Cap. 1 Symmetric group.

July 2003

When we introduce a new symbol or definition we will use the convenient form $:=$ which means that the term introduced at its left is defined by the expression at its right. A typical example could be $P := \{x \in \mathbb{N} | 2 \text{ divides } x\}$ which stands for: *P is by definition the set of all natural numbers x such that 2 divides x.*

The symbol $\pi : A \to B$ denotes a mapping named π from the set A to the set B.

Most of our work will be for algebras over the field of real or complex numbers, sometimes we will take a more combinatorial point of view and analyze some properties over the integers. Associative algebras will implicitely be assumed to have a unit elment. When we discuss matrices over a ring A we always identify A with the scalar matrices (constant multiples of the identity matrix).

We use the standard notations:

$$\mathbb{N},\ \mathbb{Z},\ \mathbb{Q},\ \mathbb{R},\ \mathbb{C}$$

for the natural numbers (including 0), the integers, rational, real and complex numbers.

Typeset by $\mathcal{AMS}$-TEX

Symmetric group

July 2003

Contents

Introduction In this chapter we will develop the basic facts of the representation theory of the symmetric group.

1 Symmetric functions

1.1 The theory of symmetric functions is a classical theory developed (by Lagrange, Ruffini, Galois and others) in connection with the theory of algebraic equations in one variable and the classical question of resolution by radicals.

The main link are the formulas expressing the coefficients of a polynomial through its roots. A formal approach is the following.

Consider polynomials in variables $x_1, x_2, \ldots, x_n$ and an extra variable t over the ring of integers. The elementary symmetric functions $e_i := e_i(x_1, x_2, \ldots, x_n)$ are implicitely defined by the formula:

(1.1.1)
$$p(t) := \prod_{i=1}^{n}(1 + tx_i) := 1 + \sum_{i=1}^{n} e_i t^i.$$

More explicitely $e_i(x_1, x_2, \ldots, x_n)$ is the sum of $\binom{n}{i}$ terms, the products, over all subsets of $\{1, 2, \ldots, n\}$ with i elements, of the variables with indeces in that subset.

(1.1.2)
$$e_i = \sum_{1 \le a_1 < a_2 < \cdots < a_i \le n} x_{a_1} x_{a_2} \ldots x_{a_i}.$$

If σ is a permutation of the indeces we obviously have

$$\prod_{i=1}^{n}(1+tx_i)=\prod_{i=1}^{n}(1+tx_{\sigma i})$$

and thus the elements e_i are invariant under permutation of the variables.

Of course the polynomial $t^n p(-\frac{1}{t})$ has the elements x_i as its roots.

Definition. *A polynomial in the variables $(x_1, x_2, \ldots, x_n)$, invariant under permutation of these variables, is called a symmetric function.*

The functions e_i are called **elementary symmetric functions.**

There are several obviously symmetric functions. The power sums $\psi_k := \sum_{i=1}^n x_i^k$ and the funtions S_k defined as the sum of all monomials of degree k.

These are particular cases of the following general construction.

Consider the basis of the ring of polynomials given by the monomials which is permuted by the symmetric group.

By Proposition 2.4 we have:

A basis of the space of symmetric functions is given by the sums of monomials in the same orbit, for all orbits.

Orbits correspond to non increasing vectors $(h_1 \geq h_2 \geq \ldots \geq h_n),\ h_i \in \mathbb{N}$ and we may set $\Sigma_{(h_1,h_2,\ldots,h_n)}$ to be the sum of monomials in the corresponding orbit.

As we will see soon there are also some subtler symmetric functions (the Schur functions) that will play an important role in the sequel.

We can start with a first important fact, the explicit connection between the functions e_i and the ψ_k.

To do this we will perform the next computations in the ring of formal power series, although the series that we will consider have also a meanining as convergent series.

Start from the identity $\prod_{i=1}^n(tx_i+1)=\sum_{i=0}^n e_i t^i$ and take the logaritmic derivative (relative to the variable t) of both sides. We use the fact that such an operator transforms products into sums to get

$$\sum_{i=1}^{n}\frac{x_i}{(tx_i+1)}=\frac{\sum_{i=1}^n ie_it^{i-1}}{\sum_{i=0}^n e_it^i}.$$

The left hand side of this formula can be developed as

$$\sum_{i=1}^{n}x_i\sum_{h=0}^{\infty}(-tx_i)^h=\sum_{h=0}^{\infty}(-t)^h\psi_{h+1}.$$

From this we get the identity

$$(\sum_{h=0}^{\infty}(-t)^h\psi_{h+1})(\sum_{i=0}^{n}e_it^i)=(\sum_{i=1}^{n}ie_it^{i-1})$$

which gives, equating coefficients:

$$(-1)^m\psi_{m+1}+\sum_{i=1}^{m}(-1)^i\psi_i e_{m+1-i}=\sum_{i+j=m}(-1)^i\psi_{i+1}e_j=(m+1)e_{m+1}$$

where we intend $e_i=0$ if $i>n$.

It is clear that these formulas give recursive ways of expressing the ψ_i in terms of the e_j with integral coefficients, on the other hand they can also be used to express the e_i in terms of the ψ_j, but in this case it is necessary to perform some divisions and the coefficients are rational and usually not integers.[1]

It is useful to give a second proof, consider the map:

$$\pi_n:\mathbb{Z}[x_1,x_2,\dots,x_n]\to\mathbb{Z}[x_1,x_2,\dots,x_{n-1}]$$

given by evaluating x_n in 0.

Lemma. *The intersection of $Ker(\pi_n)$ with the space of symmetric functions of degree $<n$ is reduced to 0.*

Proof. Consider $\Sigma_{(h_1,h_2,\dots,h_n)}$, a sum of monomials in an orbit, if the degree is less than n we have $h_n=0$; under π_n we get $\pi_n(\Sigma_{(h_1,h_2,\dots,h_n)})=\Sigma_{(h_1,h_2,\dots,h_{n-1})}$ thus if the degree is less than n the map π_n maps these basis elements into distinct basis elements.

Now the second proof. In the identity $\prod_{i=1}^n(t-x_i):=\sum_{i=0}^n(-1)^ie_it^{n-i}$ substitute t with x_i and then sum over all i we get:

$$0=\sum_{i=0}^{n}(-1)^ie_i\psi_{n-i},\ \text{or}\ \psi_n=\sum_{i=1}^{n}(-1)^{i-1}e_i\psi_{n-i}.$$

By the previous lemma this identity remains valid also for symmetric functions in more than n variables and gives the required recursion.

It is in fact a general fact that symmetric functions can be expressed as polynomials in the elementary ones, we will now discuss an algorithmic proof.

To make the proof transparent let us stress in our formulas also the number of variables and denote by $e_i^{(k)}$ the i^{th} elementary symmetric function in the variables $x_1,\dots,x_k$. Since:

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

we have:

$e_i^{(n)}=e_{i-1}^{(n-1)}x_n+e_i^{(n-1)}$ or $e_i^{(n-1)}=e_i^{(n)}-e_{i-1}^{(n-1)}x_n$.

In particular, in the homomorphism $\pi:\mathbb{Z}[x_1,\dots,x_n]\to\mathbb{Z}[x_1,\dots,x_{n-1}]$ given by evaluating x_n in 0 we have that symmetric functions map to symmetric functions and

$$\pi(e_i^{(n)})=e_i^{(n-1)},\ i<n,\ \pi(e_n^{(n)})=0.$$

[1] These formulas were found by Newton, hence the name *Newton functions* for the ϕ_k.

Given a symmetric polynomial $f(x_1, \ldots, x_n)$ we evaluate it at $x_n = 0$, if the resulting polynomial $\overline{f}(x_1, \ldots, x_{n-1})$ is 0 then f is divisible by x_n.

If so, by symmetry it is divisible by all of the variables and hence by the function e_n. We perform the division and pass to another symmetric function of lower degree.

Otherwise by induction there exists a polynomial p in $n-1$ variables which, evaluated in the $n-1$ elementary symmetric functions of $x_1, \ldots, x_{n-1}$, gives $f(x_1, \ldots, x_{n-1}, 0)$. Thus $f - p(e_1, e_2, \ldots, e_{n-1})$ is a symmetric function vanishing at $x_n = 0$.

We are back to the previous step.

The uniqueness is implicit in the algorithm which can be used to express any symmetric polynomial as a unique polynomial in the elementary symmetric functions.

Theorem. *A symmetric polynomial is a polynomial, in a unique way, in the elementary symmetric functions.*

It is quite useful, in view of the previous lemma and theorem, to think in a stable way to symmetric functions in larger and larger sets of variables, one the constructs a *limit ring*, which one calls just the **ring of symmetric functions** $\mathbb{Z}[e_1, \ldots, e_i, \ldots]$ and can be thought as the polynomial ring in infinitely many variables e_i where formally we give degree (or weight) i to e_i. The ring of symmetric funcions in $n-$variables is obtained by setting $e_i = 0$, $\forall i > n$. One often develops formal identities in this ring with the idea that, in order to verify an identity whic is homogeneous of some degree m it is enough to do it for symmetric functions in $m-$variables.

1.2 In the same way the reader may discuss the following fact.

Consider the $n!$ monomials

$$x_1^{h_1} \ldots x_{n-1}^{h_{n-1}}, \ 0 \le h_i \le n - i.$$

Theorem. *The previous monomials are a basis of* $\mathbb{Z}[x_1, \ldots, x_n]$ *over* $\mathbb{Z}[e_1, \ldots, e_n]$.

Remark. The same theorem is clearly true if we replace the coefficient ring $\mathbb{Z}$ by any commutative ring A. In particular we will use it when A is itself a polynomial ring.

2 Resultant, discriminant, Bezoutiante

2.1 In order to understand the importance of theorem 1.1 on elementary symmetric functions and also the classical point of view let us develop a geometric picture.

Consider the space $\mathbb{C}^n$ and the space $P_n := \{t^n + b_1 t^{n-1} + \ldots + b_n\}$ of monic polynomials (which can be identified to $\mathbb{C}^n$ by the use of the coefficients).

Consider next the map $\pi : \mathbb{C}^n \to P_n$ given by:

$$\pi(\alpha_1, \ldots, \alpha_n) := \prod_{i=1}^{n} (t - \alpha_i).$$

We thus obtain a polynomial $t^n - a_1 t^{n-1} + a_2 t^{n-2} + \cdots + (-1)^n a_n = 0$ with roots $\alpha_1, \ldots, \alpha_n$ (and the coefficients a_i are the elementary symmetric functions in the roots), any monic polynomial is obtained in this way (fundamental theorem of Algebra).

Two points in $\mathbb{C}^n$ project to the same point in P_n if and only if they are in the same orbit under the symmetric group, i.e. P_n parametrizes the S_n orbits.

Suppose we want to study a property of the roots which can be verified by evaluating some symmetric polynomials in the roots, this will usually be the case for any condition on the set of all roots. Then one can perform the computation without expliciting the roots, since one has only to study the formal symmetric polynomial expression and, using the previous or another algorithm express the value of a symmetric function of the roots through the coefficients.

In other words a polynomial function f on $\mathbb{C}^n$ which is symmetric, factors through the map π giving rise to an effectively computable[2] polynomial function $\overline{f}$ on P_n such that $f = \overline{f}\pi$.

A classical example is given by the discriminant.

The condition that the roots be distinct is clearly that $\prod_{i<j}(\alpha_i - \alpha_j) \neq 0$. The polynomial $V(x) := \prod_{i<j}(x_i - x_j)$ is in fact not symmetric. It is the value of the Vandermonde determinant, i.e. the determinant of the matrix:

$$A := \begin{pmatrix} x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ x_1 & x_2 & \ldots & x_n \\ 1 & 1 & \ldots & 1 \end{pmatrix} \tag{2.1.1}$$

Proposition. *$V(x)$ is antisymmetric, i.e. permuting the variables it is multiplied by the sign of the permutation.*

Remark. The theory of the sign of permutations can be deduced by analyzing the Vandermonde. In fact since for a transposition τ it is clear that $V(x)^\tau = -V(x)$ it follows that $V(x)^\sigma = V(x)$, or $-V(x)$ according to whether σ is a product of an even or an odd number of permutations. The sign is then clearly a homomorphism.

We also see immediately that V^2 is a symmetric polynomial.

We can compute it in terms of the functions ψ_i as follows. Consider the matrix $B := AA^t$, clearly in the ij position of B we find the symmetric function $\psi_{2n-(i+j)}$ and its determinant is V^2.

[2] i.e. computable without solving the equation

The matrix B (or rather the one reordered with ψ_{i+j-2} in the ij position) is classically known as the Bezoutiante and it carries some further information on the roots. We shall see for it a different determinant formula involving directly the elementary symmetric functions.

We write V^2 as a polynomial $D(e_1, e_2, \dots, e_n)$ in the elementary symmetric functions.

Definition. *The polynomial D is called the discriminant.*

Since this is an interesting example we will pursue it a bit further.

Let us assume that F is a field, $f(t)$ a monic polynomial (of degree n) with coefficients in F and let $R := F[t]/(f(t))$. R is an algebra over F of dimension n.

For any finite dimensional algebra A over a field F we can perform the following construction.

Any element a of A induces a linear transformation $L_a : x \to ax$ on A (and also a right one). We define $tr(a) := tr(L_a)$, the trace of the operator L_a.

We consider next the bilinear form $(a, b) := tr(ab)$ this is the *trace form* of A. It is symmetric and *associative* in the sense that $(ab, c) = (a, bc)$.

We compute it first for $R := F[t]/(t^n)$ using the fact that t is nilpotent we see that $tr(t^k) = 0$ if $k > 0$ and so the trace form has rank 1 with kernel the ideal generated by t.

To compute it for the algebra $R := F[t]/(f(t))$ we pass to the algebraic closure $\overline{F}$ and compute in $\overline{F}[t]/(f(t))$.

We split the polynomial with respect to its distinct roots $f(t) = \prod_{i=1}^k (t - \alpha_i)^{h_i}$ and $\overline{F}[t]/(f(t)) = \oplus_{i=1}^k \overline{F}[t]/(t - \alpha_i)^{h_i}$.

Thus the trace of an element mod $f(t)$ is the sum of its traces mod $(t - \alpha_i)^{h_i}$.

Let us compute the trace of t^k mod $(t - \alpha_i)^{h_i}$ we claim that it is $h_i \alpha_i^k$. In fact in the basis $1, (t - \alpha_i), (t - \alpha_i)^2, \dots, (t - \alpha_i)^{h_i - 1}$ (mod $(t - \alpha_i)^{h_i}$) the matrix of t is upper triangular with constant eigenvalue α_i on the diagonal and so the claim follows.

Adding all the contributions we see that, in $F[t]/(f(t))$ the trace of multiplication by t^k is $\sum_i h_i \alpha_i^k$.

As a consequence we see that the matrix of the trace form, in the basis $1, t, \dots, t^{n-1}$ is the Bezoutiante of the roots. Since for a given block $\overline{F}[t]/(t - \alpha_i)^{h_i}$ the ideal generated by $(t - \alpha_i)$ is nilpotent of codimension 1, we see that it is exactly the radical of the block and the kernel of its trace form. It follows that:

Proposition. *The rank of the Bezoutiante equals the number of distinct roots.*

Given a polynomial $f(t)$ let $\overline{f}(t)$ denote the polynomial with the same roots as $f(t)$ but all distinct. It is the generator of the radical of the ideal generated by $f(t)$. In characteristic zero this polynomial is obtained dividing $f(t)$ by the G.C.D. between $f(t)$ and its derivative.

Let us consider now the algebra $R := F[t]/(f(t))$ its radical N and $\overline{R} := R/N$. By the previous analysis it is clear that $\overline{R} = F[t]/(\overline{f}(t))$.

Consider now the special case in which $F = \mathbb{R}$ is the field of real numbers. Then we can divide the distinct roots into the real roots $\alpha_1, \alpha_2, \ldots, \alpha_k$ and the complex ones $\beta_1, \overline{\beta}_1, \beta_2, \overline{\beta}_2, \ldots, \beta_h, \overline{\beta}_h$.

The algebra $\overline{R}$ is isomorphic to the direct sum of k copies of $\mathbb{R}$ and h copies of $\mathbb{C}$, its trace form is the orthogonal sum of the corresponding trace forms. On $\mathbb{R}$ the trace form is just x^2 but on $\mathbb{C}$ we have $tr((x+iy)^2) = 2(x^2 - y^2)$. We deduce that:

Theorem. *The number of real roots of $f(t)$ equals the signature of its Bezoutiante.*

There are simple variations on this theme, for instance if we consider the quadratic form $Q(x) := tr(tx^2)$ we see that its matrix is again easily computed in terms of the ψ_k and its signature equals the number of real positive minus the number of real negative roots. In this way one can also determine the number of real roots in any interval.

These results are Sylvester's variations on Sturm's theorem. They can be found in the paper in which he discusses the law of Inertia which now bears his name (cf. [Si]).

2.2 Let us go back to the roots, if $x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m$ are two sets of variables consider the polynomial

$$A(x, y) := \prod_{i=1}^{n} \prod_{j=1}^{m} (x_i - y_j).$$

This is clearly symmetric, separately in the variables x and y, if we evaluate it in numbers it vanishes if and only if one of the values of the x's coincides with a value of the y's, conversely any polynomial in these two sets of variables which has this property is a multiple of A.

By the general theory A can be expressed as a polynomial R in elementary symmetric functions.

Let us denote by $a_1, a_2, \ldots, a_n$ the elementary symmetric functions in the x_i's and $b_1, \ldots, b_m$ the ones in the y_j's. Thus $A(x, y) = R(a_1, \ldots, a_n, b_1, \ldots, b_m)$.

The polynomial R is called the *resultant*.

When we evaluate the variables x and y to be the roots of two monic polynomials $f(t)$, $g(t)$ of degrees n, m respectively we see that the value of A can be computed evaluating R in the coefficients (with some signs) of these polynomials. Thus the resultant is the polynomial in their coefficients, vanishing when the two polynomials have a common root.

There is a more general classical expression as determinant, and we drop the condition that the polynomials be monic. The theory is the following.

Let $f(t) := a_0t^n + a_1t^{n-1} + \cdots + a_n$ $g(t) := b_0t^n + b_1t^{n-1} + \cdots + b_m$ and let us denote by P_h the $h+1$ dimensional space of all polynomials of degree $\leq h$.

Consider the linear transformation:

$$T_{f,g} : P_{m-1} \oplus P_{n-1} \to P_{m+n-1} \text{ given by } T_{f,g}(a, b) := fa + gb.$$

This is a transformation between two $n+m$ dimensional spaces and in the bases $(1,0), (t,0), \ldots, (t^{m-1}, 0), (0,1), (0,t), \ldots, (0, t^{n-1})$ and $1, t, t^2, \ldots, t^{n+m-1}$ it is quite easy to write down its square matrix $R_{f,g}$:

$$\left(\begin{array}{cccccccccc}
a_n & 0 & 0 & \ldots & 0 & b_m & 0 & \ldots & 0 & 0 \\
a_{n-1} & a_n & 0 & \ldots & 0 & b_{m-1} & b_m & \ldots & \ddots & 0 \\
a_{n-2} & a_{n-1} & a_n & 0 & \ldots & b_{m-2} & b_{m-1} & b_m \ddots & & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
a_1 & a_2 & a_3 & & & & & & \ddots & \vdots \\
a_0 & a_1 & a_2 & & & & & & & \\
0 & a_0 & a_1 & & & & & & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \vdots & & & & b_0 & b_1 & b_2 & \ddots & \vdots \\
0 & 0 & \ldots & & & 0 & b_0 & b_1 & \ddots & \vdots \\
0 & 0 & & & & 0 & 0 & b_0 & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & b_0 & \vdots \\
0 & 0 & 0 & \ldots & a_0 & 0 & \ldots & \ldots & 0 & b_0
\end{array}\right) \tag{2.2.1}$$

Proposition. *If $a_0b_0 \neq 0$, the rank of $T_{f,g}$ equals $m+n-d$ where d is the degree of $h := G.C.D(f,g)$*[3] .

Proof. By Euclid's algorithm the image of $T_{f,g}$ consists of all polynomials of degree $\leq n+m-1$ and multiples of h, its kernel of pairs $(sg', -sf')$ where $f = hf', g = hg'$, hence the claim.

As a corollary we have that the determinant $R(f,g)$ of $R_{f,g}$ vanishes exactly when the two polynomials have a common root. This gives us a second definition of resultant.

Definition. *The polynomial $R(f,g)$ is called the resultant of the two polynomials $f(t), g(t)$.*

If we consider the coefficients of f and g as variables we can still think of $T_{f,g}$ as a map of vector spaces, except that the base field is the field of rational functions in the given variables.

[3] $G.C.D(f,g)$ is the greatest common divisor of f, g.

Then we can solve the equation $fa + gb = 1$ by Cramer's rule and we see that the coefficients of the polynomials a, b are given by the cofactors of the first row of the matrix $R_{f,g}$ divided by the resultant, in particular we can write $R = Af(t) + Bg(t)$ where A, B are polynomials in t of degrees $m-1, n-1$ respectively and with coefficients polynomials in the variables $(a_0, a_1, \ldots, a_n, b_0, b_1, \ldots, b_m)$.

This can also be understood as follows. In the matrix $R_{f,g}$ we add to the first row the second multiplied by t the third multiplied by t^2 and so on, we see that the first row becomes $f(t), f(t)t, f(t)t^2, \ldots, f(t)t^{m-1}, g(t), g(t)t, g(t)t^2, \ldots, g(t)t^{n-1}$. Under these operations of course the determinant does not change and we see that developing it along the first row we get the desired identity.

Exercise Consider the two polynomials as $a_0 \prod_{i=1}^{n}(t - x_i)$, $b_0 \prod_{j=1}^{m}(t - y_j)$ and thus substitute in R to the variables a_i the element $(-1)^i a_0 e_i(x_1, \ldots, x_n)$ and to b_i the element $(-1)^i b_0 e_i(y_1, \ldots, y_m)$. The polynomial we obtain is $a_0^m b_0^n A(x, y)$.

2.3 In the special case when we take $g(t) = f'(t)$, the derivative of $f(t)$, we have that the vanishing of the resultant is equivalent to the existence of multiple roots. We have already seen that the vanishing of the discriminant implies the existence of multiple roots, it is now easy to connect the two approaches.

The resultant $R(f, f')$ is considered as a polynomial in the variables $(a_0, a_1, \ldots, a_n)$, if we substitute in $R(f, f')$ to the variables a_i the element $(-1)^i a_0 e_i(x_1, \ldots, x_n)$ we have a polynomial in the x with coefficients involving a_0 which vanishes whenever two x's coincides.

Thus it is divisible by the discriminant of these variables. A degree computation shows in fact that it is a constant (with respect to the x) multiple cD. The constant c can be evaluated easily for instance specializing to the polynomial $x^n - 1$, this polynomial has as roots the n^{th} roots $e^{\frac{2\pi i k}{n}}$, $0 \leq k < n$ of 1. The Newton functions

$$\psi_h := \sum_{i=0}^{n-1} e^{\frac{2\pi i h k}{n}} = \begin{cases} 0 & \text{if} \quad h \not| n \\ n & \text{if} \quad h \,|\, n \end{cases}$$

hence the Bezoutiante is $-(-n)^n$ and the computation of the Resultant is n^n so the constant is $(-1)^{n-1}$.

3 Schur functions

3.1 It is important to discuss along symmetric, also alternating functions, we assume to work on integral polynomials.

Definition. *A polynomial f in the variables $(x_1, x_2, \ldots, x_n)$, is called an alternating func-*

tion if, for every permutation σ of these variables

$$f^\sigma = f(x_{\sigma(1)}, x_{\sigma(2}, \dots, x_{\sigma(n)}) = \epsilon_\sigma f(x_1, x_2, \dots, x_n),$$

ϵ_σ *being the sign of the permutation.*

We have seen the Vandermonde determinant as a basic alternating polynomial, $V(x) := \prod_{i<j}(x_i - x_j)$.

The main remark on alternating functions is the following.

Proposition. *A polynomial $f(x)$ is alternating if and only if it is of the form $f(x) = V(x)g(x)$ with $g(x)$ a symmetric polynomial.*

Proof. Substitute, in an alternating polynomial f to a variable x_j a variable x_i for $i \neq j$. We get the same polynomial if we first exchange x_i and x_j in f. Since this changes the sign it means that, under this substitution f becomes 0.

This means in turn, that f is divisible by $x_i - x_j$; since i, j are arbitrary f is divisible by $V(x)$. Writing $f = V(x)g$ it is clear that g is symmetric.

Let us be more formal, let A, S denote the sets of **antisymmetric** and **symmetric** polynomials. We have seen that:

Proposition. *The space A of antisymmetric polynomials is a free rank 1 module over the ring S of symmetric polynomials generated by $V(x)$ or $A = V(x)S$.*

In particular any integral basis of A gives, dividing by $V(x)$, an integral basis of S. In this way we will presently obtain the *Schur functions.*

To understand the construction let us make a fairly general discussion. In the ring of polynomials $Z[x_1, x_2, \dots, x_n]$ let us consider the basis given by the monomials (which are permuted by S_n).

Recall that the orbits of monomials are indexed by non increasing sequences of integers. To $m_1 \geq m_2 \geq m_3 \cdots \geq m_n \geq 0$ corresponds the orbit of the monomial $x_1^{m_1} x_2^{m_2} x_3^{m_3} \dots x_n^{m_n}$.

Let f be an antisymmetric polynomial and (ij) a transposition. Applying this transposition to f it changes sign while the transposition fixes all monomials in which x_i, x_j have the same exponent.

It follows that all the monomials which have non 0 coefficient in f must have distinct exponents. Given a sequence of exponents $m_1 > m_2 > m_3 > \cdots > m_n \geq 0$ the coefficients of the monomial $x_1^{m_1} x_2^{m_2} x_3^{m_3} \dots x_n^{m_n}$ and of $x_{\sigma(1)}^{m_1} x_{\sigma(2)}^{m_2} x_{\sigma(3)}^{m_3} \dots x_{\sigma(n)}^{m_n}$ differ by the sign of σ.

It follows that:

Theorem. *The functions:*

$$A_{m_1 > m_2 > m_3 > \cdots > m_n \geq 0}(x) := \sum_{\sigma \in S_n} \epsilon_\sigma x_{\sigma(1)}^{m_1} x_{\sigma(2)}^{m_2} \dots x_{\sigma(n)}^{m_n},$$

are an integral basis of the space of antisymmetric functions.

It is often useful, when computing with alternating functions, to use a simple device. Consider the subspace SM spanned by the set of *standard monomials* $x_1^{k_1}x_2^{k_2}\dots x_n^{k_n}$ with $k_1 > k_2 > k_3 \dots > k_n$ and the linear map L from the space of polynomials to SM which is 0 on the non standard monomials and it is the identity on SM. Then $L(\sum_{\sigma\in S_n}\epsilon_\sigma x_{\sigma(1)}^{m_1}x_{\sigma(2)}^{m_2}\dots x_{\sigma(n)}) = x_1^{m_1}x_2^{m_2}\dots x_n^{m_n}$ thus L establishes a linear isomorphism between the space of alternating polymonials and SM which maps the basis of the theorem in the standard monomials.

3.2 It is convenient to use the following conventions. Consider the sequence

$$\varrho := (n-1, n-2, \dots, 2, 1, 0),$$

Lemma. *The map*

$$\lambda = (p_1, p_2, p_3, \dots, p_n) \longrightarrow \lambda + \varrho = (p_1 + n - 1, p_2 + n - 2, p_3 + n - 3, \dots, p_n)$$

is a a bijiective correspondence between decreasing and strictly decreasing sequences.

We thus indicate by $A_{\lambda+\varrho}$ the corresponding antisymmetric function. We can express it also as a determinant of the matrix M_λ having in the i, j position the element $x_j^{p_i+n-i}$.[4]

We next set $S_\lambda(x) := A_{\lambda+\varrho}/V(x)$ the *Schur function* associated to λ, when there is no ambiguity we will drop the variables symbol and speak of S_λ.

We can identify λ to a *partition* with $n-$parts, of the integer $\sum p_i$ and write $\lambda \vdash \sum_i p_i$.

We have thus that (with the notations of 1.1):

Theorem. *The functions S_λ, with $\lambda \vdash m$ and $ht(\lambda) \le n$ are an integral basis of the part of degree m of the ring of symmetric functions.*

Notice that the Vandermonde determinant is the alternating function A_ϱ and $S_0 = 1$.

Several interesting combinatorial facts are associated to these functions, we will see some of them in the next section. The main significance of the Schur functions is in the representation theory of the linear group as we will see later in Chapter 3.

If $\lambda = (p_1, p_2, p_3, \dots, p_n)$ is a partition and a a positive integer let us denote by $\underline{a}$ the partition $(a, a, a, \dots, a)$ then from 6.1.1 follows that

$$A_{\lambda+\varrho+\underline{a}} = (x_1x_2\dots x_n)^a A_{\lambda+\varrho},\ S_{\lambda+\underline{a}} = (x_1x_2\dots x_n)^a S_\lambda. \tag{3.2.1}$$

We let n be the number of variables and want to understand, given a Schur function $S_\lambda(x_1, \dots, x_n)$ the form of $S_\lambda(x_1, \dots, x_{n-1}, 0)$ as symmetric function in $n-1$ variables.

Let $\lambda := h_1 \ge h_2 \ge \dots \ge h_n \ge 0$, we have seen that, if $h_n > 0$ then $S_\lambda(x_1, \dots, x_n) = \prod_{i=1}^n x_i S_{\overline{\lambda}}(x_1, \dots, x_n)$ where $\overline{\lambda} := h_1 - 1 \ge h_2 - 1 \ge \dots \ge h_n - 1$.

[4] It is conventional to drop the numbers equal to 0 in a decreasing sequence.

In this case, clearly $S_\lambda(x_1,\dots,x_{n-1},0)=0$.

Assume now $h_n=0$ and denote by the same symbol λ the sequence $h_1\geq h_2\geq\dots\geq h_{n-1}$. Let us start from the Vandermonde determinant $V(x_1,\dots,x_{n-1},x_n)=\prod_{i<j\leq n}(x_i-x_j)$ and set $x_n=0$ getting

$$V(x_1,\dots,x_{n-1},0)=\prod_{i=1}^{n-1}x_i\prod_{i<j\leq n-1}(x_i-x_j)=\prod_{i=1}^{n-1}x_iV(x_1,\dots,x_{n-1}).$$

Now consider the alternating function $A_{\lambda+\varrho}(x_1,\dots,x_{n-1},x_n)$.

Set $\ell_i:=h_i+n-i$ so that $\ell_n=0$ and

$$A_{\lambda+\varrho}(x_1,\dots,x_{n-1},x_n)=\sum_{\sigma\in S_n}\epsilon_\sigma x_1^{\ell_{\sigma(1)}}\dots x_n^{\ell_{\sigma(n)}},$$

setting $x_n=0$ we get the sum restricted only on the terms for which $\sigma(n)=n$ or

$$A_{\lambda+\varrho}(x_1,\dots,x_{n-1},0)=\sum_{\sigma\in S_{n-1}}\epsilon_\sigma x_1^{\ell_{\sigma(1)}}\dots x_{n-1}^{\ell_{\sigma(n-1)}}$$

now $\ell_i=h_i+n-i=(h_i+1)+(n-1)-i$ and so in $(n-1)-$ variables:

$$A_{\lambda+\varrho}(x_1,\dots,x_{n-1},0)=A_{\lambda+\varrho+\underline{1}}(x_1,\dots,x_{n-1})=\prod_{i=1}^{n-1}x_iA_\lambda(x_1,\dots,x_{n-1}),$$

and so $S_\lambda(x_1,\dots,x_{n-1},0)=S_\lambda(x_1,\dots,x_{n-1})$. Thus we see that:

Proposition. *Under the evaluation of x_n to 0 the Schur functions S_λ vanish, if height$(\lambda)=n$ otherwise they map to the corresponding Schur functions in $(n-1)-$variables.*

One uses these remarks as follows. Consider a fixed degree n, for any m let S_m^n be the space of symmetric functions of degree n in m variables.

From the theory of Schur functions the space S_m^n has as basis the functions $S_\lambda(x_1,\dots,x_m)$ where $\lambda\vdash n$ has heigth $\leq m$. Under the evaluation $x_m\to 0$ we have a map $S_m^n\to S_{m-1}^n$. We have proved that this map is an isomorphism as soon as $m>n$ hence all identities which we prove for symmetric functions in n variables of degree n are valid in any number of variables, we have proved.

Theorem. *The formal ring of symmetric functions in infinitely many variables has as basis all Schur functions S_λ, restriction to symmetric functions in $m-$variables sets to 0 all S_λ with height $>m$.*

When using partitions it is often more useful to describe a partition by expliciting the number of parts with 1 element, the number of parts with 2 elements and so on. Thus one writes a partition as $1^{a_1}2^{a_2}\dots i^{a_i}\dots$.

We want to prove now that for the elementary symmetric functions we have

$$e_h=S_{1^h}. \tag{3.2.2}$$

According to our previous discussion we can set all the variables $x_i,\ i > h$ to 0. Then e_h reduces to $\prod_{i=1}^h x_i$ as well as S_{1^h} from 6.2.1.

3.3 Next we want to discuss the value of $S_\lambda(1/x_1, 1/x_2, \dots, 1/x_n)$. We see that substituting x_i with $1/x_i$ in the matrix M_λ and multiplying the j^{th} column by $x_j^{m_1+n-1}$ we obtain a matrix which equals, up to rearranging the rows, that of the partition $\lambda' := m_1', m_2', \dots, m_n'$ where $m_i + m_{n-i+1}' = m_1$. Up to a sign thus:

$$(x_1x_2\dots x_n)^{m_1+n-1}A_{\lambda+\varrho}(1/x_1, \dots, 1/x_n) = A_{\lambda'+\varrho}.$$

For the Schur function we have to apply the procedure to both numerator and denominator so that the signs cancel and we get $S_\lambda(1/x_1, 1/x_2, \dots, 1/x_n) = (x_1x_2\dots x_n)^{-m_1}S_{\lambda'}$.

If we use the diagram notation for partitions we easily visualize λ' by inserting λ in a rectangle of base m_1 and then taking its complement.

4 Cauchy formulas

4.1 The formulas we want to discuss have important applications in representation theory, for the moment we wish to present them as purely combinatorial identities.

$$\text{(C1)} \qquad \prod_{i,j=1,n} \frac{1}{1 - x_iy_j} = \sum_\lambda S_\lambda(x)S_\lambda(y)$$

the right hand side is the sum over all partitions.

$$\text{(C2)} \qquad \prod_{i\leq j=1,n} \frac{1}{1 - x_ix_j} = \sum_{\lambda\in\Lambda_{ec}} S_\lambda(x),$$

if n is even

$$\text{(C3)} \qquad \prod_{i<j=1,n} \frac{1}{1 - x_ix_j} = \sum_{\lambda\in\Lambda_{er}} S_\lambda(x),$$

Here Λ_{ec}, resp. Λ_{er} indicates the set of diagrams with rows (resp. columns) of even length.

$$\text{(C4)} \qquad \prod_{i=1,\ j=1}^{n,\ m} (1 + x_iy_j) = \sum_\lambda S_\lambda(x)S_{\tilde\lambda}(y).$$

where $\tilde\lambda$ denotes the dual partition (1.1) obtained exchanging rows and columns.

We prove only the first one. We offer two proofs:

1. It can be deduced (in a way similar to the computation of the Vandermonde determinant) considering the determinant of the $n \times n$ matrix:

$$A := (a_{ij}), \text{ with } a_{ij} = \frac{1}{1 - x_i y_j}.$$

Claim

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

Subtracting the first row to the i^{th} one has a new matrix (b_{ij}) where:

$$b_{1j} = a_{1j}, \text{ and for } i > 1, \; b_{ij} = \frac{1}{1 - x_i y_j} - \frac{1}{1 - x_1 y_j} = \frac{(x_i - x_1) y_j}{(1 - x_i y_j)(1 - x_1 y_j)}$$

thus from the i^{th} row $i > 1$ one can extract from the determinant the factor $x_i - x_1$ and from the j^{th} columnn the factor $\frac{1}{1-x_1 y_j}$.

Thus the given determinant is the product of $\prod_{i=2}^{n} \frac{(x_i - x_1)}{(1 - x_1 y_i)}$ with the determinant

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ \frac{y_1}{1-x_2 y_1} & \frac{y_2}{1-x_2 y_2} & \dots & \dots & & \frac{y_n}{1-x_2 y_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{y_1}{1-x_n y_1} & \frac{y_2}{1-x_n y_2} & \dots & \dots & & \frac{y_n}{1-x_n y_n} \end{pmatrix} \tag{4.1.1}$$

subtracting the first column to the i^{th} we get the terms $\frac{y_i - y_1}{(1-x_j y_1)(1-x_j y_i)}$ thus we end extracting the product $\prod_{i=2}^{n} \frac{(y_i - y_1)}{(1 - x_i y_1)}$ and we are left with the determinant of the same type of matrix but without the variables x_1, y_1, we can thus finish by induction.

Now we can develop the determinant by developing each element $\frac{1}{1-x_i y_j} = \sum_{k=0}^{\infty} x_i^k y_j^k$ or in matrix form each row (resp. column) as a sum of infinitely many rows (or columns).

By multilinearity in the rows the determinant is a sum of determinants of matrices:

$$\sum_{k_1=0}^{\infty} \dots \sum_{k_n=0}^{\infty} det(A_{k_1,k_2,\dots,k_n}), \; A_{k_1,k_2,\dots,k_n} := ((x_i y_j)^{k_i}).$$

Clearly $det(A_{k_1,k_2,\dots,k_n}) := \prod_i x_i^{k_i} det(y_j^{k_i})$. This is zero if the k_i are not distinct, otherwise we reorder the sequence k_i so to be decreasing at at the same time we must introduce a sign, collecting all the terms in which the k_i are a permutation of a given sequence $\lambda + \rho$ we get the term $A_{\lambda+\varrho}(x) A_{\lambda+\varrho}(y)$. Finally:

$$\frac{V(x)V(y)}{\prod_{i,j=1,n}(1 - x_i y_j)} = \sum_{\lambda} A_{\lambda+\varrho}(x) A_{\lambda+\varrho}(y).$$

From this the required identity follows.

2. Change the matrix to $\frac{1}{x_i - y_j}$ using the fact that

$$V(x_1^{-1}, \dots, x_n^{-1}) = (-\prod_i x_i)^{1-n} V(x_1, \dots, x_n)$$

and develop the determinant as sum of fractions $\frac{1}{\prod(x_i - y_{\sigma(i)})}$. Write it as a rational function $\frac{f(x,y)}{\prod_{i,j=1,n}(x_i - y_j)}$ we see immediately that $f(x,y)$ is alternating in both x, y of total degree $n^2 - n$, hence $f(x,y) = cV(x)V(y)$ for some constant c, which will appear in the formula $C1$. Comparing in degree 0 we see that $C1$ holds.

Let us remark that Cauchy formula holds also when $m \leq n$ so that $\prod_{i=1}^n \prod_{j=1}^m \frac{1}{1 - x_i y_j t}$ is obtained from $\prod_{i=1}^n \prod_{j=1}^m \frac{1}{1 - x_i y_j t}$ by setting $y_j = 0$, $\forall m < j \leq n$. By 6.2 we get:

$$\prod_{i=1}^n \prod_{j=1}^m \frac{1}{1 - x_i y_j t} = \sum_{\lambda \vdash n,\, ht(\lambda) \leq m} S_\lambda(x_1, \dots, x_n) S_\lambda(y_1, \dots, y_m).$$

5 Characters of the symmetric group

The character theory of the symmetric is developed in a combinatorial way due to Frobenius.

5.1 The conjugacy classes of S_n are parametrized by partitions of n. We want to parametrize also irreducible characters by partitions. Thus obtaining a character table consisting, given two partitions λ, μ, to compute the value $c_\lambda(\mu)$ of the character of an element of the conjugacy class C_μ on an irreducible representation M_λ parametrized by the partition λ.

The final result of this analysis is expressed in compact form by the theory of symmeric functions.

Recall first that we denote $\psi_k(x) = \sum_{i=1}^n x_i^k$. For a partition $\mu \vdash n := k_1, k_2, \dots, k_n$ denote by:

$$\psi_\mu(x) := \psi_{k_1}(x) \psi_{k_2}(x) \dots \psi_{k_n}(x).$$

Using the fact that the Schur functions are an integral basis of the symmetric functions there exist (unique) integers $c_\lambda(\mu)$ for which:

$$\psi_\mu(x) = \sum_\lambda c_\lambda(\mu) S_\lambda(x). \tag{5.1.1}$$

We interpret these numbers as class functions c_λ on the symmetric group

$$c_\lambda(C_\mu) := c_\lambda(\mu)$$

and have.

Theorem Frobenius. *$c_\lambda(\mu)$ is the table of irreducible characters of S_n.*

Step 1 First we shall prove that the class functions c_λ are orthonormal.

Step 2 Next we shall express these functions as integral linear combinations of permutation characters.

Step 3 Finally we shall be able to conclude our theorem.

Step 1 In order to follow Frobenius approach we go back to symmetric functions in n variables $x_1, x_2, \dots, x_n$. We shall freely use the Schur functions and the Cauchy formula for symmetric functions:

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

proved in §4.1.

We change its right hand side as follows. Compute:

$$log(\prod_{i,j=1}^{n} \frac{1}{1 - x_i y_j}) = \sum_{i,j=1}^{n} \sum_{h=1}^{\infty} \frac{(x_i y_j)^h}{h} =$$

(5.1.2) $$\sum_{h=1}^{\infty} \sum_{i,j=1}^{n} \frac{(x_i y_j)^h}{h} = \sum_{h=1}^{\infty} \frac{\psi_h(x)\psi_h(y)}{h}.$$

Taking the exponential we get the following expression:

(5.1.3) $$exp(\sum_{h=1}^{\infty} \frac{\psi_h(x)\psi_h(y)}{h}) = \sum_{k=0}^{\infty} \frac{1}{k!} (\sum_{h=1}^{\infty} \frac{\psi_h(x)\psi_h(y)}{h})^k =$$

(5.1.4) $$\sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\sum_{i=1}^{\infty} k_i = k} \binom{k}{k_1\, k_2\, \dots} \frac{\psi_1(x)^{k_1}\psi_1(y)^{k_1}}{1} \frac{\psi_2(x)^{k_2}\psi_2(y)^{k_2}}{2^{k_2}} \frac{\psi_3(x)^{k_3}\psi_3(y)^{k_3}}{3^{k_3}} \dots .$$

Let us further manipulate this expression, remark that a way to present a partition is to give the number of times that each number i appears.

If i appears k_i times in a partition μ, the partition is indicated by:

(5.1.5) $$\mu := 1^{k_1} 2^{k_2} 3^{k_3} \dots i^{k_i} \dots .$$

Let us indicate by

$$n(\mu) = a(\mu)b(\mu) := k_1! 1^{k_1} k_2! 2^{k_2} k_3! 3^{k_3} \dots k_i! i^{k_i} \dots$$

(5.1.6) $$a(\mu) := k_1!k_2!k_3!\dots k_i!\dots,\ b(\mu) := 1^{k_1}2^{k_2}3^{k_3}\dots i^{k_i}\dots$$

then 5.1.4 becomes

$$\sum_\mu \frac{1}{n(\mu)}\psi_\mu(x)\psi_\mu(y).$$

5.2 We need to interpret now the number $n(\mu)$:

Proposition. *If $s \in C_\mu$, $n(\mu)$ is the order of the centralizer G_s of s and $|C_\mu|n(\mu) = n!$.*

Proof. Let us write the permutation s as a product of a list of cycles c_i. If g centralizes s we have that the cycles gc_ig^{-1} are a permutation of the given list of cycles.

It is clear that in this way we get all possible permutations of the cycles of equal length.

Thus we have a surjective homomorphism of G_s to a product of symmetric groups $\prod S_{k_i}$, its kernel H is formed by permutations which fix each cycle.

A permutation of this type is just a product of permutations, each on the set of indeces appearing in the corresponding cycle, and fixing it.

For a full cycle the centralizer is the cyclic group generated by the cycle, so H is a product of cyclic groups of order the length of each cycle. The formula follows. □

Let us now substitute in the identity:

$$\sum_{\mu\vdash n} \frac{1}{n(\mu)}\psi_\mu(x)\psi_\mu(y) = \sum_{\lambda\vdash n} S_\lambda(x)S_\lambda(y)$$

the expression $\psi_\mu = \sum_\lambda c_\lambda(\mu)S_\lambda$ and get:

(5.2.1) $$\sum_{\mu\vdash n} \frac{1}{n(\mu)}c_{\lambda_1}(\mu)c_{\lambda_2}(\mu) = \begin{cases} 0 & \text{if} \quad \lambda_1 \neq \lambda_2 \\ 1 & \text{if} \quad \lambda_1 = \lambda_2 \end{cases}.$$

We have thus that the class functions c_λ are an orthonormal basis completing Step 1.

5.3 *Step 2* We consider now some permutation characters (cf. 2.2). Take a partition $\lambda := h_1, h_2, \dots, h_k$ of n. Consider then the subgroup $S_\lambda := S_{h_1} \times S_{h_2} \times \dots \times S_{h_k}$ and the permutation representation on:

(5.3.1) $$S_n/S_{h_1} \times S_{h_2} \times \dots \times S_{h_k}$$

we will indicate by β_λ the corresponding character.

This character is computed with 2.3.2 $\chi(g) = \sum_i \frac{|G(g)|}{|H(g_i)|}$ of Chapter 5 applied to the case $G/H = S_n/S_{h_1} \times S_{h_2} \times \dots \times S_{h_k}$ and for a permutation g relative to a partition $\mu := 1^{p_1}2^{p_2}3^{p_3}\dots i^{p_i}\dots n^{p_n}$.

A conjugacy class in $S_{h_1} \times S_{h_2} \times \ldots \times S_{h_k}$ is given by k partitions $\mu_i \vdash h_i$ of the numbers $h_1, h_2, \ldots, h_k$; the conjugacy class of type μ intersected with $S_{h_1} \times S_{h_2} \times \ldots \times S_{h_k}$ gives all possible k tuple of partitions $\mu_1, \mu_2, \ldots, \mu_k$ of type

$$\mu_h := 1^{p_{1h}} 2^{p_{2h}} 3^{p_{3h}} \ldots i^{p_{ih}} \ldots$$

and:

$$\sum_{h=1}^{k} p_{ih} = p_i.$$

In a more formal way we may define a sum of two partitions $\lambda = 1^{p_1} 2^{p_2} 3^{p_3} \ldots i^{p_i} \ldots$, $\mu = 1^{q_1} 2^{q_2} 3^{q_3} \ldots i^{q_i} \ldots$ as the partition:

$$\lambda + \mu := 1^{p_1+q_1} 2^{p_2+q_2} 3^{p_3+q_3} \ldots i^{p_i+q_i} \ldots$$

and remark that, with the notations of 5.1.7 $b(\lambda + \mu) = b(\lambda) b(\mu)$.

We are thus decomposing $\mu = \sum_{i=1}^{k} \mu_i$ and we have $b(\mu) = \prod b(\mu_i)$.

The cardinality $m_{\mu_1, \mu_2, \ldots, \mu_k}$ of the conjugacy class $\mu_1, \mu_2, \ldots, \mu_k$ in $S_{h_1} \times S_{h_2} \times \ldots \times S_{h_k}$ is:

$$m_{\mu_1, \mu_2, \ldots, \mu_k} = \prod_{j=1}^{k} \frac{h_j!}{n(\mu_j)} = \prod_{j=1}^{k} \frac{h_j!}{a(\mu_j)} \frac{1}{b(\mu)}$$

Now

$$\prod_{j=1}^{k} a(\mu_j) = \prod_{h=1}^{k} (\prod_{i=1}^{n} p_{ih}!)$$

So we get:

$$m_{\mu_1, \mu_2, \ldots, \mu_k} = \frac{1}{n(\mu)} \prod_{j=1}^{k} h_j! \prod_{i=1}^{n} \binom{p_i}{p_{i1} p_{i2} \ldots p_{ik}}.$$

Finally for the number $\beta_\lambda(\mu)$ we have:

$$\beta_\lambda(\mu) = \frac{n(\mu)}{\prod_{i=1}^{k} h_i!} \sum_{\mu = \sum_{i=1}^{k} \mu_i,\ \mu_i \vdash h_i} m_{\mu_1, \mu_2, \ldots, \mu_k} = \sum_{\mu = \sum_{i=1}^{k} \mu_i,\ \mu_i \vdash h_i} \prod_{i=1}^{n} \binom{p_i}{p_{i1} p_{i2} \ldots p_{ik}}.$$

This sum is manifestly the coefficient of $x_1^{h_1} x_2^{h_2} \ldots x_k^{h_k}$ in the symmetric function $\psi_\mu(x)$.

In fact when we expand

$$\psi_\mu(x) = \psi_1(x)^{p_1} \psi_2(x)^{p_2} \ldots \psi_i(x)^{p_i} \ldots$$

for each factor $\psi_k(x) = \sum_{i=1}^{n} x_i^k$ one selects the index of the variable chosen and constructs a corrresponding product monomial.

For each such monomial denote by p_{ij} the number of choices of the term x_j^i in the p_i factors $\psi_i(x)$, we have $\prod_i \binom{p_i}{p_{i1} p_{i2} \ldots p_{ik}}$ such choices and they contribute to the monomial $x_1^{h_1} x_2^{h_2} \ldots x_k^{h_k}$ if and only if $\sum_i i p_{ij} = h_j$.

Thus if Σ_λ denotes the sum of all monomials in the orbit of $x_1^{h_1}x_2^{h_2}\dots x_k^{h_k}$ we get the formula:

$$\psi_\mu(x)=\sum_\lambda \beta_\lambda(\mu)\Sigma_\lambda(x). \tag{5.3.2}$$

5.4 We wish to expand now the basis $\Sigma_\lambda(x)$ in terms of the basis $S_\lambda(x)$ and conversely:

$$\Sigma_\lambda(x)=\sum_\mu p_{\lambda,\mu}S_\mu(x),\ S_\lambda(x)=\sum_\mu k_{\lambda,\mu}\Sigma_\mu(x) \tag{5.4.1}$$

In order to explicit some information about the matrices:

$$(p_{\lambda,\mu}),\ (k_{\lambda,\mu})$$

recall that the partitions are totally ordered by lexicographic ordering.

We also order the monomials by the lexicographic ordering of the sequence of exponents $h_1,h_2,\dots,h_n$ of the variables $x_1,x_2,\dots,x_n$.

We remark that the ordering of monomials has the following immediate property:

If M_1,M_2,N are 3 monomials and $M_1<M_2$ then $M_1N<M_2N$.

For any polynomial $p(x)$ we can thus select the leading monomial $l(p)$ and for two polynomials $p(x),q(x)$ we have:

$$l(pq)=l(p)l(q).$$

For a partition $\mu\vdash n:=h_1\geq h_2\geq\dots\geq h_n$ the leading monomial of Σ_μ is

$$x^\mu:=x_1^{h_1}x_2^{h_2}\dots x_n^{h_n}.$$

Similarly the leading monomial of the alternating function $A_{\mu+\varrho}(x)$ is:

$$x_1^{h_1+n-1}x_2^{h_2+n-2}\dots x_n^{h_n}=x^{\mu+\varrho}.$$

We compute now the leading monomial of the Schur function S_μ, using all the definitions and notations of §6.1, since

$$x^{\mu+\varrho}=l(A_{\mu+\varrho}(x))=l(S_\mu(x)V(x))=l(S_\mu(x))x^\varrho$$

we deduce that:

$$l(S_\mu(x))=x^\mu.$$

This computation has the following immediate consequence:

Proposition. *The matrices $P:=(p_{\lambda,\mu}),\ Q:=(k_{\lambda,\mu})$ are upper triangular with 1 on the diagonal.*

Proof. A symmetric polynomial with leading coefficient x^μ is clearly equal to Σ_μ plus a linear combination of the $\Sigma_\lambda,\ \lambda<\mu$ this proves the claim for the matrix Q; the matrix P is the inverse of Q and the claim follows. $\square$

We can now conclude:

Theorem. *i)* $\beta_\lambda = c_\lambda + \sum_{\phi<\lambda} k_{\phi,\lambda} c_\phi,\ k_{\phi,\lambda} \in \mathbb{N}$.
$c_\lambda = \sum_{\mu \geq \lambda} p_{\mu\lambda} b_\mu$.

ii) The functions $c_\lambda(\mu)$ are a list of the irreducible characters of the symmetric group.

iii) (Frobenius Theorem) $\chi_\lambda = c_\lambda$.

Proof. From the various definitions we get:

(5.4.2)
$$c_\lambda = \sum_\phi p_{\phi,\lambda} b_\phi,\ \beta_\lambda = \sum_\phi k_{\phi,\lambda} c_\phi,$$

therefore the functions c_λ are virtual characters. Since they are orthonormal they are $\pm$ the irreducible characters.

From the recursive formulas it follows that $\beta_\lambda = c_\lambda + \sum_{\phi<\lambda} k_{\phi,\lambda} c_\phi,\ m_{\lambda,\phi} \in \mathbb{Z}$. Since β_λ is a character it is a positive linear combination of the irreducible characters, it follows that each c_λ is an irreducible character and that the coefficients $k_{\phi,\lambda} \in \mathbb{N}$ represent the multiplicities of the decomposition of the permutation representation into irreducible components.[5]

□

Remark The basic formula $\psi_\mu(x) = \sum_\lambda c_\lambda(\mu) S_\lambda(x)$ can be multiplied by the Vandermonde determinant getting

(5.4.3)
$$\psi_\mu(x) V(x) = \sum_\lambda c_\lambda(\mu) A_{\lambda+\varrho}(x)$$

now we may apply the leading monomial theory and deduce that $c_\lambda(\mu)$ is the coefficient in $\psi_\mu(x)V(x)$ belonging to the leading monomial $x^{\lambda+\rho}$ of $A_{\lambda+\varrho}$.

This furnishes a possible algorithm, we will discuss later some features of this formula.

5.5 There is a nice interpretation of the theorem of Frobenius which I want to describe.

Definition. *The linear isomorphism between characters of S_n and symmetric functions of degree n which assigns to χ_λ the Schur function S_λ is called the* **Frobenius character.** *It is denoted by $\chi \to F(\chi)$.*

By Frobenius theorem the Frobenius character can be computed by the formula (cf. 5.1.2):

$$F(\chi) = \frac{1}{n!} \sum_{\sigma \in S_n} \chi(\sigma) \phi_{\mu(\sigma)}(x) = \sum_{\mu \vdash n} \frac{\chi(\mu)}{n(\mu)} \phi_\mu(x)$$

Recall that $n(\mu)$ is the order of the centralizer of a permutation with cycle structure μ. This shows the following important multiplicative behaviour of the Frobenius character.

[5]The numbers $k_{\phi,\lambda}$ are called Kostka numbers. As we shall see they count some combinatorial objects called semistandard tableaux.

Theorem. *Given two representations V, W of S_m, S_n respectively we have:*

$$F(Ind_{S_n \times S_m}^{S_{n+m}}(V \otimes W)) = F(V)F(W). \tag{5.5.1}$$

Proof. Let us denote by χ the character of $Ind_{S_m \times S_n}^{S_{m+n}}(V \otimes W)$. Recall the discussion of induced characters in Chap 5, and formula 2.3.1 $\chi(g) = \sum_i \frac{|G(g)|}{|H(g_i)|}\chi_V(g_i)$, where $|G(g)|$ is the order of the centralizer of g in G the O_i are the conjugacy classes in H in the conjugacy class of g in G. In our case for $Ind_{S_m \times S_n}^{S_{m+n}}(V \otimes W)$ we deduce that $\chi(\sigma) = 0$ unless σ is conjugate to an element (a, b) of $S_n \times S_m$. In term of partitions the partitions $\nu \vdash n + m$ which contribute to the characters are the ones of type $\lambda \oplus \mu$. In the language of permutations formula 2.3.1 becomes:

$$\chi(\nu) = \sum_{\nu = \lambda + \mu} \frac{n(\lambda + \mu)}{n(\lambda)n(\mu)} \chi_V(\lambda)\chi_W(\mu)$$

since $\phi_{\lambda \oplus \mu} = \phi_\lambda \phi_\mu$ we obtain for $F(\chi)$:

$$\sum_{\nu \vdash m+n} \frac{\chi(\nu)\phi_\nu}{n(\nu)} = \sum_\nu \frac{\phi_\nu}{n(\nu)} \sum_{\nu = \lambda + \mu} \frac{n(\lambda + \mu)}{n(\lambda)n(\mu)} \chi_V(\lambda)\chi_W(\mu) = \sum_{\lambda \vdash m, \mu \vdash n} \frac{\chi_V(\lambda)\chi_W(\mu)}{n(\lambda)n(\mu)} \phi_\lambda \phi_\mu$$

□

5.6 We discuss here a complement to the representation theory of S_n.

It will be necessary to work formally with symmetric functions in infinitely many variables, a formalism which has been justified in 1.1.

With this in mind we think of the identities 5.1.1, 5.3.4, 5.4.1 etc. as identities in infinitely many variables.

First of all a convention. If we are given a representation of a group on a graded vector space $U := \{U_i\}_{i=0}^\infty$ (i.e. a representation on each U_i) its character is usually written as a power series with coefficients in the character ring in a variable q:[6]

$$\chi_U(t) := \sum_i \chi_i q^i. \tag{5.6.1}$$

Where χ_i is the character of the representation U_i.

Definition. *The expression 5.6.1 is called a graded character.*

Graded characters have some formal similarities with characters. Given two graded representations $U = \{U_i\}_i, V = \{V_i\}_i$ we have their direct sum, and their tensor product

$$(U \oplus V)_i := U_i \oplus V_i, \quad (U \otimes V)_i := \oplus_{h=0}^i U_h \otimes V_{i-h}.$$

[6]It is now quite usual to use q as variable since it often appears coming from computations on finite fields where $q = p^r$ or as quantum deformation parameter

For the graded characters we have clearly:

(5.6.2) $$\chi_{U\oplus V}(q) = \chi_U(q) + \chi_V(q),\ \ \chi_{U\otimes V}(q) = \chi_U(q)\chi_V(q).$$

Let us consider a simple example.[7]

Lemma Molien's formula. *Given a linear operator A on a vector space U its action on the symmetric algebra $S(U)$ has as graded character:*

(5.6.3) $$\sum_{i=0}^{\infty} tr(S^i(A))q^i = \frac{1}{det(1-qA)}$$

Its action on the exterior algebra $\wedge U$ has as graded character:

(5.6.4) $$\sum_{i=0}^{\dim U} tr(\wedge^i(A))q^i = \det(1+qA)$$

Proof. Since for every symmetric power $S^k(U)$ the character of the operator induced by A is a polynomial in A it is enough to prove the formula, by continuity and invariance, when A is diagonal.

Take a basis of eigenvectors $u_i,\ i = 1, \dots, n$ with eigenvalue λ_i.

Then $S[U] = S[u_1] \otimes S[u_2] \otimes \dots \otimes S[u_n]$ and $S[u_i] = \sum_{h=0}^{\infty} Fu_i^h$.

The graded character of $S[u_i]$ is $\sum_{h=0}^{\infty} \lambda_i^h q^h = \frac{1}{1-\lambda_i q}$ hence:

$$\chi_{S[U]}(q) = \prod_{i=1}^{n} \chi_{S[u_i]}(q) = \frac{1}{\prod_{i=1}^{n}(1-\lambda_i q)} = \frac{1}{\det(1-qA)}.$$

Similarty $\wedge U = \wedge[u_1] \otimes \wedge[u_2] \otimes \dots \otimes \wedge[u_n]$ and $\wedge[u_i] = F \oplus Fu_i$ hence

$$\chi_{\wedge[U]}(q) = \prod_{i=1}^{n} \chi_{\wedge[u_i]}(q) = \prod_{i=1}^{n}(1+\lambda_i q) = \det(1+qA).$$

□

We apply the previous discussion to S_n acting on the space $\mathbb{C}^n$ permuting the coordinates and the representation that it induces on the polynomial ring $\mathbb{C}[x_1, x_2, \dots, x_n]$.

We denote by $\sum_{i=0}^{\infty} \chi_i q^i$ the corresponding graded character.

If σ is a permutation with cycle decomposition of lengths $\mu(\sigma) = \mu := m_1, m_2, \dots m_k$ the standard basis of $\mathbb{C}^n$ decomposes into $k-$cycles each of length m_i. On the subspace relative to a cycle of length m, σ acts with eigenvalues the m-roots of 1 and

$$\det(1-q\sigma) = \prod_i \prod_{j=1}^{m}(1 - e^{\frac{j2\pi\sqrt{-1}}{m}} q) = \prod_i (1-q^m)$$

[7]strictly speaking we are not treating now a group, but the set of all matrices under multiplication, which is only a semigroup, for this set tensor product of representation makes sense but not duality.

Thus the graded character of σ acting on the polynomial ring is

$$\frac{1}{\det(1-q\sigma)} = \prod_i \sum_{j=0}^{\infty} q^{jm_i} = \prod_i \psi_{m_i}(1,q,q^2,\dots,q^k,\dots) =$$
$$\psi_\mu(1,q,q^2,\dots,q^k,\dots) = \sum_{\lambda\vdash n} \chi_\lambda(\sigma) S_\lambda(1,q,q^2,\dots,q^k,\dots)$$

To summarize

Theorem. *The graded character of S_n acting on the polynomial ring is*

$$\sum_{\lambda\vdash n} \chi_\lambda S_\lambda(1,q,q^2,\dots,q^k,\dots)$$

We have a corollary of this formula. If $\lambda = h_1 \geq h_2 \cdots \geq h_n$, the term of lowest degree in q in $S_\lambda(1,q,q^2,\dots,q^k,\dots)$ is clearly given by the leading term $x_1^{h_1} x_2^{h_2} \dots x_n^{h_n}$ computed in $1,q,q^2,\dots,q^n$ and this gives $q^{h_1+2h_2+3h_3+\dots+nh_n}$. We deduce that the representation M_λ of S_n appears for the first time in degree $h_1+2h_2+3h_3+\dots+nh_n$ and in this degree it appears with multiplicity 1. This particular submodule of $\mathbb{C}[x_1,x_2,\dots,x_n]$ is called the *Specht module* and it plays an important role.[8]

Now we want to discuss another related representation.

Recall first that $\mathbb{C}[x_1,x_2,\dots,x_n]$ is a free module over the ring of symmetric functions $\mathbb{C}[\sigma_1,\sigma_2,\dots,\sigma_n]$ of rank $n!$. It follows that, for every choice of the numbers $\underline{a} := a_1,\dots,a_n$ the ring $R_{\underline{a}} := \mathbb{C}[x_1,x_2,\dots,x_n]/ < \sigma_i - a_i >$ constructed from $\mathbb{C}[x_1,x_2,\dots,x_n]$ modulo the ideal generated by the elements $\sigma_i - a_i$, is of dimension $n!$ and a representation of S_n.

We claim that it is always the regular representation.

First we prove it in the case in which the polynomial $t^n - a_1 t^{n-1} + a_2 t^{n-2} - \dots + (-1)^n a_n$ has distinct roots $\alpha_1,\dots,\alpha_n$, this means that the ring $\mathbb{C}[x_1,x_2,\dots,x_n]/ < \sigma_i - a_i >$ is the coordinate ring of the set of $n!$ points $\alpha_{\sigma(1)},\dots,\alpha_{\sigma(n)}$, $\sigma \in S_n$ and this is clearly the regular representation.

The condition for a polynomial to have distinct roots is open in the coefficients and given by the non vanishing of the discriminant.

It is easily seen that the character of $R_{\underline{a}}$ is continuous in $\underline{a}$ and, since the characters of a finite group are a discrete set this implies that the character is constant.

It is of particular interest (combinatorial and geometric) to analyze the special case $\underline{a} = 0$ and the ring $R := \mathbb{C}[x_1,x_2,\dots,x_n]/ < \sigma_i >$ which is a graded algebra affording the regular representation.

Thus the graded character $\chi_R(q)$ of R is a graded form of the regular representation.

To compute it notice that as a graded representation we have an isomorphism

$$\mathbb{C}[x_1,x_2,\dots,x_n] = R \otimes \mathbb{C}[\sigma_1,\sigma_2,\dots,\sigma_n]$$

[8] it appears in the Springer representation for instance, cf. .

and thus an identity of graded characters.

The ring $\mathbb{C}[\sigma_1, \sigma_2, \ldots, \sigma_n]$ has the trivial representation, by definition, and generators in degree $1, 2, \ldots, n$ so its graded character is just $\prod_{i=1}^n (1-q^i)^{-1}$ and finally we deduce:

Theorem.

$$\chi_R(q) = \sum_{\lambda \vdash n} \chi_\lambda S_\lambda(1, q, q^2, \ldots, q^k, \ldots) \prod_{i=1}^n (1-q^i)$$

Notice then that the series $S_\lambda(1, q, q^2, \ldots, q^k, \ldots) \prod_{i=1}^n (1-q^i)$ represent the multiplicities of χ_λ in the various degrees of R and thus are polynomials with positive coefficients with sum the dimension of χ_λ.

Exercise Prove that the Specht module has non zero image in the quotient ring $R := \mathbb{C}[x_1, x_2, \ldots, x_n]/ < \sigma_i >$.

The ring $R := \mathbb{Z}[x_1, x_2, \ldots, x_n]/ < \sigma_i >$ has an interesting geometric interpretation as the *cohomology algebra* of the flag variety. This can be understood as the space of all decompositions $\mathbb{C}^n = V_1 \perp V_2 \perp \cdots \perp V_n$ into orthogonal 1-dimensional subspaces. The action of the symmetric group is iduced by the topological action permuting the summands of the decomposition.

6 The hook formula

6.1 We want to deduce now a formula, due to Frobenius, for the dimension $d(\lambda)$ of the irreducible representation M_λ of the symmetric group.

From 5.4.3 applied to the partition 1^n, corresponding to the conjugacy class of the identity, we obtain:

(6.1.1)
$$(\sum_{i=1}^n x_i)^n V(x) = \sum_\lambda d(\lambda) A_{\lambda+\varrho}(x)$$

Write the development of the Vandermonde determinant as $\sum_{\sigma \in S_n} \epsilon_\sigma \prod_{i=1}^n x_i^{\sigma(n-i+1)-1}$.

Letting $\lambda + \rho = \ell_1 > \ell_2 > \cdots > \ell_n$ the number $d(\lambda)$ is the coefficient of $\prod_i x_i^{\ell_i}$ in

$$(\sum_{i=1}^n x_i)^n \sum_{\sigma \in S_n} \epsilon_\sigma \prod_{i=1}^n x_i^{\sigma(n-i+1)-1}.$$

Thus a term $\epsilon_\sigma \binom{n}{k_1\, k_2 \ldots k_n} \prod_{i=1}^n x_i^{\sigma(n-i+1)-1+k_i}$ contributes to $\prod_i x_i^{\ell_i}$ if and only if $k_i = \ell_i - \sigma(n-i+1) + 1$. We deduce

$$d(\lambda) = \sum_{\substack{\sigma \in S_n \mid \forall i \\ \ell_i - \sigma(n-i+1)+1 \geq 0}} \epsilon_\sigma \frac{n!}{\prod_{i=1}^n (\ell_i - \sigma(n-i+1)+1)!}$$

We change the term

$$n!\prod_{i=1}^{n}\frac{1}{(\ell_i-\sigma(n-i+1)+1)!}=\frac{n!}{\prod_{i=1}^{n}\ell_i!}\prod_{i=1}^{n}\prod_{\substack{0\le k\le\\ \sigma(n-i+1)-2}}(\ell_i-k)$$

and remark that this formula makes sense, and it is 0, if σ does not satisfy the restriction $\ell_i-\sigma(n-i+1)+1\ge 0$.

Thus

$$d(\lambda)=\frac{n!}{\prod_{i=1}^{n}\ell_i!}\overline{d}(\lambda),\quad \overline{d}(\lambda)=\sum_{\sigma\in S_n}\epsilon_\sigma\prod_{i=1}^{n}\prod_{\substack{0\le k\le\\ \sigma(n-i+1)-2}}(\ell_i-k)$$

$\overline{d}(\lambda)$ is the value of the determinant of a matrix with $\prod_{0\le k\le j-2}(\ell_i-k)$ in the $n-i+1,j$ position.

$$\begin{vmatrix}
1 & \ell_n & \ell_n(\ell_n-1) & \dots & \prod_{0\le k\le n-2}(\ell_n-k)\\
\dots & \dots & \dots & \dots & \dots\\
\dots & \dots & \dots & \dots & \dots\\
1 & \ell_i & \ell_i(\ell_i-1) & \dots & \prod_{0\le k\le n-2}(\ell_i-k)\\
\dots & \dots & \dots & \dots & \dots\\
\dots & \dots & \dots & \dots & \dots\\
1 & \ell_1 & \ell_1(\ell_1-1) & \dots & \prod_{0\le k\le n-2}(\ell_1-k)
\end{vmatrix}$$

this determinant, by elementary operations on the columns, reduces to the Vandermonde determinant in the ℓ_i with value $\prod_{i<j}(\ell_i-\ell_j)$. Thus we obtain the formula of Frobenius:

$$d(\lambda)=n!\prod_{i=1}^{n}\frac{\prod_{i<j}(\ell_i-\ell_j)}{\ell_i!} \tag{6.1.2}$$

6.2 We want to give a combinatorial interpretation of 6.1.2. Notice that in $\frac{\prod_{i<j}(\ell_i-\ell_j)}{\ell_i!}$ the $i-1$ factors of the numerator cancel the corresponding factors in the denominator leaving ℓ_i-i+1 factors. In all $\sum_i\ell_i-\sum_{i=1}^{n}(i-1)=n$ are left.

These factors can be interpreted has the *hook lengths* of the boxes of the corresponding diagram.

More precisely given a box x of a French diagram its **hook** is the set of elements of the diagram which are either on top or to the right of x, including x. E.g. we mark the hooks of $1,2;\ 2,1;\ 2,2$ in 4,3,1,1

$$\begin{array}{llll}
\cdot & & & \\
\cdot & & & \\
\cdot & \square & \cdot & \\
\cdot & \square & \square & \square
\end{array}
\qquad\qquad
\begin{array}{llll}
\square & & & \\
\square & & & \\
\square & \square & \square & \\
\cdot & \cdot & \cdot & \cdot
\end{array}
\qquad\qquad
\begin{array}{llll}
\cdot & & & \\
\cdot & & & \\
\cdot & \square & \square & \\
\cdot & \cdot & \cdot & \cdot
\end{array}$$

The total number of boxes in the hook of x is the *hook length* of x, denoted by h_x.

Now let us show that the factors in the factorial $\ell_i!$ which are not cancelled are the hooks of the boxes in the i^{th} row.

In fact let $h_i = \ell_i + i - n$ be the length of the i^{th} row, given $k > i$ let us consider the $h_{k-1} - h_k$ numbers strictly between $\ell_i - \ell_{k-1} = h_i - h_{k-1} + k - i - 1$ and $\ell_i - \ell_k = h_i - h_k + k - i$.

$h_{k-1} - h_k$ is the number of cases in the i^{th} row for which the hoook ends vertically on the $k-1$ row, and it is easily seen, since the vertical *leg* of each such hook has length $k-i$ and the horizontal *arm* length goes from $h_i - h_k$ to $h_i - h_{k-1} + 1$ that the lengths of these hooks vary between $k - i + h_i - h_k - 1$ and $k - i + h_i - h_{k-1}$ and are these given numbers.

Frobenius formula becomes thus the *hook formula*, denote by $B(\lambda)$ the set of boxes of a diagram of shape λ:

$$d(\lambda) = \frac{n!}{\prod_{x \in B(\lambda)} h_x}. \tag{6.2.1}$$

6.3 We wish to describe now a fairly simple recursive algorithm, due to Murnhagam, to compute the numbers c_λ. It is based on the knowledge of the multiplication of $\psi_k S_\lambda$ in the ring of symmetric functions.

We assume the number n of variables to be more than $k + |\lambda|$, i.e. to be in a stable range for the formula.

Let h_i denote the rows of λ, we may as well compute $\psi_k S_\lambda V(x) = \psi_k(x) A_{\lambda+\varrho}(x)$:

$$\psi_k(x) A_{\lambda+\varrho}(x) = (\sum_{i=1}^{n} x_i^k)(\sum_{s \in S_n} \epsilon_s x_{s1}^{h_1+n-1} x_{s2}^{h_1+n-2} \dots x_{sn}^{h_n}). \tag{6.3.1}$$

Indicate by $k_i = h_i + n - i$. We inspect the monomials appearing in the alternating function which is at the right of 6.3.1.

Each term is a monomial with exponents obtained from the sequence k_i by adding to one of them say k_j the number k.

If the resulting sequence has two numbers equal it cannot contribute a term to an alternating sum and so it must be dropped, otherwise reorder it getting a sequence:

$$k_1 > k_2 > \dots k_i > k_j + k > k_{i+1} > \dots k_{j-1} > k_{j+1} > \dots > k_n.$$

Then we see that the partition $\lambda' : h'_i$ associated to this sequence is:

$$h'_t = h_t, \quad \text{if } t \leq i \text{ or } t > j, \; h'_t = h_{t-1} + 1 \text{ if } i + 2 \leq t \leq j, \; h'_{i+1} = h_j + k + j - i - 1.$$

The coefficient of $S_{\lambda'}$ in $\psi_k(x) S_\lambda(x)$ is $(-1)^{j-i}$ by reordering the rows.

There is a simple way of visualizing the various partitions λ' which arise in this way.

Notice that we have modified a certain number of consecutive rows, adding a total of k new boxes. Each row except the bottom row, has been replaced by the row immediately below it plus one extra box.

This property appears saying that the new diagram λ' is thus any diagram which contains the diagram λ and such that their difference is connected, made of k boxes and it is like a "slinky"[9] i.e. it is part of the *rim* of the diagram λ' (which are the points on its boundary or the points i, j for which there is no point h, k with $i < h, j < k$ lying in the diagram).

So one has to think of a slinky made of k boxes, sliding in all possible ways down the diagram.

The sign to attribute to such a configuration si $+1$ if the number of rows occupied is odd, -1 otherwise.

For instance we can visualize $\psi_3 S_{327}$ as:

$$+\begin{array}{lll} \circ & & \\ \circ & & \\ \circ & & \\ \cdot & & \\ \cdot & \cdot & \\ \cdot & \cdot & \cdot \end{array} \qquad -\begin{array}{lll} \circ & \circ & \\ \cdot & \circ & \\ \cdot & \cdot & \\ \cdot & \cdot & \cdot \end{array} \qquad -\begin{array}{lll} \cdot & \circ & \circ \\ \cdot & \cdot & \circ \\ \cdot & \cdot & \cdot \end{array}$$

$$-\begin{array}{llll} \cdot & & & \\ \cdot & \cdot & \circ & \circ \\ \cdot & \cdot & \cdot & \circ \end{array} \qquad +\begin{array}{llllll} \cdot & & & & & \\ \cdot & \cdot & & & & \\ \cdot & \cdot & \cdot & \circ & \circ & \circ \end{array}$$

Formally one can define a $k-$slinky as a walk in the plane $\mathbb{N}^2$ made of $k-$steps, and each step is either one step down or one step right. The sign of the slinky is -1 if it occupies an even number of rows, $+1$ otherwise.

Next one defines a striped tableau of type $\mu := k_1, k_2, \ldots, k_t$ to be a tableau filled, for each $i = 1, \ldots, t$ with exactly k_i entries of the number i subject to fill a k_i-slinky. Moreover we assume that the set of boxes filled with the numbers up to i, for each i is still a diagram. E.g. a 3,4,2,5,6,3,4,1 striped diagram:

$$\begin{array}{lllllllll} 8 & & & & & & & & \\ 4 & & & & & & & & \\ 4 & 4 & 4 & 5 & & & & & \\ 3 & 3 & 4 & 5 & & & & & \\ 1 & 2 & 2 & 5 & 5 & 7 & 7 & 7 & 7 \\ 1 & 1 & 2 & 2 & 5 & 5 & 6 & 6 & 6 \end{array}$$

to such a striped tableau we associate as sign the product of the signs of all its slinkies. In our case it is the sign pattern $- - + + - + ++$ for a total $-$ sign.

Murnagham's rule can be formulated as:

$c_\lambda(\mu)$ equals the number of striped tableaux of type μ and shape λ each counted with its sign.

[9]this was explained to me by A. Garsia.

Notice that, when $\mu = 1^n$ the slinky is only one box and the condition is that the diagram is filled with all the distinct numbers $1, \dots, n$ and the filling is increasing from left to rigth and from the bottom to the top. This is the definition of a standard tableau. Its sign is always 1 and so $d(\lambda)$ equals the number of standard tableaux of shape λ.

6.4 We want to draw another important consequence of the previous multiplication formula between Newton functions and Schur functions.

Consider a module M_λ for S_n and consider $S_{n-1} \subset S_n$, we want to analyze M_λ as a representation of the subgroup S_{n-1}. For this we perform a character computation.

We introduce first a simple notation, given two partitions $\lambda \vdash m$, $\mu \vdash n$ we say that $\lambda \subset \mu$ if we have an inclusion of the corresponding Ferrer's diagrams or equivalently if each row of λ is less or equal of the corresponding row of μ.

If $n = m + 1$ we will also say that λ, μ are **adjacent**, in this case clearly μ is obtained from λ removing a box lying in a corner.

With these remarks we notice a special case of 8.1:

$$\psi_1 S_\lambda = \sum_{\mu \vdash |\lambda|+1, \lambda \subset \mu} S_\mu. \tag{6.4.1}$$

Consider now an element of S_{n-1} to which is associated a partition ν; the same element as permutation in S_n has associated partition $\nu 1$ so computing characters we have:

$$\sum_{\lambda \vdash n} c_\lambda(\nu 1) S_\lambda = \psi_{\nu 1} = \psi_1 \psi_\nu = \sum_{\tau \in \vdash (n-1)} c_\tau(\nu) \psi_1 S_\tau$$

$$= \sum_{\tau \in \vdash (n-1)} c_\tau(\nu) \sum_{\mu \in \vdash n,\, \tau \subset \mu} S_\mu. \tag{6.4.2}$$

In other words:

$$c_\lambda(\nu 1) = \sum_{\mu \in \vdash (n-1),\, \mu \subset \lambda} c_\mu(\nu), \tag{6.4.3}$$

This identity between characters becomes in module notations:

Theorem Branching rule for the symmetric group. *When restricting from S_n to S_{n-1} we have:*

$$M_\lambda = \oplus_{\mu \in \vdash (n-1),\, \mu \subset \lambda} M_\mu. \tag{6.4.4}$$

A remarkable feature of this decomposition is that each irreducible S_{n-1} module appearing in M_λ has multiplicity 1, which implies in particular that the decomposition 8.2.3 is unique.

A very convenient way to record a partition μ obtained from λ by removing a box is given marking this box with the number n. We can repeat now the branching to S_{n-2}

and get:

$$M_\lambda = \bigoplus_{\mu_2 \vdash= n-2,\ \mu_1 \vdash = n-1, \mu_2 \subset \mu_1 \subset \lambda} M_{\mu_2}. \tag{6.4.5}$$

Again we mark a pair $\mu_2 \in \vdash (n-2),\ \mu_1 \vdash (n-1), \mu_2 \subset \mu_1 \subset \lambda$ by marking the first box removed to get μ_1 with n and the second box with $n-1$.

From 4261, branching once:

$$\begin{array}{cccc} . & . & . & 8 \\ . & . & & \\ . & & & \\ . & & & \end{array} \quad + \quad \begin{array}{cccc} . & . & . & . \\ . & 8 & & \\ . & & & \\ . & & & \end{array} \quad + \quad \begin{array}{cccc} . & . & . & . \\ . & . & & \\ . & & & \\ 8 & & & \end{array}$$

Banching twice:

$$\begin{array}{cccc} . & . & 7 & 8 \\ . & . & & \\ . & & & \\ . & & & \end{array} \quad + \quad \begin{array}{cccc} . & . & . & 8 \\ . & 7 & & \\ . & & & \\ . & & & \end{array} \quad + \quad \begin{array}{cccc} . & . & . & 8 \\ . & . & & \\ . & & & \\ 7 & & & \end{array} \quad + \quad \begin{array}{cccc} . & . & . & 7 \\ . & 8 & & \\ . & & & \\ . & & & \end{array}$$

$$+ \quad \begin{array}{cccc} . & . & . & . \\ . & 8 & & \\ . & & & \\ 7 & & & \end{array} \quad + \quad \begin{array}{cccc} . & . & . & 7 \\ . & . & & \\ . & & & \\ 8 & & & \end{array} \quad + \quad \begin{array}{cccc} . & . & . & . \\ . & 7 & & \\ . & & & \\ 8 & & & \end{array} \quad + \quad \begin{array}{cccc} . & . & . & . \\ . & . & & \\ 7 & & & \\ 8 & & & \end{array}$$

In general we give the following definitions:

Given $\mu \subset \lambda$ two diagrams, the complement of μ in λ is called a **skew diagram** indicated by λ/μ. A **standard skew tabluau** of shape λ/μ consists of filling the boxes of λ/μ with distinct numbers such that each row and each column is strictly increasing.

Example of a skew tableau of shape 6527/326:

$$\begin{array}{cccccc} . & . & . & 1 & 2 & 4 \\ . & . & 2 & 3 & 4 & \\ . & 2 & & & & \\ 6 & 7 & & & & \end{array}$$

Notice that we have placed some dots in the position of the partition 327 which has been removed.

If $\mu = \emptyset$ we speak of a standard tableau. We can easily convince ourselves that, if $\lambda \vdash= n$, $\mu \vdash= n-k$ and $\mu \subset \lambda$ there is a 1-1 correspondence between:

1) Sequences $\mu = \mu_k \subset \mu_{k-1} \subset \mu_{k-2} \ldots \subset \mu_1 \subset \lambda$ with $\mu_i \vdash n-i$.

2) Standard skew diagrams of shape λ/μ filled with the numbers

$$n-k+1, n-k+2, \ldots, n-1, n.$$

The correspondence is established by associating to a standard skew tableau T the sequence of diagrams μ_i where μ_i is obtained from λ by removing the boxes occupied by the numbers $n, n-1, \ldots, n-i+1$.

When we apply the branch rule several times, passing from S_n to S_{n-k} we obtain a decomposition of M_λ into a sum of modules indexed by all possible skew standard tableaux of shape λ/μ filled with the numbers $n-k+1, n-k+2, \ldots, n-1, n$.

In particular, for a given shape $\mu \vdash n-k$ the multiplicity of M_μ in M_λ equals the number of such tableaux.

In particular we may go all the way down to S_1 and obtain a canonical decomposition of M_λ into 1-dimensional spaces indexed by all the standard tableaux of shape λ. We recover in a more precise way what we discussed in the previous paragraph.

Proposition. *The dimension of M_λ equals the number of standard tableaux of shape λ.*

It is of some interest to discuss the previous decomposition in the following way.

For every k let S_k be the symmetric group on k elements contained in S_n, so that $\mathbb{Q}[S_k] \subset \mathbb{Q}[S_n]$ as subalgebra.

Let Z_k be the center of $\mathbb{Q}[S_k]$. The algebras $Z_k \subset \mathbb{Q}[S_n]$ generate a commutative subalgebra C. In fact for every k we have that the center of $\mathbb{Q}[S_k]$ has a basis of idempotents u_λ indexed by the partitions of k.

On any irreducible representation this subalgebra, by the analysis made has a basis of common eigenvectors given by the decomposition into 1 dimensional spaces previously described.

EXERCISE Prove that the common eigenvalues of the u_λ are distinct and so this decomposition is again unique.

Remark. The decomposition just obtained is almost equivalent to selecting a basis of M_λ indexed by standard diagrams. Fixing an invariant scalar product in M_λ we immediately see by induction that the decomposition is orthogonal (because non isomorphic representations are necessarily orthogonal). If we work over $\mathbb{R}$ we can select thus a vector of norm 1 in each summand. This leaves still some sign ambiguity which can be resolved by suitable conventions. The selection of a standard basis is in fact a rather fascinating topic, it can be done in several quite unequivalent ways suggested by very different considerations, we will see some in the next chapters.

A possible goal is to exhibit not only an explicit basis but also explicit matrices for the permutations of S_n or at least for a set of generating permutations (usually one chooses the Coxeter generators $(i\ i+1),\ i = 1, \ldots, n-1$).

We will discuss this question when we will deal in a more systematic way with standard tableaux.