operator following the end of this proof. We now define the Macdonald polynomials as the eigenfunctions of D. In particular, we will show that for each partition λ with $l(\lambda) \leq n$, there is a unique symmetric polynomial $P_\lambda \in \Lambda_{n,F}$ satisfying the conditions

(a) $P_\lambda = \sum\limits_{\mu \leq \lambda} u_{\lambda\mu} m_\mu$ with $u_{\lambda\mu} \in F$ and $u_{\lambda\lambda} = 1$, and

(c) $DP_\lambda = c_{\lambda\lambda} P_\lambda$.

Indeed, (a) together with (i) implies that

$$DP_\lambda = \sum_{\mu \leq \lambda} u_{\lambda\mu} Dm_\mu = \sum_{\nu \leq \mu \leq \lambda} u_{\lambda\mu} c_{\mu\nu} m_\nu$$

and by (a) we have $c_{\lambda\lambda} P_\lambda = \sum\limits_{\nu \leq \lambda} c_{\lambda\lambda} u_{\lambda\nu} m_\nu$. Consequently (a) and (c) are satisfied if and only if

$$c_{\lambda\lambda} u_{\lambda\nu} = \sum_{\nu \leq \mu \leq \lambda} u_{\lambda\mu} c_{\mu\nu}$$

or equivalently

$$(c_{\lambda\lambda} - c_{\nu\nu}) u_{\lambda\nu} = \sum_{\nu < \mu \leq \lambda} u_{\lambda\mu} c_{\mu\nu}$$

whenever $\nu < \lambda$. In this case $c_{\lambda\lambda} \neq c_{\nu\nu}$ by (iii) and so the equation above uniquely determines $u_{\lambda\nu}$ in terms of the $u_{\lambda\mu}$ such that $\nu < \mu \leq \lambda$. Hence the coefficients $u_{\lambda\mu}$ in (a) are uniquely determined after normalizing so that $u_{\lambda\lambda} = 1$.

Having proven that there are symmetric polynomial eigenfunctions P_λ of D which satisfy (a), it only remains to show that they are pairwise orthogonal. By the self-adjointness condition (ii) we have

$$c_{\lambda\lambda} \langle P_\lambda, P_\mu \rangle_{q,t} = \langle DP_\lambda, P_\mu \rangle_{q,t} = \langle P_\lambda, DP_\mu \rangle_{q,t} = c_{\mu\mu} \langle P_\lambda, P_\mu \rangle_{q,t}$$

Since $c_{\lambda\lambda} \neq c_{\mu\mu}$ when $\lambda \neq \mu$, then $\langle P_\lambda, P_\mu \rangle_{q,t} = 0$ for $\lambda \neq \mu$. Hence the P_λ satisfy (b), which completes the proof in $\Lambda_{n,F}$. In order to pass to Λ_F, one must modify the operator D so that it is compatible with the projection homomorphisms π_n of §2.7. The necessary modifications can be found in [6, Ch. VI, §4], where the operator $D : \Lambda_{n,F} \to \Lambda_{n,F}$ is denoted by D_n^1. The eigenfunctions of the modified operator E_n are precisely the Macdonald polynomials we have already constructed, and consequently the P_λ are well-defined as elements of Λ_F. $\square$

The existence proof of the Macdonald polynomials which we have presented ultimately rests on the existence of a linear operator D satisfying the conditions (i)-(iii). We will now explicitly define this operator. We will not verify that it satisfies the properties (i)-(iii), as the proofs are rather technical and are not particularly important for our development of the theory of Macdonald polynomials. The cautious reader can find the necessary details in [5, §2]. In order to define the operator D, we fix the following notation.

Notation. We denote by V the Vandermonde determinant $V(x) = A_\delta(x) = \prod_{1 \leq i < j \leq n}(x_i - x_j)$ in n variables, as before. For each $u \in F$, $1 \leq i \leq n$, we define the "shift" operator $T_{u,x_i} : F[x_1, \ldots, x_n] \to F[x_1, \ldots, x_n]$ by

$$(T_{u,x_i} f)(x_1, \ldots, x_n) := f(x_1, \ldots, ux_i, \ldots, x_n)$$

Definition 6.3. We define the operator D as follows:

$$D := V^{-1}\sum_{i=1}^{n}(T_{t,x_i}V)T_{q,x_i} = \sum_{i=1}^{n}\left(\prod_{i\neq j}\frac{tx_i - x_j}{x_i - x_j}\right)T_{q,x_i} \tag{6.2.1}$$

Using the fact that $V = A_\delta = \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)x^{\sigma\delta}$, we may equivalently write D as

$$D = V^{-1}\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\sum_{i=1}^{n}t^{(\sigma\delta)_i}x^{\sigma\delta}T_{q,x_i} \tag{6.2.1'}$$

It is also worth noting here that the eigenvalues $c_{\lambda\lambda}$ of D are given by (2.11) in [5] as

$$c_{\lambda\lambda} = \sum_{i=1}^{n}q^{\lambda_i}t^{n-i} \tag{6.2.2}$$

Having now proved the existence of the Macdonald polynomials, we will dedicate the remainder of this section to establishing some of their basic algebraic properties. Many of these properties generalize analogous properties of the Schur polynomials that we encountered in §3, for as we remarked in the previous section, the Macdonald polynomials reduce to the Schur polynomials by letting $q = t$.

In light of Proposition 6.2 and the orthogonality condition (b) in Theorem 6.5, we have an analogue of the Cauchy identity (3.4.1) for the Macdonald functions. In order to state the identity, we first introduce the following "normalized" Macdonald functions.

Notation. Let $b_\lambda = b_\lambda(q,t) := \langle P_\lambda, P_\lambda\rangle_{q,t}^{-1}$, and let $Q_\lambda := b_\lambda P_\lambda$.

Clearly $\langle P_\lambda, Q_\mu\rangle = \delta_{\lambda\mu}$ so by our previous remark we obtain the following identity:

$$\Pi(x,y;q,t) = \sum_{\lambda}P_\lambda(x;q,t)Q_\lambda(y;q,t) \tag{6.2.3}$$

As we remarked earlier, the parameters q and t can be thought of as real variables, and in particular one need only consider real variables in the interval (0,1]. This is reflected in the following proposition, which states that the Macdonald polynomials P_λ are invariant under taking reciprocals in q and t.

Proposition 6.6. *The Macdonald functions satisfy the following identities when evaluated at reciprocals of q and t:*

(i) $P_\lambda(x;q^{-1},t^{-1}) = P_\lambda(x;q,t)$, *and ,*

(ii) $Q_\lambda(x;q^{-1},t^{-1}) = (qt^{-1})^{|\lambda|}Q_\lambda(x;q,t)$.

Proof. Both identities in follow from Proposition 6.1. In particular, since the inner products defined by (q,t) and (q^{-1},t^{-1}) differ by a scalar factor on each Λ_F^n, we have (i). The scalar factor appearing in (ii) follows by computing $b_\lambda(q^{-1},t^{-1})$ using the scalar factor appearing in Proposition 6.1. □

As we have also mentioned, the Macdonald polynomials generalize both the Schur and Jack polynomials, which can be obtained by taking the appropriate limits in q and t. By instead considering particular choices of partition λ, we also have the following interesting special cases.

Proposition 6.7. *For partitions consisting of either a single row or a single column, the corresponding Macdonald functions are given by*

(i) $P_{(k)}(x;q,t) = \frac{(q;q)_k}{(t;q)_k} g_k(x;q,t)$ *where* $(a;q)_k := \prod_{r=0}^{k-1}(1-aq^r)$,

(ii) $Q_{(k)}(x;q,t) = g_k(x;q,t)$, *and*

(iii) $P_{(1^k)}(x;q,t) = e_k(x)$.

Proof. By Proposition 6.3, g_k is orthogonal to all m_μ with $\mu \neq (k)$. Consequently, g_k is orthogonal to all P_μ except for $\mu = (k)$ and hence g_k is a scalar multiple of $P_{(k)}$. The particular scalar that appears in (i) is due to Example 1 of [6, Ch. VI, §2]. The identity (iii) follows from the fact that (1^k) is minimal in the dominance order, and hence $P_{(1^k)} = m_{(1^k)} = e_k$. The proof of (ii) will appear at the end of this section, after we have presented the duality result (6.2.5). $\square$

The following two propositions establish two further properties of the Macdonald polynomials which generalize analogous properties of the Schur polynomials. In particular, the Macdonald polynomials also possess the vanishing property stated in Lemma 3.7.

Proposition 6.8. *If λ is a partition with $l(\lambda) > n$, then $P_\lambda(x_1,\dots,x_n;q,t) = 0$.*

Proof. By (1.11) of [6, Ch. I], the dominance ordering satisfies the property that $\mu \leq \lambda$ if and only if $\lambda' \leq \mu'$. If $l(\lambda) > n$ then for all $\mu \leq \lambda$ we have $l(\mu) = \mu_1' \geq \lambda_1' = l(\lambda) > n$. Hence $m_\mu(x_1,\dots,x_n) = 0$ for each $\mu \leq \lambda$, and so $P_\lambda(x_1,\dots,x_n;q,t) = 0$. $\square$

The following generalization of Proposition 3.6 also holds for the Macdonald polynomials.

Proposition 6.9. *Let a be a positive integer. Then*

$$P_{\lambda+(a^n)}(x_1,\dots,x_n;q,t) = (x_1\cdots x_n)^a P_\lambda(x_1,\dots,x_n;q,t) \tag{6.2.4}$$

Proof. It suffices to show that the identity holds for $a = 1$. We have $T_{q,x_i}(x_1\cdots x_n P_\lambda) = qx_1\cdots x_n P_\lambda$ and consequently

$$D(x_1\cdots x_n P_\lambda) = qx_1\cdots x_n D(P_\lambda) = q\left(\sum_{i=1}^n q^{\lambda_i}t^{n-i}\right)x_1\cdots x_n P_\lambda$$

by (6.2.2). Thus $P_{\lambda+(1^n)}$ and $x_1\cdots x_n P_\lambda$ are both eigenfunctions of D with eigenvalue $\sum_{i=1}^n q^{\lambda_i+1}t^{n-i}$. Since the eigenvalues of D are distinct, then the polynomials $P_{\lambda+(1^n)}$ and $x_1\cdots x_n P_\lambda$ only differ by a scalar. Since they each have the leading term $x^{\lambda+(1^n)}$ then they are in fact equal. $\square$

We will conclude this section by considering the action of the automorphism $\omega_{q,t}$ on the Macdonald polynomials. The principal result in this vein are the identities (5.1) in [6, Ch. VI, §5], which generalize the identity (3.3.3) as follows:

$$\begin{aligned}\omega_{q,t}(P_\lambda(x;q,t)) &= Q_{\lambda'}(x;t,q) \\ \omega_{q,t}(Q_\lambda(x;q,t)) &= P_{\lambda'}(x;t,q)\end{aligned} \tag{6.2.5}$$

As usual, by taking the specialization $q = t$ we obtain the identity (3.3.3) for Schur functions, namely the fact that $\omega(S_\lambda) = S_{\lambda'}$. From (6.2.5) we may easily deduce that $b_\lambda(q,t)b_{\lambda'}(t,q) = 1$. By applying $\omega_{q,t}$ to the Cauchy identity (6.2.3), we also obtain the dual version, reminiscent of (3.4.2):

$$\begin{aligned}\prod_{i,j}(1 + x_i y_j) &= \sum_\lambda P_\lambda(x;q,t)P_{\lambda'}(y;t,q) \\ &= \sum_\lambda Q_\lambda(y;q,t)Q_{\lambda'}(x;t,q)\end{aligned} \tag{6.2.6}$$

We may now conclude by completing the proof of Proposition 6.7. By (iii) of the same proposition we have $Q_{(k)}(x;q,t) = \omega_{t,q}(P_{(1^k)}(x;t,q)) = \omega_{t,q}(e_k(x))$. Using Proposition 6.4 we have $\omega_{t,q}(e_k(x)) = \omega_{q,t}^{-1}(e_k(x)) = g_k(x;q,t)$, which proves (ii).

6.3 Pieri Formulas for the Macdonald Polynomials

The central result of Section 5 was Theorem 5.12, and in particular its specialization to Schur functions. This combinatorial expression for the Schur functions established a deep correspondence between the algebraic properties of symmetric functions . The purpose of this section and the next is to work towards an analogous combinatorial expression for the Macdonald polynomials. A crucial step in working towards this goal, which is interesting in its own right, is to generalize the Pieri formulas of Proposition 5.14 to the Macdonald setting; we will do so in $\Lambda_{n,F}$. This will require the following definitions and notation.

Definition 6.4. For each partition μ with $l(\mu) \le n$, define a homomorphism $u_\mu : F[x_1, \ldots, x_n] \to F$ by $u_\mu(x_i) := q^{\mu_i} t^{n-i}$, $1 \le i \le n$.

Remark. For any polynomial f we have $u_0(f) = f(t^\delta) = f(t^{n-1}, t^{n-2}, \ldots, t, 1)$.

Definition 6.5. Let μ, ν be partitions of length at most n such that $\mu \subset \nu$ and $\theta = \nu/\mu$ is a vertical strip. Define

$$B_{\nu/\mu} := t^{n(\nu)-n(\mu)} \prod_{1\le i<j\le n} \frac{1 - q^{\mu_i-\mu_j} t^{j-i+\theta_i-\theta_j}}{1 - q^{\mu_i-\mu_j} t^{j-i}} \tag{6.3.1}$$

where $n(\mu) := \sum_{i\ge1}(i-1)\mu_i$.

Notation. Write $\tilde{P}_\lambda := \frac{P_\lambda}{u_0(P_\lambda)}$. With this notation, we are now ready to provide our initial formulation of the Pieri formula for Macdonald polynomials.

Proposition 6.10. *For all partitions λ, μ of length at most n, we have*

$$u_\lambda(\tilde{P}_\mu) = u_\mu(\tilde{P}_\lambda) \tag{6.3.2}$$

For all partitions σ of length at most n and integers $r > 0$, we have

$$\tilde{P}_\sigma e_r = \sum_\nu B_{\nu/\sigma} \tilde{P}_\nu \tag{6.3.3}$$

where the sum is taken over all partitions $\nu \supset \sigma$ of length at most n such that ν/σ is a vertical r-strip.

Proof. We will prove (6.3.2) and (6.3.3) simultaneously by induction on $|\mu|$. First consider the case where λ is any partition and $\mu = 0$ is the empty partition. Then

$$u_\lambda(\tilde{P}_0) = u_\lambda(1) = 1 = \frac{u_0(P_\lambda)}{u_0(P_\lambda)} = u_0(\tilde{P}_\lambda)$$

and so (6.3.2) is easily seen to hold. Now let $\mu \neq 0$ be a partition of length at most n and suppose for our induction hypothesis that (6.3.2) holds for all partitions λ, π such that $|\pi| < |\mu|$ or $|\pi| = |\mu|$ and $\pi < \mu$.

Let $r \geq 1$ be such that $\mu_r > \mu_{r+1}$ and let $\sigma := \mu - 1^r$. To show that (6.3.3) holds in general, it suffices to prove that it holds for this particular choice of σ, as any choice of partition σ and integer $r \geq 1$ can be realized in this way for a suitable partition μ (namely $\mu = \sigma + 1^r$). We will now prove that (6.3.3) does in fact hold for this σ.

Consider the product $\tilde{P}_\sigma e_r$. The leading monomial in $\tilde{P}_\sigma$ is x^σ and the leading monomial in e_r is $x_1 \cdots x_r = x^{(1^r)}$. By definition of σ, the leading monomial in $\tilde{P}_\sigma e_r$ is therefore x^μ. Consequently

$$\tilde{P}_\sigma e_r = \sum_{\nu \leq \mu} a_\nu \tilde{P}_\nu \tag{6.3.4}$$

for suitable coefficients a_ν. For any partition π, the result (6.5) from [6, Ch. VI] states that:

$$u_\pi(e_r) u_\sigma(\tilde{P}_\pi) = \sum_\nu B_{\nu/\sigma} u_\nu(\tilde{P}_\pi) \tag{6.3.5}$$

where the sum is taken over all partitions $\nu \supset \sigma$ such that $l(\nu) \leq n$ and ν/σ is a vertical r-strip.

Suppose that $|\pi| = |\mu|$ and $\pi \leq \mu$. By our inductive hypothesis, $u_\sigma(\tilde{P}_\pi) = u_\pi(\tilde{P}_\sigma)$ since $|\sigma| < |\mu| = |\pi|$. Likewise, $u_\nu(\tilde{P}_\pi) = u_\pi(\tilde{P}_\nu)$ for all $\nu \leq \mu$; this is guaranteed by our inductive hypothesis if $\pi < \nu$ or $\nu < \mu$, and the only other possibility is that $\pi = \nu = \mu$. Consequently, we may rewrite (6.3.5) in the following form:

$$u_\pi(e_r \tilde{P}_\sigma) = \sum_\nu B_{\nu/\sigma} u_\pi(\tilde{P}_\nu) \tag{6.3.6}$$

On the other hand, we may apply the map u_π to both sides of (6.3.4) to obtain

$$u_\pi(e_r\tilde{P}_\sigma) = \sum_{\nu\leq\mu} a_\nu u_\pi(\tilde{P}_\nu) \tag{6.3.7}$$

where the sum is indexed as before. Using the result (6.8) from [6, Ch. VI], which states that $\det(u_\pi(\tilde{P}_\nu))_{\pi,\nu\leq\mu} \neq 0$, then by comparing (6.3.6) and (6.3.7) we conclude that $a_\nu = B_{\nu/\sigma}$ if ν/σ is a vertical r-strip, and $a_\nu = 0$ otherwise. In this way (6.3.4) becomes

$$\tilde{P}_\sigma e_r = \sum_\nu B_{\nu/\sigma}\tilde{P}_\nu$$

with the sum taken over all partitions $\nu \supset \sigma$ of length at most n such that ν/σ is a vertical r-strip. This completes the proof of the Pieri formula (6.3.3).

We will now complete the inductive step in order to finish the proof of the symmetry result (6.3.2). Let λ be any partition with $l(\lambda) \leq n$. Applying u_λ to both sides of the Pieri formula (6.3.3) gives

$$u_\lambda(\tilde{P}_\sigma)u_\lambda(e_r) = \sum_\nu B_{\nu/\sigma}u_\lambda(\tilde{P}_\nu)$$

By our induction hypothesis we have $u_\lambda(\tilde{P}_\sigma) = u_\sigma(\tilde{P}_\lambda)$ since $|\sigma| < |\mu|$, and $u_\lambda(\tilde{P}_\nu) = u_\nu(\tilde{P}_\lambda)$ if $\nu \neq \mu$. Hence the previous equation may be written as

$$u_\sigma(\tilde{P}_\lambda)u_\lambda(e_r) = B_{\mu/\sigma}u_\lambda(\tilde{P}_\mu) + \sum_{\nu<\mu} B_{\nu/\sigma}u_\nu(\tilde{P}_\lambda)$$

On the other hand, from (6.5) of [6, Ch. VI] we have

$$u_\sigma(\tilde{P}_\lambda)u_\lambda(e_r) = B_{\mu/\sigma}u_\mu(\tilde{P}_\lambda) + \sum_{\nu<\mu} B_{\nu/\sigma}u_\nu(\tilde{P}_\lambda)$$

Since $B_{\mu/\sigma} \neq 0$, then comparing the previous two equations shows that $u_\lambda(\tilde{P}_\mu) = u_\mu(\tilde{P}_\lambda)$, which completes the inductive step. $\square$

With a slight adjustment in notation, we may put the Pieri formula (6.3.3) into a more convenient form.

Notation. Let λ, μ be partitions. Define

$$\Psi'_{\lambda/\mu} := \frac{B_{\lambda/\mu}u_0(P_\mu)}{u_0(P_\lambda)} \tag{6.3.8}$$

Then we may rewrite the Pieri formula as

$$P_\mu e_r = \sum_\lambda \Psi'_{\lambda/\mu}P_\lambda \tag{6.3.3'}$$

summed over all partitions $\lambda \supset \mu$ of length at most n such that λ/μ is a vertical r-strip.

Given the explicit definition of $B_{\lambda/\mu}$, one may deduce an explicit expression for the coefficient $\Psi'_{\lambda/\mu}$. In particular, one has the identity (6.11) from [6, Ch. VI]:

$$u_0(P_\lambda) = P_\lambda(1, t, \ldots, t^{n-1}; q, t) = t^{n(\lambda)} \frac{u_\lambda(\Delta^+)}{u_0(\Delta^+)} \tag{6.3.9}$$

where we define

$$\Delta^+ = \Delta^+(q, t) := \prod_{1 \le i < j \le n} \frac{(x_i x_j^{-1}; q)_\infty}{(t x_i x_j^{-1}; q)_\infty}$$

From this one can derive the identity (6.13) (*loc. cit.*):

$$\Psi'_{\lambda/\mu} = \prod \frac{\left(1 - q^{\mu_i - \mu_j} t^{j-i-1}\right)\left(1 - q^{\lambda_i - \lambda_j} t^{j-i+1}\right)}{\left(1 - q^{\mu_i - \mu_j} t^{j-i}\right)\left(1 - q^{\lambda_i - \lambda_j} t^{j-i}\right)} \tag{6.3.10}$$

where the product is taken over all pairs (i, j) such that $i < j$, $\lambda_i = \mu_i$, and $\lambda_j = \mu_j + 1$.

Although (6.3.10) does provide an explicit expression for the coefficients $\Psi'_{\lambda/\mu}$, it is rather complicated to work with, especially from a theoretical perspective. As such, we will work towards an alternate expression which perhaps uses more information about the geometry of the diagrams of λ and μ, and is slightly more amenable to work with. A necessary step in obtaining this formula is to provide an explicit expression for the scalar $b_\lambda(q, t) := \langle P_\lambda, P_\lambda \rangle_{q,t}^{-1}$. This requires the following definition.

Definition 6.6. Let u be an indeterminate. Define a homomorphism $\varepsilon_{u,t} : \Lambda_F \to F$ by

$$\varepsilon_{u,t}(p_r) := \frac{1 - u^r}{1 - t^r}$$

for each integer $r \ge 1$.

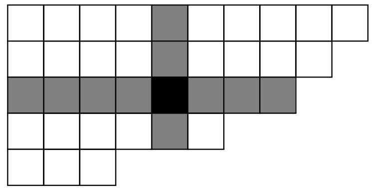
Remark. Note in particular that $\varepsilon_{t^n,t}(p_r) = \frac{1-t^{nr}}{1-t^r} = p_r(1, t, \ldots, t^{n-1})$. Hence for $f \in \Lambda_F$ we have $\varepsilon_{t^n,t}(f) = f(1, t, \ldots, t^{n-1}) = u_0(f)$.

In order to consider the action of $\varepsilon_{u,t}$ on the Macdonald polynomials, we must first introduce several pieces of terminology related to Young diagrams.

Definition 6.7. Let λ be a partition and let $s = (i, j) \in \lambda$ be a box in the Young diagram of λ. The *arm length* of s in λ is given by $a(s) = a_\lambda(s) := \lambda_i - j$. The *leg length* s in λ is given by $l(s) = l_\lambda(s) := \lambda'_j - i$. The *arm colength* and *leg colength* of s in λ are given by $a'(s) := j - 1$ and $l'(s) := i - 1$, respectively.

Geometrically, the arm length, leg length, arm colength, and leg colength of s in λ are the number of boxes sharing a row or column with s which lie strictly to the east, south, west, or north of s, respectively.

Example. Let $\lambda = (10, 9, 8, 6, 3)$, $s = (3, 5)$.



From the diagram above, one can see that $a(s) = 3$, $l(s) = 1$, $a'(s) = 4$, and $l'(s) = 2$.

With this terminology, we may present the following identity (6.11') from [6, Ch. VI], which gives an explicit expression for $u_0(P_\lambda)$:

$$u_0(P_\lambda) = P_\lambda(1, t, \ldots, t^{n-1}; q, t) = t^{n(\lambda)} \prod_{s \in \lambda} \frac{1 - q^{a'(s)} t^{n - l'(s)}}{1 - q^{a(s)} t^{l(s)+1}} \tag{6.3.9'}$$

We may now consider how $\varepsilon_{u,t}$ acts on the Macdonald polynomials; this is summarized in the following proposition.

Proposition 6.11.

$$\varepsilon_{u,t}(P_\lambda) = \prod_{s \in \lambda} \frac{t^{l'(s)} - q^{a'(s)} u}{1 - q^{a(s)} t^{l(s)+1}} \tag{6.3.11}$$

Proof. By (6.3.9'), together with the fact that $\sum_{s \in \lambda} l'(s) = n(\lambda)$, the conclusion holds when $u = t^n$ for each $n \geq l(\lambda)$. Both sides of (6.3.11) are polynomials in u over F which agree for infinitely many values of u and hence they are equal. □

In order to provide an explicit expression for the scalar $b_\lambda(q, t)$, we require one additional lemma.

Lemma 6.12. *Let $f \in \Lambda_F$ be homogeneous of degree k. Then*

$$\varepsilon_{u,t} \omega_{t,q}(f) = (-q)^{-r} \varepsilon_{u,q^{-1}}(f) \tag{6.3.12}$$

Proof. Since $\varepsilon_{u,t}$ and $\omega_{t,q}$ are both algebra homomorphisms, it suffices to verify the identity on a basis, say the p_λ with $|\lambda| = k$. In fact, by the multiplicativity of both maps it suffices to verify the identity on the factors p_{λ_i}. In this case we have

$$\varepsilon_{u,t} \omega_{t,q}(p_{\lambda_i}) = \varepsilon_{u,t} \left((-1)^{\lambda_i - 1} \frac{1 - t^{\lambda_i}}{1 - q^{\lambda_i}} p_\lambda \right) = (-1)^{\lambda_i} \frac{1 - u^{\lambda_i}}{q^{\lambda_i} - 1} = (-q)^{-\lambda_i} \frac{1 - u^{\lambda_i}}{1 - q^{-\lambda_i}} = (-q)^{-\lambda_i} \varepsilon_{u,q^{-1}}(p_{\lambda_i})$$

which is the desired identity. □

We are now ready to prove the following explicit expression for the $b_\lambda(q, t)$:

Proposition 6.13.

$$b_\lambda(q, t) = \prod_{s \in \lambda} \frac{1 - q^{a(s)} t^{l(s)+1}}{1 - q^{a(s)+1} t^{l(s)}} \tag{6.3.13}$$

Proof. Using the previous Lemma, we have

$$\begin{aligned} \varepsilon_{u,t}P_\lambda(q,t) &= \varepsilon_{u,t}\omega_{t,q}Q_{\lambda'}(t,q) \\ &= (-q)^{-|\lambda|}\varepsilon_{u,q^{-1}}Q_{\lambda'}(t,q) \\ &= (-q)^{-|\lambda|}b_{\lambda'}(t,q)\varepsilon_{u,q^{-1}}P_{\lambda'}(t^{-1},q^{-1}) \end{aligned}$$

Consequently we have

$$b_{\lambda'}(t,q) = \frac{(-q)^{-|\lambda|}\varepsilon_{u,t}P_\lambda(q,t)}{\varepsilon_{u,q^{-1}}P_{\lambda'}(t^{-1},q^{-1})}$$

Applying Proposition 6.11 to the denominator gives

$$\varepsilon_{u,q^{-1}}P_{\lambda'}(t^{-1},q^{-1}) = \prod_{s\in\lambda'}\frac{q^{-l'(s)} - t^{-a'(s)}u}{1-t^{-a(s)}q^{-l(s)-1}} = (-q)^{|\lambda|}\prod_{s\in\lambda}\frac{t^{l'(s)} - q^{a'(s)}u}{1-q^{a(s)+1}t^{l(s)}}$$

Applying Proposition 6.11 to the numerator then gives

$$b_{\lambda'}(t,q) = \left(\prod_{s\in\lambda}\frac{t^{l'(s)} - q^{a'(s)}u}{1-q^{a(s)}t^{l(s)+1}}\right)\left(\prod_{s\in\lambda}\frac{t^{l'(s)} - q^{a'(s)}u}{1-q^{a(s)+1}t^{l(s)}}\right)^{-1} = \prod_{s\in\lambda}\frac{1-q^{a(s)+1}t^{l(s)}}{1-q^{a(s)}t^{l(s)+1}}$$

As we remarked at the end of Section 6.2, we have $b_\lambda(q,t)b_{\lambda'}(t,q) = 1$ and hence the conclusion follows by taking the reciprocal. □

We will conclude this section first by revisiting the coefficients $\Psi'_{\lambda/\mu}$ appearing in the Pieri formula. As we remarked earlier, our aim will be to provide an explicit expression for these coefficients different from (6.3.10) which will involve the scalars b_λ. Using this result, we will also revisit Proposition 6.10 and give a more complete statement of the Macdonald polynomial Pieri formulas.

Notation. Let λ be a partition, $s = (i,j)$ a square (not necessarily in λ). We define

$$b_\lambda(s) = b_\lambda(s;q,t) := \begin{cases} \frac{1-q^{a(s)}t^{l(s)+1}}{1-q^{a(s)+1}t^{l(s)}} & s\in\lambda \\ 1 & \text{otherwise} \end{cases}$$

so that $b_\lambda = \prod_s b_\lambda(s)$. If $s = (i,j)\in\lambda$, let $s' = (j,i)\in\lambda'$. It is clear from the definitions that $a_\lambda(s) = l_{\lambda'}(s)$ and $l_\lambda(s) = a_{\lambda'}(s)$. It is also clear that $b_\lambda(s;q,t) = b_{\lambda'}(s';t,q)^{-1}$.

Proposition 6.14. *Let $\lambda\supset\mu$ be partitions and denote by $C_{\lambda/\mu}$ (resp. $R_{\lambda/\mu}$) the union of the columns (resp. rows) that intersect λ/μ. Then*

$$\Psi'_{\lambda/\mu} = \prod_{s\in(C_{\lambda/\mu})\setminus(R_{\lambda/\mu})}\frac{b_\lambda(s)}{b_\mu(s)} \tag{6.3.14}$$

Proof. Given a product of the form $P = \prod_{a,b\geq 0}(1-q^a t^b)^{n_{ab}}$, we may define a map L by $L(P) := \sum_{a,b\geq 0} n_{ab} q^a t^b$. The map L is in fact an injection, so it suffices to show that the image of both sides of the equation (6.3.14) under L are the same. By (6.3.10) we have

$$L(\Psi'_{\lambda/\mu}) = (t-q)\sum q^{\mu_i-\mu_j-1}(t^{j-i}-t^{j-i-1})$$

where the summation is over all pairs (i,j) such that $i<j$, $\lambda_i=\mu_i$, and $\lambda_j=\mu_j+1$. To each such pair we may associate the square $s=(i,\lambda_j)\in C_{\lambda/\mu}\setminus R_{\lambda/\mu}$. The contribution to $L(\Psi'_{\lambda/\mu})$ from the pairs (i,j) such that $(i,\lambda_j)=s$ is equal to

$$(t-q)q^{a_\lambda(s)}(t^{l_\lambda(s)}-t^{l_\mu(s)})$$

and hence $L(\Psi'_{\lambda/\mu})$ is the sum of these expressions for all $s\in\mu\cap(C_{\lambda/\mu}\setminus R_{\lambda/\mu})$. This in turn is equal to the image of the righthand side of (6.3.14) under L, as required. $\square$

We are now ready to present the complete Pieri formulas for the Macdonald polynomials.

Theorem 6.15. *Let μ be a partition, $r>0$ an integer. Then*

(i) $P_\mu g_r = \sum_\lambda \Phi_{\lambda/\mu}P_\lambda$

(ii) $Q_\mu g_r = \sum_\lambda \Psi_{\lambda/\mu}Q_\lambda$

(iii) $Q_\mu e_r = \sum_\lambda \Phi'_{\lambda/\mu}Q_\lambda$

(iv) $P_\mu e_r = \sum_\lambda \Psi'_{\lambda/\mu}P_\lambda$

In (i) and (ii) (resp. (iii) and (iv)) the sum is taken over all partitions $\lambda\supset\mu$ such that λ/μ is a horizontal (resp. vertical) r-strip. The coefficients are given by

(i) $\Phi_{\lambda/\mu} = \prod_{s\in C_{\lambda/\mu}} \frac{b_\lambda(s)}{b_\mu(s)}$

(ii) $\Psi_{\lambda/\mu} = \prod_{s\in(R_{\lambda/\mu})\setminus(C_{\lambda/\mu})} \frac{b_\mu(s)}{b_\lambda(s)}$

(iii) $\Phi'_{\lambda/\mu} = \prod_{s\in R_{\lambda/\mu}} \frac{b_\mu(s)}{b_\lambda(s)}$

(iv) $\Psi'_{\lambda/\mu} = \prod_{s\in(C_{\lambda/\mu})\setminus(R_{\lambda/\mu})} \frac{b_\lambda(s)}{b_\mu(s)}$

Proof. The Pieri formula (iv) is simply a restatement of Proposition 6.10 with coefficients given by Proposition 6.14. Applying the duality result (6.2.5) to (iv) gives (ii), with the coefficients given by

$$\Psi_{\lambda/\mu}(q,t) = \Psi'_{\lambda'/\mu'}(t,q)$$

By virtue of the fact that $b_\lambda(s;q,t)=b_{\lambda'}(s';t,q)^{-1}$, we see that $\Psi_{\lambda/\mu}(q,t)$ is given by the expression (ii) above. The formulas (i) and (iii) follow from (ii) and (iv) by taking the appropriate normalizations. In particular, the coefficients appearing in (i) and (iii) are

$$\Phi_{\lambda/\mu} = b_\lambda b_\mu^{-1}\Psi_{\lambda/\mu} \qquad \text{and} \qquad \Phi'_{\lambda/\mu} = b_\lambda^{-1}b_\mu\Psi'_{\lambda/\mu}$$

and (6.3.13) shows that they are given by the formulas above. $\square$

Remark. By comparing the expressions in Theorem 6.15, we see that

$$\Phi'_{\lambda/\mu}(q,t) = \Phi_{\lambda'/\mu'}(t,q) \qquad \text{and} \qquad \Psi'_{\lambda/\mu}(q,t) = \Psi_{\lambda'/\mu'}(t,q)$$

6.4 Skew Macdonald Polynomials

The purpose of this section will be to obtain combinatorial expressions for the Macdonald polynomials which generalize the tableau monomial formulas for the Schur polynomials (cf. Corollary 5.13). As with the Schur functions, we will follow the approach of Macdonald in [6] by developing a basic theory of skew Macdonald polynomials and obtaining the formulas for Macdonald polynomials as a special case. The development is not entirely the same in the Macdonald setting as in the Schur setting, although many of the definitions and results should be reminiscent of those encountered in Sections 5.2 and 5.3.

Recall that we defined the Littlewood-Richardson coefficients $c^\nu_{\lambda\mu}$ to be the structure constants for the Schur functions in Λ:

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

Recall as well that we defined the skew Schur function $S_{\nu/\lambda}$ by setting $\langle S_{\nu/\lambda}, S_\mu\rangle = \langle S_\nu, S_\lambda S_\mu\rangle$ so that $S_{\nu/\lambda} = \sum_\mu c^\nu_{\lambda\mu} S_\mu$. In direct analogy, we provide the following definition in the setting of Λ_F.

Definition 6.8. Let λ, μ, ν be partitions, and let

$$f^\nu_{\lambda\mu} = f^\nu_{\lambda\mu}(q,t) = \langle Q_\nu, P_\lambda P_\mu\rangle_{q,t}$$

Equivalently, we define the coefficients $f^\nu_{\lambda\mu}$ to be the structure constants for the Macdonald polynomials in Λ_F, namely by setting $P_\lambda P_\mu = \sum_\nu f^\nu_{\lambda\mu} P_\nu$.

Remark. Clearly $f^\nu_{\lambda\mu}(t,t) = c^\nu_{\lambda\mu}$. By Proposition 6.6 we also have $f^\nu_{\lambda\mu}(q,t) = f^\nu_{\lambda\mu}(q^{-1},t^{-1})$. As with the Littlewood-Richardson coefficients, it is also clear that $f^\nu_{\lambda\mu} = f^\nu_{\mu\lambda}$ by commutativity in Λ_F. Unlike the Littlewood-Richardson coefficients, $f^\nu_{\lambda\mu} \neq f^{\nu'}_{\lambda'\mu'}$ in general. The correct identity in the general case is as follows.

Proposition 6.16. *For all partitions λ, μ, ν we have*

$$f^\nu_{\lambda\mu}(q,t) = f^{\nu'}_{\lambda'\mu'}(t,q)\frac{b_\nu(q,t)}{b_\lambda(q,t)b_\mu(q,t)}$$

Proof. Applying the automorphism $\omega_{q,t}$ to both sides of the equation $P_\lambda P_\mu = \sum_\nu f^\nu_{\lambda\mu} P_\nu$ yields

$$Q_{\lambda'}(t,q)Q_{\mu'}(t,q) = \sum_\nu f^\nu_{\lambda\mu}(q,t)Q_{\nu'}(t,q)$$

or equivalently $Q_\lambda(q,t)Q_\mu(q,t) = \sum_\nu f^{\nu'}_{\lambda'\mu'}(t,q)Q_\nu(q,t)$. The conclusion follows by dividing both sides through by $b_\lambda(q,t)b_\mu(q,t)$ and comparing the coefficients of $P_\nu(q,t)$ with those appearing in the original equation $P_\lambda P_\mu = \sum_\nu f^\nu_{\lambda\mu} P_\nu$. □

The following proposition lists some conditions for the polynomials $f^\nu_{\lambda\mu}$ to be nonzero; these conditions are shared with the Littlewood-Richardson coefficients.

Proposition 6.17. *Let λ, μ, ν be partitions. Then*

(i) $f^\nu_{\lambda\mu} = 0$ *unless* $|\nu| = |\lambda| + |\mu|$, *and*

(ii) $f^\nu_{\lambda\mu} = 0$ *unless* $\lambda \subset \nu$ *and* $\mu \subset \nu$.

Proof. The condition (i) follows by degree considerations, as P_λ is homogeneous of degree $|\lambda|$. To prove (ii), we define the following subspace of Λ_F. For each partition μ, let I_μ denote the subspace of Λ_F generated by the P_ν such that $\nu \supset \mu$. By the Pieri formula (i), we see that $g_r I_\mu \subset I_\mu$ for each $r \geq 1$. Since the g_r generate Λ_F as an F-algebra, then I_μ is an ideal of Λ_F. Consequently $P_\lambda P_\mu \in I_\lambda \cap I_\mu$. By definition of the subspaces I_λ and I_μ, this implies (ii) above. $\square$

Although the coefficients $f^\nu_{\lambda\mu}$ share many properties with the Littlewood-Richardson coefficients $c^\nu_{\lambda\mu}$, explicit expressions have proven considerably more difficult to obtain outside of certain special cases. There is in fact a "Littlewood-Richardson rule" for Macdonald polynomials due to Yip [11], which provides a combinatorial method to compute the product of two Macdonald polynomials, although the details of this result are beyond the scope of the present discussion. Having exhausted all that we need to say about the $f^\nu_{\lambda\mu}$ for our purposes, we will now move on to developing a basic theory of skew Macdonald polynomials, which we define as follows.

Definition 6.9. Let λ, ν be partitions. We define the *skew Macdonald function* $Q_{\nu/\lambda}$ by setting $Q_{\nu/\lambda} := \sum_\mu f^\nu_{\lambda\mu} Q_\mu$, or equivalently by requiring

$$\langle Q_{\nu/\lambda}, P_\mu \rangle = \langle Q_\nu, P_\lambda P_\mu \rangle$$

We may also define the skew Macdonald function $P_{\nu/\lambda}$ by exchanging the P's and Q's in the previous definition:

$$\langle P_{\nu/\lambda}, Q_\mu \rangle = \langle P_\nu, Q_\lambda Q_\mu \rangle$$

Remark. Evidently we have $Q_{\nu/\lambda} = b_\nu b_\lambda^{-1} P_{\nu/\lambda}$. Moreover, from Proposition 6.17 we see that $Q_{\nu/\lambda} = 0$ unless $\nu \supset \lambda$, and that if $\nu \supset \lambda$ then $Q_{\nu/\lambda}$ is homogeneous of degree $|\nu| - |\lambda|$. By linearity, we also have the identity

$$\langle Q_{\nu/\lambda}, f \rangle = \langle Q_\nu, P_\lambda f \rangle \tag{6.4.1}$$

for all $f \in \Lambda_F$.

As in the case of Schur functions, the skew Macdonald functions provide us with simple expressions for Macdonald functions in two independent sets of variables (cf. Proposition 5.9).

Proposition 6.18. *Let x and y be independent sets of variables. We have*

$$Q_\lambda(x, y) = \sum_\mu Q_{\lambda/\mu}(x) Q_\mu(y) \tag{6.4.2}$$

$$P_\lambda(x, y) = \sum_\mu P_{\lambda/\mu}(x) P_\mu(y)$$

Proof. Introduce a third independent set of variables z. We have

$$\begin{aligned}\sum_{\lambda} Q_{\lambda/\mu}(x)P_\lambda(z) &= \sum_{\lambda,\nu} f^\lambda_{\mu\nu} Q_\nu(x)P_\lambda(z) && \text{(Definition 6.9)}\\ &= \sum_\nu Q_\nu(x)P_\mu(z)P_\nu(z) && \text{(Definition 6.8)}\\ &= P_\mu(z)\Pi(x,z;q,t) && (6.2.3)\end{aligned}$$

Consequently we have

$$\begin{aligned}\sum_{\lambda,\mu} Q_{\lambda/\mu}(x)P_\lambda(z)Q_\mu(y) &= \sum_\mu P_\mu(z)Q_\mu(y)\Pi(x,z;q,t)\\ &= \Pi(x,z;q,t)\Pi(z,y;q,t)\\ &= \sum_\lambda Q_\lambda(x,y)P_\lambda(z)\end{aligned}$$

The first identity in (6.4.2) follows by comparing the coefficients of $P_\lambda(z)$. The second identity follows from the first by dividing both sides by b_λ. $\square$

We are now equipped to provide combinatorial formulas for the skew Macdonald functions. Before presenting the theorem statement, we first review the following notation.

Notation. Let λ, μ be partitions and let T be a skew tableau on λ/μ, say with entries in $[n]$. We may think of T as a sequence of partitions

$$\mu = \lambda^{(0)} \subset \lambda^{(1)} \subset \cdots \subset \lambda^{(n)} = \lambda$$

such that each skew diagram $\lambda^{(i)}/\lambda^{(i-1)}$ is a horizontal strip. Let

$$\Phi_T(q,t) := \prod_{i=1}^n \Phi_{\lambda^{(i)}/\lambda^{(i-1)}}(q,t) \tag{6.4.3}$$

where as in Theorem 6.15 we define

$$\Phi_{\lambda^{(i)}/\lambda^{(i-1)}}(q,t) = \prod_{s\in C_{\lambda^{(i)}/\lambda^{(i-1)}}} \frac{b_{\lambda^{(i)}}(s;q,t)}{b_{\lambda^{(i-1)}}(s;q,t)}$$

Similarly, let

$$\Psi_T(q,t) := \prod_{i=1}^n \Psi_{\lambda^{(i)}/\lambda^{(i-1)}}(q,t) \tag{6.4.4}$$

where we define

$$\Psi_{\lambda^{(i)}/\lambda^{(i-1)}}(q,t) = \prod_{s\in(R_{\lambda^{(i)}/\lambda^{(i-1)}})\setminus(C_{\lambda^{(i)}/\lambda^{(i-1)}})} \frac{b_{\lambda^{(i-1)}}(s;q,t)}{b_{\lambda^{(i)}}(s;q,t)}$$

Theorem 6.19. *Let λ, μ be partitions. We have*

$$Q_{\lambda/\mu} = \sum_T \Phi_T(q,t) x^T \tag{6.4.5}$$
$$P_{\lambda/\mu} = \sum_T \Psi_T(q,t) x^T$$

where the summation is over all skew tableaux T on λ/μ, and x^T denotes the tableau monomial corresponding to T (cf. 4.4.3).

Proof. Since the bases of Λ_F given by the m_ν and g_ν are dual to one another, we have

$$Q_{\lambda/\mu} = \sum_\nu \langle Q_{\lambda/\mu}, g_\nu \rangle m_\nu = \sum_\nu \langle Q_\lambda, P_\mu g_\nu \rangle m_\nu \tag{6.4.6}$$

The righthand side follows by applying (6.4.1) with $f = g_\nu$. We proceed now by expanding the coefficients $\langle Q_\lambda, P_\mu g_\nu \rangle$. Repeated iterations of the Pieri formula (i) of Theorem 6.15 yield

$$P_\mu g_\nu = \sum_\pi \left(\sum_T \Phi_T \right) P_\pi \tag{6.4.7}$$

where the outer sum is over all partitions $\pi \supset \mu$ such that $|\pi/\mu| = |\nu|$, and the inner sum is over all skew tableaux T on π/μ with content ν. We now substitute (6.4.7) into (6.4.6) to obtain

$$Q_{\lambda/\mu} = \sum_\nu \left\langle Q_\lambda, \sum_\pi \left(\sum_T \Phi_T \right) P_\pi \right\rangle m_\nu = \sum_{\nu,\pi} \left(\sum_T \Phi_T \right) \langle Q_\lambda, P_\pi \rangle m_\nu$$

By orthogonality, the inner products appearing in this sum are equal to 1 if $\pi = \lambda$, and are equal to 0 otherwise. Hence

$$Q_{\lambda/\mu} = \sum_\nu \sum_T \Phi_T \, m_\nu$$

where the inner sum is taken over all skew tableaux T on λ/μ with content ν. The identity (6.4.5) follows by rearranging the sum in terms of individual monomials.

The proof of the combinatorial formula for $P_{\lambda/\mu}$ is similar, albeit using (iii) of Theorem 6.15 and interchanging P's and Q's as necessary. □

Of course, as with the analogous Theorem 5.12 for skew Schur functions, we obtain combinatorial formulas for the Macdonald functions by setting $\mu = 0$. In this case the resulting combinatorial formulas are the same, but with the sums taken over all tableaux T on λ. Likewise, formulas for (skew) Macdonald polynomials in n variables are obtained by restricting the sums to only contain tableaux with entries in $[n]$.

The difficulty in determining the coefficients $f^\nu_{\lambda\mu}$ is made clear by the formulas (6.4.5). Since the coefficients $\Psi_T(q,t)$ appearing in these formulas are relatively complicated in nature (in contrast with the Schur functions, for which each monomial was monic), we cannot reduce the problem of counting the number of times a monomial x^T appears in the product $P_\lambda P_\mu$ to counting the number of ways the tableau T factors into tableaux on λ and μ. Hence the approach taken in Section 5.4 to prove the Littlewood-Richardson rule does not carry over to the Macdonald setting.

We now conclude with some simple corollaries of Theorem 6.19. In the case of Macdonald polynomials in n variables, we have the following vanishing condition for $Q_{\lambda/\mu}(x)$ (cf. Proposition 5.7). This is in fact a condition on whether the set of skew tableaux on λ/μ with entries in $[n]$ is nonempty.

Corollary 6.20. *If λ, μ are partitions, then $Q_{\lambda/\mu}(x_1, \dots, x_n; q, t) \neq 0$ if and only if $0 \leq \lambda_i' - \mu_i' \leq n$ for each $i \geq 1$.*

In the case where $n = 1$, so that $x = (x_1)$ is a single variable, the skew Macdonald polynomials are given as follows.

Corollary 6.21. *Let $x = (x_1)$ be a single variable. We have*

$$Q_{\lambda/\mu}(x; q, t) = \begin{cases} \Phi_{\lambda/\mu}(q,t) x^{|\lambda/\mu|} & \text{if } \lambda/\mu \text{ is a horizontal strip} \\ 0 & \text{otherwise} \end{cases}$$

and similarly for $P_{\lambda/\mu}(x; q, t)$, albeit with $\Phi_{\lambda/\mu}(q,t)$ replaced by $\Psi_{\lambda/\mu}(q,t)$.

Finally, the duality result (6.2.5) extends to the skew Macdonald functions:

$$\begin{aligned} \omega_{q,t} P_{\lambda/\mu}(q,t) &= Q_{\lambda'/\mu'}(t,q) \\ \omega_{q,t} Q_{\lambda/\mu}(q,t) &= P_{\lambda'/\mu'}(t,q) \end{aligned} \tag{6.4.8}$$

Proof. We have

$$\begin{aligned} \omega_{q,t} Q_{\lambda/\mu}(q,t) &= \omega_{q,t} \sum_\nu f^\lambda_{\mu\nu} Q_\nu(q,t) \\ &= \sum_\nu f^\lambda_{\mu\nu} P_{\nu'}(t,q) \\ &= \sum_\nu f^{\lambda'}_{\mu'\nu'} b_{\mu'}(t,q) b_{\lambda'}(t,q)^{-1} Q_{\nu'}(t,q) \\ &= b_{\mu'}(t,q) b_{\lambda'}(t,q)^{-1} Q_{\lambda'/\mu'}(t,q) \\ &= P_{\lambda'/\mu'}(t,q) \end{aligned}$$

The proof of the first identity is similar. $\square$

References

[1] Mengyuan Cao. “Representation theory of Lie colour algebras and its connection with Brauer algebras”. MA thesis. Ottawa, Canada: University of Ottawa, 2018.

[2] William Fulton. *Young Tableaux: With Applications to Representation Theory and Geometry*. London Mathematical Society Student Texts 35. Cambridge University Press, 1997.

[3] Mark Haiman. “Macdonald Polynomials and Geometry”. In: *New Perspectives in Algebraic Combinatorics*. Ed. by L. J. Billera et al. Mathematical Sciences Research Institute Publications 38. Cambridge University Press, 1999, pp. 207–254.

[4] Alexander A. Kirillov Jr. “Lectures on Affine Hecke Algebras and Macdonald’s Conjectures”. In: *Bulletins of the American Mathematical Society* 34.3 (1997), pp. 251–292.

[5] I.G. Macdonald. “A New Class of Symmetric Functions”. In: *Séminaire Lotharingien de Combinatoire* 20 (1988), pp. 131–171.

[6] I.G. Macdonald. *Symmetric Functions and Hall Polynomials*. 2nd ed. Oxford University Press, 1995.

[7] Claudio Procesi. *Lie Groups: An Approach through Invariants and Representations*. Springer, 2007.

[8] Robert A. Proctor. “Equivalence of the Combinatorial and the Classical Definitions of Schur Functions”. In: *Journal of Combinatorial Theory* A51 (1989), pp. 135–137.

[9] Bruce E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*. 2nd ed. Graduate Texts in Mathematics 203. Springer, 2001.

[10] Richard P. Stanley. *Enumerative Combinatorics*. Vol. 2. Cambridge Studies in Advanced Mathematics 62. Cambridge University Press, 1999.

[11] Martha Yip. “A Littlewood-Richardson rule for Macdonald polynomials”. In: *Mathematische Zeitschrift* 272.3–4 (2012), pp. 1259–1290.