SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Symmetric Polynomials and Tableaux

av

Emil Sjöholm

2014 - No 30

Symmetric Polynomials and Tableaux

Emil Sjöholm

Självständigt arbete i matematik 15 högskolepoäng, Grundnivå

Handledare: Rikard Bögvad

2014

Symmetric Polynomials and Tableaux

Emil Sjöholm

November 7, 2014

Abstract

This paper introduces different symmetric polynomials, ways to switch between these and how they are connected to the symmetric group. Furthermore we introduce λ-tableaux and Young tableaux and why these are important for the symmetric polynomials. Schur polynomials are defined in different ways and we give some nice identities regarding them. Lastly an example is given how the symmetric polynomials can be applied to representation theory.

Contents

1 Introduction

This paper introduces symmetric polynomials and states some theorems regarding these. It also digs deeper into Schur polynomials which is a useful tool in representation theory. We talk about λ-tableux as well as Young tableaux and see how these are connected to the symmetric functions.

2 Basic knowledge

2.1 Groups and Fields

Definition 2.1.1 *The set of all bijections σ from $\{1, 2, \dots, n\}$ to itself is called the* ***symmetric group*** *and is denoted S_n. Composition is used as multiplication.*

Now we will show that S_n in fact is a group.

Theorem 2.1.1 *The symmetric group is a group.*
Proof 2.1

- We start with proving the closure axiom. If $\sigma, \theta \in S_n$ then they are both one-to-one and onto. we want to show that $\sigma\theta$ is one-to-one and onto. Let $x_1, x_2 \in S$, if $(\sigma\theta)(x_1) = (\sigma\theta)(x_2)$ then $\sigma(\theta(x_1)) = \sigma(\theta(x_2))$. Since σ, θ are one to one $x_1 = x_2$. Since they are also onto functions there exists an $x_n \in S$ such that $\theta(x_1) = x_n$ and an $x_k \in S$ such that $\sigma(x_n) = x_k$ hence $\sigma(\theta(x_1)) = x_k$ and we have shown that $\sigma\theta$ is one-to-one and onto.

- Now we prove that the composition is associative. Let $\sigma, \theta, \tau \in S_n$, then for each $x_k \in S$, we have

$$(\sigma(\theta\tau))(x_k) = \sigma\theta\tau(x_k) = (\sigma\theta)\tau(x_k).$$

  Therefore the composition is associative.

- It is clear that the identity function is used as the identity.

- Lastly we prove that the inverse exists. This is clear since every function $\sigma \in S_n$ will rearrange the posts in S. Then the inverse is the function rearranging the posts in reverse order. This will obviously be one-to-one and onto since σ is one-to-one and onto. □

If $\sigma, \theta \in S_n$ then $\sigma\theta$ is interpreted as first performing θ and then σ. The symmetric group rearranges, or permutes, variables in a set. An example of a symmetric group is $S_3 = \{(), (1, 2), (1, 3), (2, 3), (1, 2, 3), (1, 3, 2)\}$ where $(1, 2)$ means element 1 and 2 are interchanged while nothing happens with the third element. On the other hand $(1, 2, 3)$ means that $1 \to 2 \to 3 \to 1$. Note that $(1, 2, 3)$ is the same as $(1, 3)(1, 2)$.

Definition 2.1.2 *The* ***length of a cycle*** *$\sigma \in S_n$ is defined as the number of variables it contain. One say that $(1, 2, 4)(3, 5)(6)$ contain a three-cycle namely $(1, 2, 4)$ a two-cycle $(3, 5)$ and a one-cycle (6). A two-cycle is also called a* ***transposition****.*

Theorem 2.1.2 *Any permutation in S_n can be written as a product of disjoint cycles.*

Proof 2.2 Let $S = \{1, 2, \dots, n\}$ and $\sigma \in S_n$. Now we calculate the powers of $\sigma(1)$, that is $1, \sigma(1), \sigma^2(1), \sigma^3(1), \dots$ until we reach a repetition. We will of course reach a repetition since there are n elements in S. Suppose that

$\sigma^j(1) = \sigma^k(1)$ and that $k > j \geq 0$. If $j > 0$, by multiplying both sides j times with σ^{-1} we calculate $\sigma^{k-j}(1) = 1$. This will of course be an earlier repetition, hence the first repetition is when we are back at 1. Now we let m be the smallest positive integer such that $\sigma^m(1) = 1$, then all the elements $1, \sigma(1), \ldots, \sigma^{m-1}(1)$ are distinct and form a cycle of length m. If $m = n$ we are done, otherwise we choose the smallest integer a not existing in the aforementioned cycle. Now we form the cycle $a, \sigma(a), \ldots, \sigma^{r-1}(a)$ where $a = \sigma^r(a)$, note that we can do this by the same reasoning as above. If now $r + m < n$ we select a b not yet appearing and we continue doing this until we have $\sigma = (1, \sigma(1), \ldots, \sigma^m - 1(1))(a, \sigma(a), \ldots, \sigma^r - 1(a))(b, \sigma(b), \ldots, \sigma^s - 1(b)) \ldots$. Now we have written σ as a product of disjoint cycles. $\square$

Theorem 2.1.3 *Any permutation in S_n where $n \geq 2$, can be written as a product of transpositions.*

Proof 2.3 The identity can always be expressed as $1 = (1, 2)(1, 2)$. Due to 2.1.5 we only have to prove it for every cycle. Take an arbitrary cycle $\sigma \in S_n$, on the form $(a_1, a_2, \ldots, a_k)$. Now consider the transpositions $(a_k, a_1)(a_k, a_2) \ldots (a_k, a_n)$. Performing this will obviously yield $(a_k, a_1, \ldots, a_{k-1})$ which is the cycle as we wanted. $\square$

Now we want to introduce a useful function, but first we have to make sure it is well-defined.

Theorem 2.1.4 *If a permutation σ is written as a product of transpositions in two ways, then the number of transpositions is either even in both cases or odd in both cases.*

Proof 2.4 By assuming the theorem is false we construct a proof by contradiction. Then there exists a permutation σ that can be written as $\sigma = \tau_1\tau_2 \ldots \tau_{2m} = \theta_1\theta_2 \ldots \theta_{2n+1}$ where τ and θ are transposition and $m, n \in \mathbb{N}$. The inverse of the very right hand side is of course $\theta_{2n+1}\theta_{2n} \ldots \theta_1$ since every transposition is its own inverse. Using the first equality sign and multiplying both sides with σ^{-1} we get $\sigma\sigma^{-1} = \tau_1\tau_2 \ldots \tau_{2m}\theta_{2n+1}\theta_{2n} \ldots \theta_1$. Because $\sigma\sigma^{-1} = 1$ we have shown that the identity can be written as a product of an odd number of transpositions. If we can falsify this statement we are done with the proof.
Suppose that (1) $\delta_1\delta_2 \ldots \delta_r$ is the shortest product of an odd number of transpositions that is equal to the identity, of course $k > 2$. Say that $\delta_1 = (a, b)$, note that a must appear at least once more in (1) since otherwise a will go onto b and (1) can not be the identity. Among all products of length k that are equal to the identity, and such that a appears in the transposition on the left, we assume that (1) has the fewest number of $a's$
Now we show that if δ_i is the transposition of the smallest index i in which a occours, then δ_i can be moved to the left without changing the number of times a occours in (1) nor the number of transpositions in (1). Combining δ_1 with δ_i will then lead to a contradiction, namely that we did not select the product consisting of the fewest $a's$.
Let a, x, y, z be distinct. Simple computations show us that $(x, y)(a, z) = (a, z)(x, y)$ and $(x, y)(a, y) = (a, x)(x, y)$, note that those are the only cases possible containing a since $(a, y) = (y, a)$. Hence we can move δ_i to position 2 without changing the number of times a appears or the number of transpositions.

Therefore we can assume that δ_2 contains an a as well. But now we have either the case $(a,b)(a,c)\dots\delta_r$ or $(a,b)(a,b)\dots\delta_r$, in the second case the two first product cancel each other out and the assumption that we have the lowest possible r is wrong. In the first case we calculate $(a,b)(a,c) = (a,c)(b,c)$ hence we can make a sequence of fewer $a's$. Both cases gives us a contradiction therefore we can not write the identity as an odd number of transposition and the proof is complete. $\square$

This result gives sense to the following definition.

Definition 2.1.3 *A permutation $\sigma \in S_n$ is defined as even if it can be written with as a composition of an even number of transpositions. The permutation is odd if it can be written as an odd number of transpositions.*
*The **signature** of a permutation σ is defined below as:*

$$Sgn(\sigma) = \begin{cases} 1, & if\ \sigma\ is\ even \\ -1, & if\ \sigma\ is\ odd \end{cases}$$

Due to theorem 2.1.4 this is well-defined. One can note that the product of two even permutations will be even, that the product of two odd permutations will be even and that the product of an odd and even permutation will be odd, this is not of great importance in this paper.

Example 2.1 We want to calculate $Sgn(\sigma)$ if $\sigma = (1,2,3)$. We only have to find one example of a product of transpositions that is equal to σ due to 2.1.7. The shortest example is $(1,2,3) = (1,3)(1,2)$ where we have an even number of transpositions, hence $Sgn(\sigma) = 1$.
Definition 2.1.4 *The* **alternating group**, denoted A_n, is the group consisting of all even permutation in n variables.

That this is a group can be shown in a similar fashion to why S_n is a group, while remembering that the composition of two even permutation is even. [1].

2.2 Integer partitions

Definition 2.2.1 *An ordered set $\alpha = \alpha_1, \alpha_2, ..., \alpha_n$ is said to be **weakly decreasing** if $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$.*
*It is **weakly increasing** if $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$.*
It is **strictly decreasing** if $\alpha_1 > \alpha_2 > \cdots > \alpha_n$.
It is **strictly increasing** if $\alpha_1 < \alpha_2 < \cdots < \alpha_n$.

Definition 2.2.2 *Write an integer β as a sum of non-negative integers $\alpha_1 + \alpha_2 + \cdots + \alpha_n = \alpha$, then α is a **partition** of β.*
*If α is weakly decreasing then α is said to be an **ordered partition** of β and denoted $\alpha \vdash \beta$.*

Definition 2.2.3 *The* ***length of an ordered partition*** $Y = (y_1, y_2, \ldots, y_n)$, *is denoted* $\ell(Y)$ *and defined as the largest integer* k *such that* $y_k \neq 0$.

Example 2.2 For example an ordered partition Y is $(4, 2, 1, 0, 0) \vdash 7$ since it is ordered and $4 + 2 + 1 + 0 + 0 = 7$. $\ell(Y) = 3$ since $y_3 \neq 0$ and $y_4 = 0$ [2].

3 Basic symmetric Polynomials

3.1 Symmetric polynomials

A symmetric polynomial is a polynomial in several variables where if you permute some or all of the variables the polynomial always remain the same. This is useful when the specific variables does not matter, only how many there are. Below every function is a polynomial if otherwise is not stated explicitly.

Definition 3.1.1 *$f(x_1, x_2, \dots, x_n)$ is a **symmetric polynomial** if*

$$\forall \sigma \in S_n, f(x_1, x_2, \dots, x_n) = f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)})$$

From this definition it is clear that $f(x_1, x_2) = x_1x_2$ is a symmetric polynomial but $f(x_1, x_2, x_3) = x_1x_2$ is not, since interchanging x_2 with x_3 will yield a different polynomial.

Definition 3.1.2 *$A[x_1, x_2, \dots, x_n]^{S_n}$ is the **set of all symmetric polynomials** in n variables with coefficients in some field A.*

Example 3.1.1 Consider $A[x_1, x_2]^{S_n}$ where $A = \{0, 1\}$. This is an infinite set since $x_1^k + x_2^k$ is a symmetric polynomial for any k and clearly belongs to this set.

3.2 Monomials

Definition 3.2.1 *A polynomial in several variables is called a **monomial** if it has only one term.*

E.g x_1x_2 is a monomial but $x_1 + x_2$ is not, however x_1 and x_2 separately are monomials.
Monomials are in general not symmetric, the only monomial that is symmetric in n variables is

$$x_1x_2 \dots x_n$$

It is very easy to make a monomial into a symmetric polynomial as we shall see now.

3.3 Monomial symmetric polynomials

The monomial symmetric polynomials provide us with an easier notation for some symmetric polynomials. These are also the easiest of symmetric polynomials.

Definition 3.3.1 *Consider an ordered partition $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$, then the* ***monomial symmetric polynomial*** *corresponding to α is*

$$m_\alpha(X) = \sum_g x_{g(1)}^{\alpha_1} x_{g(2)}^{\alpha_2} \dots x_{g(n)}^{\alpha_n}$$

where the sum is over all $g \in S_n$. This process is called that we symmetrise a monomial.

It is fairly obvious that every symmetric polynomial can be written as sums of monomial symmetric polynomials and we will state this and prove it in chapter 3.7. Note that $\ell(\alpha)$ can equal n but that is not necessary, see the example below in three variables and consider $\alpha = \{3, 2, 0\}$.

Example 3.3.1 Consider the partition $\alpha = \{3, 2\}$ then

$$\begin{aligned} m_\alpha(x_1, x_2) &= x_1^3 x_2^2 + x_1^2 x_2^3 \\ m_\alpha(x_1, x_2, x_3) &= x_1^3 x_2^2 + x_1^3 x_3^2 + x_1^2 x_2^3 + x_1^2 x_3^3 + x_2^3 x_3^2 + x_2^2 x_3^3 \end{aligned}$$

Example 3.3.2 We want to rewrite $f(x_1, x_2) = x_1 + x_1^3 x_2 + x_2 + x_1 x_2^3$ as a sum of monomial symmetric polynomials. We start by sorting the polynomials into their symmetric parts, $f(x_1, x_2) = (x_1 + x_2) + (x_1^3 x_2 + x_1 x_2^3)$. Here we easily see that $f(x_1, x_2) = m_{(1,1)}(x_1, x_2) + m_{(3,1)}(x_1, x_2)$

It does not actually matter if α is ordered or not since it will yield the same set of symmetric polynomials.

3.4 Elementary symmetric polynomials

The elementary symmetric polynomials in n variables is a set of generators for all symmetric polynomials in n variables, in other words all symmetric polynomials can be expressed as polynomials in the elementary symmetric polynomials. This is called the fundamental theorem of symmetric polynomials and will be discussed in the next section.

Definition 3.4.1 *The* ***elementary symmetric polynomials*** *in n variables are:*

$$\begin{aligned} e_0(x_1, x_2, \dots, x_n) &= 1 \\ e_1(x_1, x_2, \dots, x_n) &= x_1 + x_2 + \dots + x_n \\ e_2(x_1, x_2, \dots, x_n) &= \sum_{1 \le i < j \le n} x_i x_j \\ &\vdots \\ e_n(x_1, x_2, \dots, x_n) &= x_1 x_2 \dots x_n \end{aligned}$$

For k$\notin \{0, 1, \ldots, n\}$ we define $e_k(x_1, x_2, \ldots, x_n) = 0$.

These can easily be expressed as monomial symmetric functions. For example

$$e_2(x_1, x_2, \ldots, x_n) = m_{(1,1)}(x_1, x_2, \ldots, x_n) \text{ and}$$
$$e_n(x_1, x_2, \ldots, x_n) = m_{\underbrace{(1,1,\ldots,1)}_{n\ ones}}(x_1, x_2, \ldots, x_n).$$

[3].

Definition 3.4.2 The **complete homogeneous symmetric polynomials** $h_k(x_1, x_2, \ldots, x_n)$ are defined exactly as the elementary symmetric polynomials except for the "$<$" being replaced with "$\leq$".

3.5 The fundamental theorem of symmetric polynomials

Theorem: 3.5.1 *Any symmetric polynomial can be expressed uniquely as a polynomial in the elementary symmetric polynomials on those variables*

There are many ways to prove the fundamental theorem of symmetric polynomials and this proof gives us a way to determine how to write our symmetric polynomial using only $e_0, e_1, \ldots, e_n$.

Before the proof we need to introduce a notation.

Defintion: 3.5.1 *The* ***lexicographic order*** *of two monomials*

$$p_1 = c_1 x_1^{g(1)} x_2^{g(2)} \ldots x_n^{g(n)} \text{ and } p_2 = c_2 x_1^{h(1)} x_2^{h(2)} \ldots x_n^{h(n)}$$

is defined as $p_1 < p_2$ *if* $\exists 1 \leq t \leq n\ \forall s > t$ *such that* $g(t) < h(t)$ *and* $g(s) = h(s)$ *and* $c_1, c_2 \in \mathbb{R} \setminus \{0\}$

In other words the monomial with the largest exponent the furthest to the right has the highest lexicographic order, for example $5x_1^2x_2^3x_3^3 < 3x_1^2x_2^2x_3^4$ since the exponent of x_3 satisfy $3 < 4$ and $5x_1^2x_2^3x_3^3 > 3x_1^2x_2^2x_3^3$ since x_3 is a tie (they have the same power). Therefore we have to check the position one step to the left, here the exponent of x_2 satisfy $3 > 2$. The lexicographic order has nothing to do with the constants c_1, c_2 as long as they are nonzero. Any polynomial is of course a sum of such monomials.

The way we rewrite our symmetric polynomial P is that we subtract a product of elementary polynomials, making the largest monomial (lexicographically speaking) disappear until there are no polynomial left. We do it in such a fashion that no monomial grows larger than the original monomial, hence making the largest monomial smaller and smaller. Below we will give the proof.

Proof: 3.1 Consider a symmetric polynomial P in n variables.
Pick the largest monomial $cx_1^{g(1)}x_2^{g(2)} \ldots x_n^{g(n)}$ in P. Notice that $g(1) \leq g(2) \leq$

$\cdots \leq g(n)$, otherwise we have not selected the largest monomial (some permutation of the variables would clearly have higher lexicographic order if this was not the case).
Let us introduce a new monomial Q, we define $Q = ce_1^{g(n)-g(n-1)} e_2^{g(n-1)-g(n-2)} \dots e_n^{g(1)}$ where e_k are the elementary symmetric polynomials in n variables. Q is written as a product of elementary symmetric polynomials.
Now we only need to prove that $cx_1^{g(1)} x_2^{g(2)} \dots x_n^{g(n)}$ is the largest monomial in Q and then $P - Q$ will be a symmetric polynomial with lower lexicographic order then P. Hence we need to iterate this step until we have $P - R - Q - S \cdots = 0$ where $Q, R, S \dots$ consists of powers of elementary symmetric polynomials. Then we can write $P = Q + R + S + \dots$.
The largest monomial of Q is

$$c * (largest\ monomial\ of\ e_1)^{g(n)-g(n-1)} \dots (largest\ monomial\ of\ e_n)^{g(1)}$$

since we have picked the largest x_n term from every monomial in e_k, $k = 1, 2, \dots, n$. In this monomial the variable x_n has the exponent $g(n)$ since $g(n) - g(n-1) + g(n-1) - g(n-2) + \dots - g(1) + g(1) = g(n)$.

The variable x_{n-1} will be in every term of the largest monomial in Q except the one from e_1. Therefore x_{n-1} will have the exponent $g(n-1) - g(n-2) + \dots - g(1) + g(1) = g(n-1)$. In much the same way x_{n-2} will be $g(n-2) - g(n-3) + \dots - g(1) + g(1)$ and so on for the rest of the variables $x_3, x_4, \dots, x_n$. Hence the largest monomial in Q is equal to the largest monomial in P. The reasoning above shows that we have eliminated the largest monomial and can continue doing this until $P - R - Q - S \cdots = 0$. What is left to prove is that this expression is unique, or even stronger, the expression for the monomial of highest order is unique. In other words, the way we define Q above is the only way to cancel out the highest order without making the lexicographic order larger. Assuming there is another way we will try to construct it, e_n can be used at most $g(1)$ times in our product since if it is used more, x_1 will be too large for our largest monomial. But e_n must be used at least $g(1)$ times in our product, since it can not get any more contributions because then the created monomial will not have the largest order. Using the same reasoning for e_{n-1} and so on we see that Q is the only way of making this reduction. $\square$

We will give an easy example on how the algorithm works.

Example 3.5.1 Say

$$P(x_1, x_2) = m_{(1,1)}(x_1, x_2) + m_2(x_1, x_2) = x_1 x_2 + x_1^2 + x_2^2.$$

The largest monomial is x_2^2, therefore we pick

$$Q = e_1^{(2-0)} e_2^0 = (x_1 + x_2)^2 = x_1^2 + 2x_1 x_2 + x_2^2$$

Note that x_2^2 is the largest monomial. Now we calculate

$$P - Q = x_1^2 + x_1 x_2 + x_2^2 - (x_1^2 + 2x_1 x_2 + x_2^2) = -x_1 x_2$$

According to our algorithm it is now time to start over on $-x_1x_2$.

$$R = -e_1^{(1-1)}e_2^1 = -e_2.$$

Calculating $P-Q-R$ and see that this is zero, we know that $P = Q+R = e_1^2-e_2$

3.6 Power-sum symmetric polynomials

Definition 3.6.1 *The* ***Power-sum symmetric polynomial*** *are denoted*

$$P_k(x_1, x_2, \dots, x_n) = x_1^k + x_2^k + \cdots + x_n^k$$

It is easily seen that $P_1 = e_1$ as long as they have the same number of variables. Because of the fundamental theorem of symmetric polynomials we know that every Power-sum symmetric polynomial can be written as a linear combination in terms of e_n in exactly one way.

Example 3.6.1 The monic polynomial in two variables $m_{2,1} = x_1^2x_2 + x_2^2x_1$ can be written as the quotient

$$\frac{P_1^3 - P_1P_2}{2} = \frac{x_1^3 + 3x_1^2x_2 + 3x_1x_2^2 + x_2^3 - x_1^3 - x_1^2x_2^1 - x_1x_2^2 - x_2^3}{2} = x_1^2x_2+x_2^2x_1$$

3.7 Newton identities

The Newton Identities gives a way to translate between the elementary symmetric polynomials and the power-sum symmetric polynomials.

Theorem 3.7.1 *The* ***Newton identities*** *in n variables is written as*

$$\sum_{i=1}^{k} \big((-1)^{i-1}e_{k-i}(x_1, x_2, \dots, x_n)P_i(x_1, x_2, \dots, x_n) - e_k(x_1, x_2, \dots, x_n)\big) = 0$$

The Newton identity gives us a tool of translating from elementary symmetric polynomials to power-sum symmetric polynomials. We will first do an example of how to apply the theorem before we prove it.

We give a table of the first five Newton identities in at least 5 variables.

$$\begin{aligned}
e_1 &= p_1 \\
2e_2 &= e_1p_1 - p_2 \\
3e_3 &= e_2p_1 - e_1p_2 + p_3 \\
4e_4 &= e_3p_1 - e_2p_2 + e_1p_3 - p_4 \\
5e_5 &= e_4p_1 - e_3p_2 + e_2p_3 - e_1p_4 + p_5
\end{aligned}$$

Dividing the second row with two, the third row with three and so on we have expressions for e_1, e_2, e_3, e_4, e_5. Switching p_2 with $2e_2$, p_3 with $3e_3$ and so on we get the following table:

$$\begin{aligned}
p_1 &= e_1 \\
p_2 &= e_1 p_1 - 2e_2 \\
p_3 &= -e_2 p_1 + e_1 p_2 + 3e_3 \\
p_4 &= e_3 p_1 - e_2 p_2 + e_1 p_3 - 4e_4 \\
p_5 &= -e_4 p_1 + e_3 p_2 - e_2 p_3 + e_1 p_4 + 5e_5
\end{aligned}$$

Here we have an expression for the power-sum symmetric polynomials using only the elementary symmetric functions and note that this is recursive; to calculate p_5 you must know p_1, p_2, p_3 *and* p_4. The recursive part makes it harder to calculate p_k the larger k is.

Here is an inductive proof of the Newton identities.

Proof 3.2 Let x be a variable and $f(x) = \prod_{i=1}^{n}(x_i - x)$. Then

$$f(x) = -x^n + b_1 x^{n-1} - b_2 x^{n-2} + \cdots + b_n \ for\ odd\ n\ and$$

$$f(x) = -x^n + b_1 x^{n-1} - b_2 x^{n-2} + \cdots - b_n \ for\ even\ n.$$

We see that

$$b_i x^{n-i} = \sum_{1 \le a < b < \cdots < i \le n} x_a x_b \ldots x_i x^{n-i} = x^{n-i} \sum_{1 \le a < b < \cdots < i \le n} x_a x_b \ldots x_i,$$

but the last sum is the definition of the elementary symmetric polynomial. Hence we know that $b_i = e_i$.
Denote the left hand side of the equation in theorem 3.7.1 by $F_n^{(k)}(x_1, \ldots, x_n)$. We want to show that $F_n^{(k)}(x_1, \ldots, x_n) = 0$ for every $n, k \in \mathbb{N}$. Firstly we assume $k \ge n$. Summing $f(x_1) + f(x_2) + \cdots + f(x_n)$ we see that each function is equal to zero and hence the summand is equal to zero, but if we sum term by term we get

$$0 = f(x_1) + f(x_2) + \cdots + f(x_n) =$$

$$= \sum_{i=1}^{k} \big((-1)^{i-1} e_{k-i}(x_1, x_2, \ldots, x_n) P_i(x_1, x_2, \ldots, x_n) - e_k(x_1, x_2, \ldots, x_n)\big)$$

which is the Newton identity. This is the same as $F_n^{(n)} = 0$. In a similar way we expand $x_1^{k-n} f(x_1) + \cdots + x_n^{k-n} f(x_n)$, since $k \ge n$, and conclude that this is 0, hence we have shown that $F_n^{(k)} = 0$ for all $k \ge n$.
Now comes the inductive step, choose $n, k \in \mathbb{N}$ such that $n > k$. Assume that $F_{n'}^{(k')} = 0$ for all n', k' such that $n' - k' < n - k$. Firstly note that if $x_n = 0$ then $F_n^{(k)}(x_1, \ldots, x_n) = F_{n-1}^{(k)}(x_1, \ldots, x_{n-1})$. By the induction assumption

$F_{n-1}^{(k)}(x_1, \ldots, x_{n-1}) = 0$, this means that $F_n^{(k)}$ is divisible by x_n. Since $F_n^{(k)}$ is a symmetric polynomial it is divisible by $x_1 x_2 \ldots x_n$ and has to be equal to zero since the degree of the polynomial is $k < n$ (a polynomial can not be divisible by another polynomial of higher degree). Thus $F_n^{(k)} = 0$, which was the claim of our Newton identities.

4 Tableaux

A tableau is a two-dimensional grid with squares that are either empty or filled with something, in this section usually numbers. We will discuss different types of tableaux, worth noting is that *tableaux* is plural while *tableau* is singular.

4.1 Young tableaux

Definition 4.1.1 *For an ordered partition $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ a **λ-tableau** is a grid with first row consisting of α_1 entries, second row consisting of α_2 entries until the nth row that consists of α_n entries. The tableau is said to have shape α. Entries in a λ-tableau are called cells, the cell in row i and column j is denoted (i, j)*

Definition 4.1.2 *A **semistandard Young tableau (SSYT)** T of the ordered partition $\alpha \vdash n$ is a filling of the cells in the λ tabelau with numbers that increase weakly when you go to the right along the rows and increase strictly down each column.*
*If the filling is made using each of the integers {1,2,...,n} exactly once then T is said to be a **standard Young tableau**.*

Example 4.1.2 Consider the partition {5,3,2}. Below is the empty λ-tableau and a semi standard Young tableau:

1	1	3	4	4
3	5	6		
6	6			

Notice how every row increases weakly and every column increases strictly, moreover the number 2 is not represented in the SSYT.

Definition 4.1.3 *Consider a SSYT T. Then x^T is defined as:*

$$x_1^{number\ of\ ones} x_2^{number\ of\ twos} \dots x_n^{number\ of\ ns}$$

It is easy to see that the sum of the powers will be equal to the partitioned number. x^T is always a monic polynomial.

Example 4.1.2 Using the SSYT from example 4.2 we get the polynomial

$$x_1^2 x_3^2 x_4^2 x_5 x_6^3$$

The sum of the powers are $2 + 2 + 2 + 1 + 3 = 10 = 5 + 3 + 2$ which is the partition of the number, that we started with.

4.2 More about tableaux

Definition 4.2.1 *The* ***weight*** μ *of a SSYT T is:*

$$\mu = \{\#1's \in T, \#s'2 \in T, \ldots, \#n's \in T\} = \{\mu_1, \mu_2, \ldots, \mu_n\}$$

Note that if T has shape λ then both λ and μ must be partitions of the same number, further note that μ does not have to be an ordered partition.

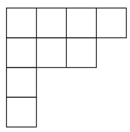
Theorem 4.2.1 If μ is a weight of a SSYT T of shape λ then $\ell(\mu) \geq \ell(\lambda)$.

Proof 4.1
We only need to show that if $\ell(\mu) < \ell(\lambda)$ then it can not be a weight of the SSYT of shape λ. Since μ has fewer elements than λ has rows, some element from μ have to occur in the first column more than once according to the pigeon hole principle, hence T cannot be semistandard. $\square$

Definition 4.2.2 The conjugate of a partition λ is denoted λ' and is the partition consisting of the column lengths of λ's shape.

It is obvious that λ' will be an ordered partition as long as λ is ordered, moreover they will be partitions of the same number.

Example 4.2.1 Consider $\lambda = (4, 3, 1, 1)$ which has the shape



Counting the cells in the columns we see that $\lambda' = (4, 2, 2, 1)$

Definition 4.2.3 *The* ***Kostka number*** $k_{\lambda\mu}$ *counts the number of SSYT of shape λ with weight μ.*

There are no easy formulas found yet how to compute the Kostka numbers.

Example 4.2.2 Consider any ordered partition λ. Then $k_{\lambda\lambda} = 1$.
Since there are λ_1 1's they all have to be in the first row since the columns have to be strictly increasing, this fills up the entire first row. Now only rows $2, 3, \ldots, n$ are left and we can use the same reasoning for row $2, 3$ until we are at row n. This shows that there is exactly one filling possible.

Example 4.2.3 We want to calculate $k_{\lambda\mu}$ where $\lambda = \{3, 2\}$ and $\mu = \{2, 1, 2\}$. First we draw the λ-tableau:

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

It is clear that the two ones has to be in the first two posts in the 1st row $\begin{array}{|c|c|c|}\hline 1 & 1 & x \\ \hline x & x \\ \cline{1-2}\end{array}$. Now we can place the 2 in two different spots and fill in the threes in the rest. This is how the two SSYT will look like

$$\begin{array}{|c|c|c|}\hline 1 & 1 & 2 \\ \hline 3 & 3 \\ \cline{1-2}\end{array} \quad \begin{array}{|c|c|c|}\hline 1 & 1 & 3 \\ \hline 2 & 3 \\ \cline{1-2}\end{array}$$

Hence we have found that $K_{(3,2)(2,1,2)} = 2$

Definition 4.2.4 *The **hook** $H_\lambda(i,j)$ of a cell in the i:th row and j:th column of a tableau λ is the set of cells (a,b) such that $a = i$ and $b \leq j$, or $a \leq i$ and $b = j$.*

Definition 4.2.5 *The **hook length** $h_\lambda(i,j)$ of a cell (i,j) in a tableau λ is the number of elements in the hook $H_\lambda(i,j)$, in other words it counts the elements strictly to the right of the given cell plus the elements strictly below the given cell plus itself.*

Theorem 4.2.2 The cell $(1,1)$ will always have the greatest hook length.

To prove this we have to state a lemma, the proof comes directly afterwards.
Lemma 1 From the cell $(1,1)$ you can walk to any other cell using only movements to the right and downwards.
Proof 4.2 If you want to walk to cell (i,j) you firstly walk downwards to row i, then walk right until you are in column j, this is always possible since there are no "holes" in the grid. $\square$

Proof 4.3 Pick an arbitrary cell (i,j). Now we want to show that $h_\lambda(i,j) > h_\lambda(i+1,j)$ and $h_\lambda(i,j) > h_\lambda(i,j+1)$ if those cells exists. This will show that going to the right or downwards will decrease the hooklength. We start to show the first inequality: $(i+1,j)$ have one less square below itself than (i,j) and at most as many squares to the right, hence the hooklength must be lower. On the other hand $(i,j+1)$ have one less square to the right then (i,j) and at most as many squares below, hence the hooklength also here has to be lower. This together with lemma 1 tells us that $(1,1)$ has larger hooklength then any other cell. $\square$

Example 4.2.4 Consider the $\lambda-$*tableau* given by the partition $\lambda = \{5,5,3,3,2\}$ Now we want to calculate $H_\lambda(i,j)$ and $h_\lambda(i,j)$ for $i = 1$ and $j = 2$. The tableau looks as follows with the square $(1,2)$ marked with x, $(1,2+1)$ marked with y and $1+1,2$ marked with z. All other points in the set $H_\lambda(i,j)$ are marked with a 1, squares not in the set are marked with a zero.

0	x	y	1	1
0	z	0	0	0
0	1	0		
0	1	0		
0	1			

Hence $H_\lambda(i,j) = \{(1,2),(1,3),(1,4),(1,5),(2,2),(2,3),(2,4),(2,5)\}$ counting the elements give us $h_\lambda(i,j)$=8

Filling all the cells of λ with their hooklength looks like:

9	8	6	3	2
8	7	5	2	1
5	4	2		
4	3	1		
2	1			

Definition 4.2.6 *A **corner** of a λ-tableau is a cell (i,j) such that $h_\lambda(i,j)$=1*

Theorem 4.2.3 *There are exactly as many corners in a λ-tableau as there are distinct integers in the ordered partition λ.*

Proof 4.4 Looking at the top row and the row below, there are two possible cases, either they contain equally many cells or the upper row contain more cells. We want to determine if there are any corners in the upper row. Only the cell most to the right can be a corner since the other cells in the row have that cell in its hook. If the row below have shorter length then we have found a corner since our cell has no cell below nor to the right. But if the row below is equally long, then our cell has at least one cell below and its hook cannot consist of only 1 cell. In this way different length of rows will create a new corner while the same length will not. Hence only the distinct numbers will create a new corner. □

Example 4.2.5 The λ-tableau of shape $\{7,7,7,4,3,3,1,1\}$ will have 4 corners since the distinct integers are $\{7,4,3,1\}$. Below the corners are marked with 1's and the rest of the cells marked with zeroes.

0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	1
0	0	0	1			
0	0	0				
0	0	1				
0						
1						

Theorem 4.2.4 *The hook length formula*

$$d_\lambda = \frac{n!}{\prod h_\lambda(i,j)}$$

where the product is over all cells in the λ-tableau, counts all possible standard

Young tableaux corresponding to λ. Note that $\lambda \vdash n$. [4]

An elementary proof is found in [5], although it is very long.

Example 4.6 Using the same tableau as in example 4.2.5 we see that

$$d_\lambda = \frac{(5+5+3+3+2)!}{9*8*6*3*2*8*7*5*2*1*5*4*2*4*3*1*2*1} = 4594590$$

Hence we can create 4593590 standard Young tableaux with the partition {5,5,3,3,2}.

5 Schur polynomials

Here we will introduce another symmetric polynomial, one that is closely related to representation theory. Since it is rather technical we have to go through some definitions first.

5.1 More tools

An antisymmetric polynomial is not a symmetric polynomial but it is used in one of the definitions of a Schur polynomial.

Definition 5.1.1 *$f(x_1, x_2, \ldots, x_n)$ is an **antisymmetric** polynomial if*

$$\forall \sigma \in S_n,\ f(x_1, x_2, \ldots, x_n) = sgn(\sigma) f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}) \text{ or equally}$$
$$f(x_1, x_2, \ldots, x_n) sgn(\sigma) = f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}).$$

From this definition follows that if $f(x_1, x_2, \ldots, x_n)$ is an antisymmetric polynomial then:

$$\forall \sigma \in A_n, f(x_1, x_2, \ldots, x_n) = f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}).$$

In some literature the word skewsymmetric is used instead of antisymmetric.

Example 5.1 Consider the polynomial

$$f(x_1, x_2, x_3) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3),$$

which is clearly antisymmetric. For example the odd permutation $\sigma = (1\ 2)$ makes the function change sign.

$$f(x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)}) = (x_{\sigma(1)} - x_{\sigma(2)})(x_{\sigma(1)} - x_{\sigma(3)})(x_{\sigma(2)} - x_{\sigma(3)}) =$$
$$= (x_2 - x_1)(x_2 - x_3)(x_1 - x_3) = sgn(\sigma) f(x_1, x_2, x_3)$$

This is an example of a Vandermonde polynomial in three variables, which we will now introduce.

5.2 Vandermonde Polynomials

The Vandermonde Polynomials are important for defining Schur polynomials, even though they are not symmetric, but antisymmetric.

Definition 5.2.1 ***Vandermonde Polynomials** are defined as the product*

$$V_n = \prod_{1 \le i < j \le n} (x_i - x_j).$$

To see that odd permutations will make the polynomial change sign and that even permutations will make the polynomial remain the same we use any transposition $\sigma \in S_n$. Say that σ permutes i and j, in the product they will both occur paired up with every other variable exactly once. Assume $i > j$, for indexes $k < i$ the product will not change and the same is true for indexes $k > j$. For indexes $i < k < j$ we will instead of $(x_i - x_k)$ and $(x_k - x_j)$ get $(x_j - x_k)$ and $(x_k - x_i)$. This rewriting is exactly the same since we take a product. Hence the only thing that will make this polynomial change is the substitution of the term $(x_i - x_j)$ to $(x_j - x_i) = -(x_i - x_j)$, therefore we see that the polynomial changes sign under an odd permutation and remains the same under an even permutation. Recall that an even permutation can always be written as a product of an even number of transpositions.

5.3 Schur definition

Firstly we introduce $A_\alpha\{x_1, x_2, \ldots, x_n\}$

Defintion 5.3.1 Consider an ordered partition

$$\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}.$$

Define $A_\alpha(x_1, x_2, \ldots, x_n) = \begin{vmatrix} x_1^{\alpha_1+n-1} & x_2^{\alpha_1+n-1} & \ldots & x_n^{\alpha_1+n-1} \\ x_1^{\alpha_2+n-2} & x_2^{\alpha_2+n-2} & \ldots & x_n^{\alpha_2+n-2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{\alpha_n} & x_2^{\alpha_n} & \ldots & x_n^{\alpha_n} \end{vmatrix}$

Note that $A_0(x_1, \ldots, x_n) = V_n$. There are many ways to define Schur polynomials, of course they all are equivalent. We will state one as the definition and an other as a theorem.

Definition 5.3.2 *Consider an ordered partition α: $\ell(\alpha) \leq n$, then the* ***Schur polynomial*** *is defined as* $S_\alpha\{x_1, x_2 \ldots, x_n\} = \sum\limits_{T \text{ of shape } \alpha} x^T$, *where the sum is over all SSYT of shape α.*

Example 5.2 Consider the partition $\alpha = (4, 1)$ and calculate the Schur polynomial in two variables. First we need to find all T of shape α there are four.

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

Now we calculate $S_{(4,1)}(x_1, x_2) = x_1^4 x_2 + x_1^3 x_2^2 + x_1^2 x_2^3 + x_1 x_2^4$

Now we will state an alternative definition of the Schur polynomials. We will do it as a theorem.

Theorem 5.3.2 Schur polynomials are symmetric polynomials.

Proof 5.1 It is enough to show that

$$(i, i+1)s_\lambda(X) = s_\lambda(X) \qquad (*)$$

for every adjacent transposition, since if we for example want to transpose i and $i+2$ we can just transpose $(i, i+1)$ then $(i+1, i+2)$ and lastly $(i, i+1)$ again. We will give an explicit proof of $(*)$.
Given T, each column contains either an $i, i+1$ pair; exactly one of $i, i+1$; or neither. Call the pairs fixed and all other i or $i+1$ free. In each row switch the number of free i's and $(i+1)$'s. Say that $i = 2$ and

$$T = \begin{ytableau} 1 & 1 & 1 & 2 & 2 & 2 & 3 \\ 2 & 2 & 3 & 3 \\ 3 \end{ytableau}$$

then the twos and threes in column $2, 3$ and $5, 6, 7$ are free and hence our new tableau, call it T' will be:

$$T' = \begin{ytableau} 1 & 1 & 1 & 2 & 2 & 3 & 3 \\ 2 & 2 & 3 & 3 \\ 3 \end{ytableau}$$

The new tableau is still semistandard by the definition of free. Since the fixed i's and $(i+1)$'s come in pairs this map has the desired exchange property. $\square$

Example 5.3 Given $X = (x_1, x_2, x_3)$ and a SSYT

$$T = \begin{ytableau} 1 & 1 & 1 & 2 \\ 2 & 2 & 3 \\ 3 \end{ytableau}$$

we want to symmetrize the polynomial $X^T = x_1^3 x_2^3 x_3^2$ by interchanging the positions in the SSYT.
We start out with two and three. The only row not containing any free two or three is the first row. Switching the number of free two's and three's we get the following SSYT

$$T_1 = \begin{ytableau} 1 & 1 & 1 & 3 \\ 2 & 2 & 3 \\ 3 \end{ytableau} .$$

Next, one and two are interchanged (note that it would not change anything if we did it in T and directly from T we cannot interchange one and three), the only free number is in row three, hence we get the tableau

$$T_2 = \begin{ytableau} 1 & 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 \end{ytableau}$$

Now we calculate

$$X^T + X^{T_1} + X^{T_2} = x_1^3x_2^3x_3^2 + x_1^3x_2^2x_3^3 + x_1^2x_2^3x_3^3$$

which clearly is a symmetric polynomial. [4]

Theorem 5.3.1 *For an ordered partition* α*,* $S_\alpha\{x_1, x_2 \ldots, x_n\} = \frac{A_\alpha(x_1,x_2,\ldots,x_n)}{V_n}$

This way of defining Schur polynomials makes it more suitable for the representation theory whilst the first definition is more closely connected to combinatorics [3]. Below is a rather involved proof using graph theory, but first we have to show a way of rewriting the right hand side of this theorem.

Proposition 5.1 For an ordered partition α, $\frac{A_\alpha(x_1,x_2,\ldots,x_n)}{V_n} = det(h_{\lambda_i} - i + j)$

To show this proposition we introduce 3 lemmas and a corollary.

Lemma 5.1 $E(t) = \sum\limits_{n=0}^{k} e_n(x_1, \ldots, x_k)t^n = \prod\limits_{i=1}^{k}(1 + x_it)$

Proof 5.2 In the expression to the left the power of t will equal the number of mixed distinct variables from e_n. Furthermore every mix of distinct variable will occur. Looking at the right hand side of the equation, the product will produce exactly one 1, exactly one x_it for all i exactly one $x_ix_jt^2$ for all $i \neq j$ and so on until we have exactly one $x_1 \ldots x_kt^k$. Hence both sides are equal.

Lemma 5.2 $H(t) = \sum\limits_{n\geq 0}^{k} h_n(x_1, \ldots, x_k)t^n = \prod\limits_{i=1}^{k} \frac{1}{(1-x_it)}$

Proof 5.3 It is well known that $\frac{1}{1-xt} = 1 + xt + (xt)^2 + (xt)^3 + \ldots$. Hence the product on the right hand side in the lemma will give us every mix of n of the variables $(x_1, x_2, \ldots, x_n)$ times t^n(variables are allowed to occur more than once), which is exactly what the left hand side is defined as.

Lemma 5.3 $H(t)E^{(i)}(-t) = \frac{1}{(1-x_it)}$

Proof 5.4 Fixing the variable i in E and using lemma 1 and lemma 2 we see that every term in the product $H(t)E^{(i)}(-t)$ cancel out except for $\frac{1}{(1-x_it)}$.

Corollary Let $\mu = (\mu_1, \ldots, \mu_l)$ be any composition. Consider the $l \times l$ matrices $A_\mu = (x_j^{\mu_i})$,$H_\mu = (h_{\mu_i - l + j})$ and $E = ((-1)^{l-i}e_{l-i}^{(j)}$. Then $A_\mu = H_\mu E$

Proof 5.5 Consider the generating function for the $e_n^{(j)}$, $E^{(j)}(t) = \sum\limits_{n=0}^{l-1} e_n^{(j)}t^n = \prod\limits_{i\neq j}(1+x_it)$. Since $H(t)E^{(j)}(-t) = \frac{1}{(1-x_it)}$ we can extract the coefficient of t^{μ_i} on both sides. This yields $\sum\limits_{k=1}^{l} h_{\mu_i - l + k}(-1)^{l-k}e_{l-k}^{(j)} = x_j^{\mu_i}$, which is what we wanted to prove.

Below is the proof of proposition 5.1.

Proof 5.6

Taking determinants in the corollary above, we obtain

$$|A_\mu| = |H\mu||E| \qquad (*),$$

where $|A_\mu| = a_\mu$. Let $\mu = \delta$, then $H_\delta = (h_{i-j})$, which is upper unitriangular. Thus $|H_\delta| = 1$ and we can rewrite $(*)$ as $|E| = a_\delta$. Now, letting $\mu = \lambda + \delta$ in the same equation, we have $\frac{a_{\lambda+\delta}}{a_\delta} = |H_{\lambda+\delta} = |h_{\lambda_i - i+j}|$. The left hand side is exactly $\frac{A_\alpha(x_1,x_2,\ldots,x_n)}{V_n}$ and we have proven proposition 5.1.

To show the equivalence of the two definitions of Schur polynomials we will introduce the Lindström-Gessel-Viennot lemma and some graph theory along with it. Firstly we need some definitions to make the LGV-lemma meaningful.

Definition Let G be a locally finite directed acyclic graph, that is each vertex has a finite degree and there are no directed cycles. To each directed edge e assign a weight ω_e. Now we can consider base vertices $A = \{a_1, a_2, \ldots, a_n\}$ and destination vertices $B = \{b_1, b_2, \ldots, b_n\}$. For each directed path P between two vertices, define $\omega(P)$ as the product of the weights of the edges of the path. Define $e(a, b) = \sum\limits_{P:a\to b} \omega(P)$ where the sum is over all paths P from a to b.

Furthermore we define a matrix $M = \begin{pmatrix} e(a_1,b_1) & e(a_1,b_2) & \ldots & e(a_1,b_n) \\ e(a_2,b_1) & e(a_2,b_2) & \ldots & e(a_2,b_n) \\ \vdots & \vdots & \ddots & \vdots \\ e(a_n,b_1) & e(a_n,b_2) & \ldots & e(a_n,b_n) \end{pmatrix}$

Definition An n-tuple of paths from A to B means an n-tuple $(P_1, P_2 ..., P_n)$ of paths in G with the property that there exists a permutation σ of $\{1, 2, ..., n\}$ such that, for every i, the path P_i is a path from a_i to $b_{\sigma(i)}$. The n-path is called non-intersecting if, whenever $i \neq j$, the paths P_i and P_j have no two vertices in common.

Definition Given an n-path $P = (P_1, P_2, ..., P_n)$, the weight $\omega(P)$ of this n-path is defined as the product $\omega(P_1)\omega(P_2)\ldots\omega(P_n)$.

Definition A **twisted** n-path P from an n-tuple $(a_1, a_2, ..., a_n)$ of vertices of G to an n-tuple $(b_1, b_2, ..., b_n)$ of vertices of G will mean an n-path from $(a_1, a_2, ..., a_n)$ to $(b_{\sigma(1)}, b_{\sigma(2)}, ..., b_{\sigma(n)})$ for some permutation σ in S_n. This permutation σ will be called the **twist** $\sigma(P)$ of this twisted n-path.

The Lindström-Gessel-Viennot lemma

$det(M) = \sum\limits_{(P_1,P_2,\ldots P_n):A\to B} sgn(\sigma(P))\omega(P)$, where the sum is taken over all non-intersecting n-paths from A to B.

Proof 5.7
We start of by rewriting the determinant of M.

$$det(M) = \sum_{\sigma \in S_n} sgn(\sigma)e(a_1, b_{\sigma(1)})e(a_2, b_{\sigma(2)}) \dots e(a_n, b_{\sigma(n)}) = \sum_{\sigma \in S_n} sgn(\sigma) \sum_P \omega(P),$$

where the sum ranges over all P such that P is an n-path from $(a_1, \dots, a_n)$ to $(b_{\sigma(1)}, \dots, b_{\sigma(n)})$. Using twisted n-paths we can rewrite the right hand side as follows $\sum_P sgn(\sigma(P))\omega(P)$. Where the sum ranges over all P such that P is a twisted n-path from $(a_1, \dots, a_n)$ to $(b_1, \dots, b_n)$. We want to prove that this is equal to the right hand side of the lemma, that is $\sum_P sgn(\sigma(P))\omega(P)$ where the sum now ranges over all P such that P is an non-intersecting twisted path from $(a_1, \dots, a_n)$ to $(b_1, \dots, b_n)$.

So, now we have to prove that

$$\sum_P sgn(\sigma(P))\omega(P) = 0 \qquad (*),$$

where the sum ranges over all P such that P is a not nonintersecting twisted path from $(a_1, \dots, a_n)$ to $(b_1, \dots, b_n)$. We create an involution on the set of all twisted not nonintersecting n-paths from $(a_1, \dots, a_n)$ to $(b_1, \dots, b_n)$ and that has the property that it flips the sign of $sgn(\sigma(P))$, while not changing $\omega(p)$. This is enough to prove $(*)$, since it implies that the sum on the right equals its negative. All we have to do now is construct the mentioned involution.
Let $P = (P_1, \dots, P_n)$ be a twisted, not nonintersecting n-path from $(a_1, \dots, a_n)$ to $(b_1, \dots, b_n)$. Let i be the smallest index such that P_i contains an intersection and let m be the first point along p_i that intersects another path (or more). Let j be the largest index such that P_j intersects P_i in the point m. Now we define out involution, call it f, as $f(P)$ and to be the same as P except for that the tails after m of P_i and P_j are switched. That is, the parts of these two paths starting at m and continuing to B are exchanged. Clearly f is an involution since i will again be the smallest index for our intersecting paths, m will be the first point of intersection and j will be the largest index containing m. Since $f(P)$ has the same multiset of edges as P we conclude that $\omega(f(P)) = \omega(P)$. Furthermore $f(P)$ is a twisted n-path and the twist $\sigma(f(P))$ differs from $\sigma(P)$ only by a transposition of $\sigma(i)$ and $\sigma(j)$. Therefore we know that $sgn(\sigma(f(P))) = -sgn(\sigma(P))$. Thus the lemma is proven.

Now we are finally prepared to prove theorem 5.3.1.

Proof 5.8 For the proof we use the LGV-lemma and an arbitrary partition $\lambda = (\lambda_1, \dots, \lambda_r)$ of n. Consider r starting points $a_i = (r + 1 - i, 1)$ and r end points $b_i = (\lambda_i + r + 1 - i, n)$ as points in a $\mathbb{Z}^2$ lattice. We define the lattice as a directed graph as follows: the only allowed directions are going one to the right or one up, the vertical edges has weight 1 and horizontal edges at height i has weight x_i. For a non-intersecting n-path $\sigma(P)$ has to be 1, since a_1 has to go to b_1 etcetera for the n-path to be non-intersecting. With this definition, r-tuples of non-intersecting paths from A to B are exactly SSYT of shape λ.

For instance, with the SSYT $\begin{array}{|c|c|c|}\hline 1 & 3 & 3 \\ \hline 2 & 4 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array}$, we get the corresponding 3-tuple

$\begin{array}{|c|}\hline 4 \\ \hline 3 \\ \hline 2 \\ \hline 1 \\ \hline\end{array}\ \begin{array}{|c|c|c|c|c|c|}\hline c & c & b & b & 0 & a \\ \hline c & 0 & b & a & a & a \\ \hline c & b & b & a & 0 & 0 \\ \hline c & b & a & a & 0 & 0 \\ \hline\end{array}$. Here the paths are denoted with a, b, c, the boxes are thought of as vertices and the lines separating the boxes are edges, the zeroes are unused vertices. In other words, c starts at $(1,1) = a_3$ and goes through $(1,1) \to (1,2) \to (1,3) \to (1,4) \to (2,4) = b_3$.
The weight of such an r-tuple is the corresponding summand in the first definition of the Schur polynomials.
On the other hand the determinant $det(M)$ from the LGV-lemma is the same the right hand side in proposition 5.1.

Below is an example of how to calculate a Schur polynomial from the two first definitions.

Example 5.4 Using the partition α=(1,1) and two variables. We get the determinant below which we shall divide with the Vandermonde polynomial.

$$\begin{vmatrix} x_1^{1+2-1} & x_2^{1+2-1} \\ x_1^{1+2-2} & x_2^{1+2-2} \end{vmatrix} = x_1^2 x_2 - x_1 x_2^2$$

Dividing this with $V_2 = x_1 - x_2$ we get our Schur polynomial.

$$S_\alpha(x_1, x_2) = \frac{x_1^2 x_2 - x_1 x_2^2}{x_1 - x_2} = x_1 x_2$$

We can check that we get the same polynomial using the first definition, clearly

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

is the only possible tableaux hence $(x_1, x_2)^T = x_1 x_2$.

Using the only other partition of 2, namely 2=2 we get the other Schur polynomial in two variables for the number 2.

$$\begin{vmatrix} x_1^{2+2-1} & x_2^{2+2-1} \\ x_1^{0+2-2} & x_2^{0+2-2} \end{vmatrix} = x_1^3 - x_2^3$$

Dividing this with $x_1 - x_2$ we get $S_2 = x_1^2 + x_1 x_2 + x_2^2$.
Now we instead have three possible tableaux

$$T_1 = \begin{array}{|c|c|}\hline 1 & 1 \\ \hline\end{array}\ \ T_2 = \begin{array}{|c|c|}\hline 2 & 2 \\ \hline\end{array}\ \ T_3 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array}.$$

They give us $S_{(2,0)} =$

$$= (x_1, x_2)^{(T_1,T_2,T_3)} = (x_1, x_2)^{T_1} + (x_1, x_2)^{T_2} + (x_1, x_2)^{T_3} = x_1^2 + x_2^2 + x_1x_2$$

Below $S_{(1,1)}$ and $S_{(2,0)}$ are rewritten in terms of the elementary symmetric polynomials.

$$S_{(1,1)} = x_1x_2 = e_2$$
$$S_{(2,0)} = x_1^2 + x_1x_2 + x_2^2 = e_1^2 - e_2$$

5.4 Some other identities

Theorem 5.3.3 Youngs Rule $S_\lambda = \sum\limits_\mu K_{\lambda\mu} m_\mu$

The proof can be found on page 355 in [3].

Hence the Kostka numbers can be defined as the constants in the Schur polynomial. Noteworthy is that there are no nice formulas knows for the Kostka numbers. That makes this definition rather hard to work with while the two first ones are simpler.

We add a final identity using the Schur polynomials.

Theorem: Cauchy Identity and dual. For two sets of variables $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$ we have

$$\sum_\lambda s_\lambda(X)s_\lambda(Y) = \prod_{i,j=1}^n \frac{1}{1 - x_iy_j}$$

and

$$\sum_\lambda s_\lambda(X)s_{\lambda'}(Y) = \prod_{i,j=1}^n \frac{1}{1 + x_iy_j}$$

It is interesting that Schur polynomials are rather abstract and not so easy to work with, but here they are written as simple products.

6 A connection to representation theory

We want to see the connection between symmetric polynomials, especially Schur polynomial, and the representation theory. We take the general linear group as example.

6.1 A short intruduction to representations

Definition 6.1.1 The **General Linear Group** of degree n, here denoted GL_n, is the set of $n \times n$ invertible matrices with elements in some field, e.g. $\mathbb{Q}$.

Recall that every invertible square matrix has a determinant different from zero, hence its rows are linearly independent as well as its columns.

Definition 6.1.2 A representation $\rho : GL_n \to GL_m$ is **polynomial** if for every $X \in GL_n$ the entries of $\rho(X)$ are polynomials in the entries of X.
See below for an example.

Definition 6.1.3 A representation $\rho : G \to GL(V)$ is said to be an **irreducible representation** if it has only trivial subrepresentations.
(A linear subspace $W \subset V$ is called G-invariant if $gw \in W$ for all $g \in G$ and all $w \in W$. The restriction of ρ to a G-invariant subspace $W \subset V$ is known as a **subrepresentation**.)

Definition 6.1.3 Let V be a finite-dimensional vector space over a field F and let $\rho : G \to GL(V)$ be a representation of a group G on V. The **character** of ρ is the function $\xi_\rho : G \to F$ given by $\xi_\rho(g) = trace(\rho(g))$. The character ξ_ρ is called **irreducible** if ρ is an irreducible representation.

We will not prove the following theorem but include it as an explanation why Schur polynomials are important.

Theorem 6.1.1 The irreducible polynomial characters of GL_n are the s_λ where λ is an ordered partition with n entries.

6.2 Example of a representation

Example 6.1
Consider $V = \mathbb{Q}e_1 \oplus \mathbb{Q}e_2$ with basis $< e_1, e_2 >$, and the action of $GL_2(\mathbb{Q})$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} e_1 = ae_1 + ce_2$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} e_2 = be_1 + de_2.$$

This induces action on polynomials in e_1, e_2, by $g(e_1^m e_2^n) = g(e_1)^m g(e_2)^n$.

Let $< e_1^m, e_1^{m-1}e_2, \ldots, e_2^m >$ be the basis for homogeneous polynomials of degree m in e_1, e_2.

As example we take $m = 2$ and therefore have the basis $< e_1^2, e_1e_2, e_2^2 >$. For the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we get that $g(e_1^2) = (ae_1 + ce_2)^2$. This maps to a new matrix in the following way:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a^2 & ? & ? \\ 2ac & ? & ? \\ c^2 & ? & ? \end{pmatrix}$$

We can calculate the question marks in the same way,

$$g(e_1e_2) = g(e_1)g(e_2) = (ae_1 + ce_2)(be_1 + de_2) = abe_1^2 + (ad + bc)e_1e_2 + cde_2^2,$$

the last calculation is analogous to the first so we just fill in the matrix

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

By looking at every diagonilized matrix

$$GL_2 \supset \begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1^2 & ? & ? \\ ? & x_1x_2 & ? \\ ? & ? & x_2^2 \end{pmatrix}$$

where the question marks are not interesting because we want to take the trace of the 3×3 matrix (they are 0). The trace is $x_1^2 + x_1x_2 + x_2^2$. From example 5.4 we know that this is a Schur polynomial, namely $s_{(1,1)}(x_1, x_2)$.
Comparing this result to theorem 5.4.1 we see that $s_{(1,1)}(x_1, x_2)$ is our irreducible polynomial character.

References

[1] Beachy JA, Blair WD. Abstract algebra. 3rd edition. Long grove: Waveland press, inc.

[2] Macdonald IG. Symmetric functions and hall polynomials. 2nd edition. Oxford: Clarendon Press, 1995.

[3] Stanley RP. Enumerative Combinatorics. 2nd volume. Cambridge: Cambridge University Press, 2012.

[4] Sagan BE. The symmetric group representations, combinatorical algorithms, and symmmetric functions. 2nd edition. New York: Springer, 2001.

[5] Bandlow J. An elementary proof of the hook formula [Dissertation]. Davis (CA) University of California-Davis; 2008

[6] Sagan BE. Schur functions. East Lansing: Michigan State University. Available from: http://www.mth.msu.edu/ sagan/Papers/Old/schur.pdf