9

Tensor Symmetry

1 Symmetry in Tensor Spaces

With all the preliminary work done this will now be a short section; it serves as an introduction to the *first fundamental theorem* of invariant theory, according to the terminology of H.Weyl.

1.1 Intertwiners and Invariants

We have seen in Chapter 1, §2.4 that, given two actions of a group G, an equivariant map is just an invariant under the action of G on maps.

For linear representations the action of G preserves the space of linear maps, so if U, V are two linear representations,

$$\hom_G(U, V) = \hom(U, V)^G.$$

For finite-dimensional representations, we have identified, in a G-equivariant way,

$$\hom(U, V) = U^* \otimes V = (U \otimes V^*)^*.$$

This last space is the space of bilinear functions on $U \times V^*$.

Explicitly, a homomorphism $f : U \to V$ corresponds to the bilinear form

$$\langle f | u \otimes \varphi \rangle = \langle \varphi | f(u) \rangle.$$

We thus have a correspondence between intertwiners and invariants.

We will find it particularly useful, according to the Aronhold method, to use this correspondence when the representations are tensor powers $U = A^{\otimes m}$; $V = B^{\otimes p}$ and $\hom(U, V) = A^{*\otimes m} \otimes B^{\otimes p}$.

In particular when $A = B$; $m = p$ we have

$$\text{(1.1.1)} \qquad \operatorname{End}(A^{\otimes m}) = \operatorname{End}(A)^{\otimes m} = A^{*\otimes m} \otimes A^{\otimes m} = (A^{*\otimes m} \otimes A^{\otimes m})^*.$$

Thus in this case we have

Proposition. *We can identify, at least as vector spaces, the G-endomorphisms of $A^{\otimes m}$ with the multilinear invariant functions on m variables in A and m variables in A^*.*

Let V be an m-dimensional space. On the tensor space $V^{\otimes n}$ we consider two group actions, one given by the linear group $GL(V)$ by the formula

$$g(v_1 \otimes v_2 \otimes \cdots \otimes v_n) := gv_1 \otimes gv_2 \otimes \cdots \otimes gv_n, \tag{1.1.2}$$

and the other by the symmetric group S_n given by

$$\sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = v_{\sigma^{-1}1} \otimes v_{\sigma^{-1}2} \otimes \cdots \otimes v_{\sigma^{-1}n}. \tag{1.1.3}$$

We will refer to this second action as the *symmetry action* on tensors. By the definition it is clear that these two actions commute.

Before we make any further analysis of these actions, recall that in Chapter 5, §2.3 we studied symmetric tensors. Let us recall the main points of that analysis. Given a vector $v \in V$ the tensor $v^n = v \otimes v \otimes v \cdots \otimes v$ is symmetric.

Fix a basis $e_1, e_2, \dots, e_m$ of V. The basis elements $e_{i_1} \otimes e_{i_2} \cdots \otimes e_{i_n}$ are permuted by S_n and the orbits are classified by the multiplicities $h_1, h_2, \dots, h_m$ with which the elements $e_1, e_2, \dots, e_m$ appear in the term $e_{i_1} \otimes e_{i_2} \dots \otimes e_{i_n}$.

The sum of the elements of the corresponding orbit are a basis of the symmetric tensors. The multiplicities $h_1, h_2, \dots, h_m$ are nonnegative integers, subject only to $\sum_i h_i = n$.

If $\underline{h} := h_1, h_2, \dots, h_m$ is such a sequence, we denote by $e_{\underline{h}}$ the sum of elements in the corresponding orbit. The image of the symmetric tensor $e_{\underline{h}}$ in the symmetric algebra is

$$\binom{n}{h_1\, h_2 \cdots h_m} e_1^{h_1} e_2^{h_2} \cdots e_m^{h_m}.$$

If $v = \sum_k x_k e_k$, we have

$$v^n = \sum_{h_1+h_2+\cdots+h_m=n} x_1^{h_1} x_2^{h_2} \cdots x_m^{h_m} e_{\underline{h}}.$$

A linear function ϕ on the space of symmetric tensors is defined by $\langle \phi | e_{\underline{h}} \rangle = a_{\underline{h}}$ and computing on the tensor v^n gives

$$\left\langle \phi \middle| \left(\sum_k x_k e_k \right)^n \right\rangle = \sum_{h_1+h_2+\cdots+h_m=n} x_1^{h_1} x_2^{h_2} \dots x_m^{h_m} a_{\underline{h}}.$$

This formula shows that the dual of the space of symmetric tensors of degree n is identified with the space of homogeneous polynomials of degree n.

Let us recall that a subset $X \subset V$ is Zariski dense if the only polynomial vanishing on X is 0. A typical example that we will use is: when the base field is infinite, the set of vectors where a given polynomial is nonzero (easy to verify).

Lemma. *(i) The elements $v^{\otimes n}$, $v \in V$, span the space of symmetric tensors.*

(ii) More generally, given a Zariski dense set $X \subset V$, the elements $v^{\otimes n}$, $v \in X$, span the space of symmetric tensors.

Proof. Given a linear form on the space of symmetric tensors we restrict it to the tensors $v^{\otimes n}$, $v \in X$, obtaining the values of a homogeneous polynomial on X. Since X is Zariski dense this polynomial vanishes if and only if the form is 0, hence the tensors $v^{\otimes n}$, $v \in X$ span the space of symmetric tensors. $\square$

Of course the use of the word symmetric is coherent with the general idea of *invariant under the symmetric group.*

1.2 Schur–Weyl Duality

We want to apply the general theory of semisimple algebras to the two group actions introduced in the previous section. It is convenient to introduce the two algebras of linear operators spanned by these actions; thus

(1) We call A the span of the operators induced by $GL(V)$ in $\mathrm{End}(V^{\otimes n})$.
(2) We call B the span of the operators induced by S_n in $\mathrm{End}(V^{\otimes n})$.

Our aim is to prove:

Proposition. *If V is a finite-dimensional vector space over an infinite field of any characteristic, then B is the centralizer of A.*

Proof. We start by identifying

$$\mathrm{End}(V^{\otimes n}) = \mathrm{End}(V)^{\otimes n}.$$

The decomposable tensor $A_1 \otimes A_2 \otimes \cdots \otimes A_n$ corresponds to the operator:

$$A_1 \otimes A_2 \otimes \cdots \otimes A_n(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = A_1v_1 \otimes A_2v_2 \otimes \cdots \otimes A_nv_n.$$

Thus, if $g \in GL(V)$, the corresponding operator in $V^{\otimes n}$ is $g \otimes g \otimes \cdots \otimes g$. From Lemma 1.1 it follows that the algebra A coincides with the symmetric tensors in $\mathrm{End}(V)^{\otimes n}$ since $GL(V)$ is Zariski dense.

It is thus sufficient to show that for an operator in $\mathrm{End}(V)^{\otimes n}$, the condition of commuting with S_n is equivalent to being symmetric as a tensor.

It is sufficient to prove that the conjugation action of the symmetric group on $\mathrm{End}(V^{\otimes n})$ coincides with the symmetry action on $\mathrm{End}(V)^{\otimes n}$.

It is enough to verify the previous statement on decomposable tensors since they span the tensor space; thus we compute:

$$\begin{aligned}
&\sigma A_1 \otimes A_2 \otimes \cdots \otimes A_n\sigma^{-1}(v_1 \otimes v_2 \otimes \cdots \otimes v_n) \\
&= \sigma A_1 \otimes A_2 \otimes \cdots \otimes A_n(v_{\sigma 1} \otimes v_{\sigma 2} \otimes \cdots \otimes v_{\sigma n}) \\
&= \sigma(A_1v_{\sigma 1} \otimes A_2v_{\sigma 2} \otimes \cdots \otimes A_nv_{\sigma n}) = A_{\sigma^{-1}1}v_1 \otimes A_{\sigma^{-1}2}v_2 \ldots A_{\sigma^{-1}n}v_n \\
&= (A_{\sigma^{-1}1} \otimes A_{\sigma^{-1}2} \ldots A_{\sigma^{-1}n})(v_1 \otimes v_2 \otimes \cdots \otimes v_n).
\end{aligned}$$

This computation shows that the conjugation action is in fact the symmetry action and finishes the proof. □

We now draw a main conclusion:

Theorem. *If the characteristic of F is 0, the algebras A, B are semisimple and each is the centralizer of the other.*

Proof. Since B is the span of the operators of a finite group it is semisimple by Maschke's theorem (Chapter 6, §1.5); therefore, by the Double Centralizer Theorem (Chapter 6, Theorem 2.5) all statements follow from the previous theorem which states that A is the centralizer of B. □

Remark. If the characteristic of the field F is not 0, in general the algebras A, B are not semisimple. Nevertheless it is still true (at least if F is infinite or big enough) that each is the centralizer of the other (cf. Chapter 13, Theorem 7.1).

1.3 Invariants of Vectors

We formulate Theorem 1.2 in a different language.

Given two vector spaces V, W we have identified $\hom(V, W)$ with $W \otimes V^*$ and with the space of bilinear functions on $W^* \times V$ by the formulas ($A \in \hom(V, W)$, $\alpha \in W^*$, $v \in V$):

$$\langle \alpha | Av \rangle. \tag{1.3.1}$$

In case V, W are linear representations of a group G, A is in $\hom_G(V, W)$ if and only if the bilinear function $\langle \alpha | Av \rangle$ is G-invariant.

In particular we see that for a linear representation V the space of G-linear endomorphisms of $V^{\otimes n}$ is identified with the space of multilinear functions of an *n covector*[66] *and n vector variables* $f(\alpha_1, \alpha_2, \dots, \alpha_n, v_1, v_2, \dots, v_n)$ which are G-invariant.

Let us see the meaning of this for $G = GL(V)$, V an m-dimensional vector space. In this case we know that the space of G-endomorphisms of $V^{\otimes n}$ is spanned by the symmetric group S_n. We want to see which invariant function f_σ corresponds to a permutation σ. By the formula 1.3.1 evaluated on decomposable tensors we get

$$\begin{aligned} f_\sigma(\alpha_1, \alpha_2, \dots, \alpha_n, v_1, v_2, \dots, v_n) &= \langle \alpha_1 \otimes \alpha_2 \otimes \cdots \otimes \alpha_n | \sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_n) \rangle \\ &= \langle \alpha_1 \otimes \alpha_2 \otimes \cdots \otimes \alpha_n | v_{\sigma^{-1}1} \\ &\qquad \otimes v_{\sigma^{-1}2} \otimes \cdots \otimes v_{\sigma^{-1}n} \rangle \\ &= \prod_{i=1}^n \langle \alpha_i | v_{\sigma^{-1}i} \rangle = \prod_{i=1}^n \langle \alpha_{\sigma i} | v_i \rangle. \end{aligned}$$

We can thus deduce:

---
[66] Covector means linear form.

Proposition. *The space of $GL(V)$ invariant multilinear functions of n covector and n vector variables is spanned by the functions*

$$f_\sigma(\alpha_1, \alpha_2, \ldots, \alpha_n, v_1, v_2, \ldots, v_n) := \prod_{i=1}^{n} \langle \alpha_{\sigma i} | v_i \rangle. \tag{1.3.2}$$

1.4 First Fundamental Theorem for the Linear Group (FFT)

Up to now we have made no claim on the linear dependence or independence of the operators in S_n or of the corresponding functions f_σ. This will be analyzed in Chapter 13, §8.

We want to drop now the restriction that the invariants be multilinear.

Take the space $(V^*)^p \times V^q$ of p covector and q vector variables as the representation of $GL(V)(\dim(V) = m)$. A typical element is a sequence

$$(\alpha_1, \alpha_2, \ldots, \alpha_p, v_1, v_2, \ldots, v_q), \ \alpha_i \in V^*, \ v_j \in V.$$

On this space consider the pq polynomial functions $\langle \alpha_i | v_j \rangle$ which are clearly $GL(V)$ invariant. We prove:[67]

Theorem (FFT First fundamental theorem for the linear group). The ring of polynomial functions on $V^{*p} \times V^q$ that are $GL(V)$-invariant is generated by the functions $\langle \alpha_i | v_j \rangle$.

Before starting to prove this theorem we want to make some remarks about its meaning.

Fix a basis of V and its dual basis in V^*. With these bases, V is identified with the set of m-dimensional column vectors and V^* with the space of m-dimensional row vectors.

The group $GL(V)$ is then identified with the group $Gl(m, \mathbb{C})$ of $m \times m$ invertible matrices. Its action on column vectors is the product Av, $A \in Gl(m, \mathbb{C})$, $v \in V$, while on the row vectors the action is by αA^{-1}.

The invariant function $\langle \alpha_i | v_j \rangle$ is then identified with the product of the row vector α_i with the column vector v_j. In other words identify the space $(V^*)^p$ of p-tuples of row vectors with the space of $p \times m$ matrices (in which the p rows are the coordinates of the covectors) and (V^q) with the space of $m \times q$ matrices. Thus our representation is identified with the space of pairs:

$$(X, Y) | X \in M_{p,m}, \ Y \in M_{m,q}.$$

The action of the matrix group is by

$$A(X, Y) := (XA^{-1}, AY).$$

[67] At this moment we are in characteristic 0, but in Chapter 13 we will generalize our results to all characteristics.

Consider the multiplication map:

(1.4.1) $$f : M_{p,m} \times M_{m,q} \to M_{p,q}, \ f(X, Y) := XY.$$

The entries of the matrix XY are the basic invariants $\langle \alpha_i | v_j \rangle$; thus the theorem can also be formulated as:

Theorem. *The ring of polynomial functions on $M_{p,m} \times M_{m,q}$ that are $Gl(m, \mathbb{C})$-invariant is given by the polynomial functions on $M_{p,q}$ composed with the map f.*

Proof. We will now prove the theorem in its first form by the Aronhold method.

Let $g(\alpha_1, \alpha_2, \dots, \alpha_p, v_1, v_2, \dots, v_q)$ be a polynomial invariant. Without loss of generality we may assume that it is homogeneous in each of its variables; then we polarize it with respect to each of its variables and obtain a new multilinear invariant of the form $\overline{g}(\alpha_1, \alpha_2, \dots, \alpha_N, v_1, v_2, \dots, v_M)$ where N and M are the total degrees of g in the α, v respectively.

First we show that $N = M$. In fact, among the elements of the linear group we have scalar matrices. Given a scalar λ, by definition it transforms v to λv and α to $\lambda^{-1}\alpha$ and thus, by the multilinearity hypothesis, it transforms the function $\overline{g}$ in $\lambda^{M-N}\overline{g}$. The invariance condition implies $M = N$.

We can now apply Proposition 1.3 and deduce that $\overline{g}$ is a linear combination of functions of the form $\prod_{i=1}^{N} \langle \alpha_{\sigma i} | v_i \rangle$.

We now apply restitution to compute g from $\overline{g}$. It is clear that g has the desired form. □

The study of the relations among invariants will be the topic of the *Second Fundamental Theorem, SFT.* Here we only remark that by elementary linear algebra, the multiplication map f has, as image, the subvariety $D_{p,q}(m)$ of $p \times q$ matrices of rank $\leq m$. This is the whole space if $m \geq \min(p, q)$; otherwise, it is a proper subvariety, called a *determinantal variety* defined, at least set theoretically, by the vanishing of the determinants of the $(m+1) \times (m+1)$ minors of the matrix of coordinate functions x_{ij} on $M_{p,q}$.

The Second Fundamental Theorem will prove that these determinants generate a prime ideal which is thus the full ideal of relations among the invariants $\langle \alpha_i | v_j \rangle$.

In fact it is even better to introduce a formal language. Suppose that V is an affine algebraic variety with the action of an algebraic group G. Suppose that $p : V \to W$ is a morphism of affine varieties, inducing the comorphism $p^* : k[W] \to k[V]$.

Definition. We say that $p : V \to W$ is a quotient under G and write $W := V//G$ if p^* is an isomorphism from $k[W]$ to the ring of invariants $k[V]^G$.

Thus the FFT says in this geometric language that the determinantal variety $D_{p,q}(m)$ is the quotient under $GL(m, \mathbb{C})$ of $(V^*)^{\oplus p} \oplus V^{\oplus q}$.

2 Young Symmetrizers

2.1 Young Diagrams

We now discuss the symmetric group. The theory of cycles (cf. Chapter 1, §2.2) implies that the conjugacy classes of S_n are in one-to-one correspondence with the isomorphism classes of $\mathbb{Z}$ actions on $[1, 2, \dots, n]$ and these are parameterized by partitions of n.

As in Chapter 1, we express that $\mu := k_1, k_2, \dots, k_n$ is a partition of n by $\mu \vdash n$.

We shall denote by $C(\mu)$ the conjugacy class in S_n formed by the permutations decomposed in cycles of length $k_1, k_2, \dots, k_n$, hence $S_n = \sqcup_{\mu \vdash n} C(\mu)$.

Consider the group algebra $R := \mathbb{Q}[S_n]$ of the symmetric group. We wish to work over $\mathbb{Q}$ since the theory has really this more arithmetic flavor. We will (implicitly) exhibit a decomposition as a direct sum of matrix algebras over $\mathbb{Q}$:[68]

$$R = \mathbb{Q}[S_n] := \bigoplus_{\mu \vdash n} M_{d(\mu)}(\mathbb{Q}). \tag{2.1.1}$$

The numbers $d(\mu)$ will be computed in several ways from the partition μ.

Recall, from the theory of group characters, that we know at least that

$$R_{\mathbb{C}} := \mathbb{C}[S_n] := \sum_i M_{n_i}(\mathbb{C}),$$

where the number of summands is equal to the number of conjugacy classes, hence the number of partitions of n. For every partition $\lambda \vdash n$ we will construct a primitive idempotent e_λ in R so that $R = \bigoplus_{\lambda \vdash n} \mathrm{R}e_\lambda R$ and $\dim_{\mathbb{Q}} e_\lambda \mathrm{R}e_\lambda = 1$. In this way the left ideals $\mathrm{R}e_\lambda$ will exhaust all irreducible representations. The description of 2.1.1 then follows from Chapter 6, Theorem 3.1 (5).

In fact we will construct idempotents e_λ, $\lambda \vdash n$ so that $\dim_{\mathbb{Q}} e_\lambda \mathrm{R}e_\lambda = 1$ and $e_\lambda \mathrm{R}e_\mu = 0$ if $\lambda \neq \mu$. By the previous results we have that R contains a direct summand of the form $\bigoplus_{\mu \vdash n} M_{n(\mu)}(\mathbb{Q})$, or $R = \bigoplus_{\mu \vdash n} M_{n(\mu)}(\mathbb{Q}) \oplus R'$. We claim that $R' = 0$; otherwise, once we complexify, the algebra $R_{\mathbb{C}} = \bigoplus_{\mu \vdash n} M_{n(\mu)}(\mathbb{C}) \oplus R'_{\mathbb{C}}$ would contain more simple summands than the number of partitions of n, a contradiction.

For a partition $\lambda \vdash n$ let B be the corresponding Young diagram, formed by n boxes which are partitioned in rows or in columns. The intersection between a row and a column is either empty or it reduces to a single box.

In a more formal language consider the set $\mathbb{N}^+ \times \mathbb{N}^+$ of pairs of positive integers. For a pair $(i, j) \in \mathbb{N} \times \mathbb{N}$ set $C_{i,j} := \{(h,k) | 1 \leq h \leq i,\ 1 \leq k \leq j\}$ (this is a *rectangle*).

These rectangular sets have the following simple but useful properties:

(1) $C_{i,j} \subset C_{h,k}$ if and only if $(i, j) \in C_{h,k}$.
(2) If a rectangle is contained in the union of rectangles, then it is contained in one of them.

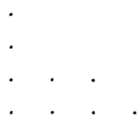
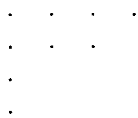
[68] So in this case all the division algebras coincide with $\mathbb{Q}$.

Definition. A Young diagram is a subset of $\mathbb{N}^+ \times \mathbb{N}^+$ consisting of a finite union of rectangles $C_{i,j}$.

In the literature this particular way of representing a Young diagram is also called a *Ferrer diagram*. Sometimes we will use this expression when we want to stress the formal point of view.

There are two conventional ways to display a Young diagram (sometimes referred to as the French and the English way) either as points in the first quadrant or in the fourth:

Example. The partition 4311:

French  English 

Any Young diagram can be written uniquely as a union of sets $C_{i,j}$ so that no rectangle in this union can be removed. The corresponding elements (i, j) will be called the *vertices* of the diagram.

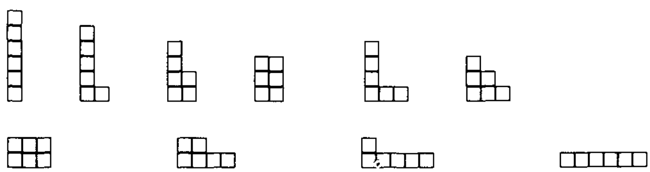
Given a Young diagram D (in French form) the set $C_i := \{(i, j) \in D\}$, i fixed, will be called the i^{th} *column*, the set $R_j := \{(i, j) \in D\}$, j fixed, will be called the j^{th} *row*.

The lengths $k_1, k_2, k_3, \dots$ of the rows are a decreasing sequence of numbers which completely determine the diagrams. Thus we can identify the set of diagrams with n boxes with the set of partitions of n; this partition is called *the row shape* of the diagram.

Of course we could also have used the column lengths and the so-called *dual* partition which is *the column shape* of the diagram.

The map that to a partition associates its dual is an involutory map which geometrically can be visualized as flipping the Ferrer diagram around its diagonal.

The elements (h, k) in a diagram will be called *boxes* and displayed more pictorially as (e.g., diagrams with 6 boxes, French display):



2.2 Symmetrizers

Definition 1. A bijective map from the set of boxes to the interval $(1, 2, 3, \dots, n-1, n)$ is called a *tableau*. It can be thought as a *filling* of the diagram with numbers. The given partition λ is called the *shape* of the tableau.

Example. The partition 4311:[69]

$$\text{French}\quad \begin{matrix} 3 & & & \\ 1 & & & \\ 5 & 2 & 7 & \\ 4 & 9 & 6 & 8 \end{matrix}\quad , \quad \begin{matrix} 7 & & & \\ 4 & & & \\ 3 & 6 & 8 & \\ 1 & 2 & 5 & 9 \end{matrix}$$

The symmetric group S_n acts on the tableaux by composition:

$$\sigma T : B \xrightarrow{T} (1, 2, 3, \dots, n-1, n) \xrightarrow{\sigma} (1, 2, 3, \dots, n-1, n).$$

A tableau induces two partitions on $(1, 2, 3, \dots, n-1, n)$:

The *row partition* is defined by: i, j are in the same part if they appear in the same row of T. The *column partition* is defined similarly.

To a partition π of $(1, 2, 3, \dots, n-1, n)$[70] one associates the subgroup S_π of the symmetric group of permutations which preserve the partition. It is isomorphic to the product of the symmetric groups of all the parts of the partition. To a tableau T one associates two subgroups $\mathrm{R}_T, \mathrm{C}_T$ of S_n.

(1) R_T is the group preserving the row partition.
(2) C_T is the subgroup preserving the column partition.

It is clear that $\mathrm{R}_T \cap \mathrm{C}_T = 1$ since each box is an intersection of a row and a column.

Notice that if $s \in S_n$, the row and column partitions associated to sT are obtained by applying s to the corresponding partitions of T. Thus

(2.2.1) $$\mathrm{R}_{sT} = s\mathrm{R}_T s^{-1}, \quad \mathrm{C}_{sT} = s\mathrm{C}_T s^{-1}.$$

We define two elements in $R = \mathbb{Q}[S_n]$:

(2.2.2) $$s_T = \sum_{\sigma \in \mathrm{R}_T} \sigma \qquad \text{the *symmetrizer on the rows*}$$

$$a_T = \sum_{\sigma \in \mathrm{C}_T} \epsilon_\sigma \sigma \qquad \text{the *antisymmetrizer on the columns.*}$$

Recall that ϵ_σ denotes the sign of the permutation. The two identities are clear:

$$s_T^2 = \prod_i h_i! \, s_T, \qquad a_T^2 = \prod_i k_i! \, a_T$$

where the h_i are the lengths of the rows and k_i are the lengths of the columns.

It is better to get acquainted with these two elements from which we will build our main object of interest.

[69] The reader will notice the peculiar properties of the right tableau, which we will encounter over and over in the future.

[70] There is an ambiguity in the use of the word *partition*. A partition of n is just a non-increasing sequence of numbers adding to n, while a partition of a set is in fact a decomposition of the set into disjoint parts.

$$(2.2.3) \qquad ps_T = s_T = s_T p, \ \forall p \in R_T; \qquad qa_T = a_T q = \epsilon_q a_T, \ \forall q \in C_T.$$

Conversely $ps_T = s_T$ or $s_p = s_T$ implies $p \in R_T$. Similarly $qa_T = \epsilon_q a_T$ or $a_T q = \epsilon_q a_T$ implies $q \in C_T$. It is then an easy exercise to check the following.

Proposition. *The left ideal $\mathbb{Q}[S_n]s_T$ has as a basis the elements gs_T as g runs over a set of representatives of the cosets gR_T and it equals, as a representation, the permutation representation on such cosets.*

The left ideal $\mathbb{Q}[S_n]a_T$ has as a basis the elements ga_T as g runs over a set of representatives of the cosets gC_T and it equals, as a representation, the representation induced to S_n by the sign representation of C_T.

Now the remarkable fact comes. Consider the product

$$(2.2.4) \qquad c_T := s_T a_T = \sum_{p \in R_T, \ q \in C_T} \epsilon_q pq.$$

We will show that:

Theorem. *There exists a positive integer $p(T)$ such that the element $e_T := \frac{c_T}{p(T)}$ is a primitive idempotent.*

Definition 2. The idempotent $e_T := \frac{c_T}{p(T)}$ is called the *Young symmetrizer* relative to the given tableau.

Remark.

$$(2.2.5) \qquad c_{sT} = s c_T s^{-1}.$$

We thus have for a given $\lambda \vdash n$ several conjugate idempotents, which we will show to be primitive, associated to tableaux of row shape λ. Each will generate an irreducible module associated to λ which will be denoted by M_λ.

For the moment, let us remark that from 2.2.5 it follows that the integer $p(T)$ depends only on the shape λ of T, and thus we will denote it by $p(T) = p(\lambda)$.

2.3 The Main Lemma

The main property of the element c_T which we will explore is the following, which is clear from its definition and 2.2.3:

$$(2.3.1) \qquad pc_T = c_T, \ \forall p \in R_T; \ c_T q = \epsilon_q c_T, \ \forall q \in C_T.$$

We need a fundamental combinatorial lemma. Consider the partitions of n as decreasing sequences of integers (including 0) and order them lexicographically.[71]

For example, the partitions of 6 in increasing lexicographic order:

$$111111, 21111, 2211, 222, 3111, 321, 411, 42, 51, 6.$$

[71] We often drop 0 in the display.

Lemma. *Let S and T be two tableaux of row shapes:*

$$\lambda = h_1 \geq h_2 \geq \ldots \geq h_n, \ \mu = k_1 \geq k_2 \geq \ldots \geq k_n$$

with $\lambda \geq \mu$. Then one and only one of the two following possibilities holds:

(i) Two numbers i, j appear in the same row in S and in the same column in T.
(ii) $\lambda = \mu$ and $pS = qT$ where $p \in \mathrm{R}_S$, $q \in \mathrm{C}_T$.

Proof. We consider the first row r_1 of S. Since $h_1 \geq k_1$, by the pigeonhole principle either there are two numbers in r_1 which are in the same column in T or $h_1 = k_1$ and we can act on T with a permutation s in C_T so that S and sT have the first row filled with the same elements (possibly in a different order).

Observe that two numbers appear in the same column in T if and only if they appear in the same column in sT or $\mathrm{C}_T = \mathrm{C}_{sT}$.

We now remove the first row in both S and T and proceed as before. At the end we are either in case (i) or $\lambda = \mu$ and we have found a permutation $q \in \mathrm{C}_T$ such that S and qT have each row filled with the same elements.

In this case we can find a permutation $p \in \mathrm{R}_S$ such that $pS = qT$.

In order to complete our claim we need to show that these two cases are mutually exclusive. Thus we have to remark that if $pS = qT$ as before, then case (i) is not verified. In fact two elements are in the same row in S if and only if they are in the same row in pS, while they appear in the same column in T if and only if they appear in the same column in qT. Since $pS = qT$ two elements in the same row of pS are in different columns of qT. □

Corollary. *(i) Given $\lambda > \mu$ partitions, S and T tableaux of row shapes λ, μ respectively, and s any permutation, there exists a transposition $u \in \mathrm{R}_S$ and a transposition $v \in \mathrm{C}_T$ such that $us = sv$.*

(ii) If, for a tableau T, s is a permutation not in $R_T C_T$, then there exists a transposition $u \in R_T$ and a transposition $v \in C_T$ such that $us = sv$.

Proof. (i) From the previous lemma there are two numbers i and j in the same row for S and in the same column for sT. If $u = (i, j)$ is the corresponding transposition, we have $u \in R_S$, $u \in C_{sT}$. We set $v := s^{-1}us$ and we have $v \in s^{-1}C_{sT}s = C_T$ by 2.2.1. By definition $sv = uv$.

(ii) The proof is similar. We consider the tableau T, construct $s^{-1}T$, and apply the lemma to $s^{-1}T, T$.

If there exists a $p' \in R_{s^{-1}T}$, $q \in C_T$ with $p's^{-1}T = qT$, since $p' = s^{-1}ps$, $p \in R_T$, we would have that $s^{-1}p = q$, $s = pq^{-1}$ against the hypothesis. Hence there is a transposition $v \in C_T$ and $v \in R_{s^{-1}T}$ or $v = s^{-1}us$, $u \in R_T$, as required. □

2.4 Young Symmetrizers 2

We now draw the conclusions relative to Young symmetrizers.

Proposition. *(i) Let S and T be two tableaux of row shapes $\lambda > \mu$.*
If an element a in the group algebra is such that

$$pa = a, \ \forall p \in R_S, \ and\, aq = \epsilon_q a, \ \forall q \in C_T,$$

then $a = 0$.
(ii) Given a tableau T and an element a in the group algebra such that

$$pa = a, \ \forall p \in R_T, \ and\, aq = \epsilon_q a, \ \forall q \in C_T,$$

then a is a scalar multiple of the element c_T.

Proof. (i) Let us write $a = \sum_{s \in S_n} a(s)s$; for any given s we can find u, v as in the previous lemma.

By hypothesis $ua = a, av = -a$. Then $a(s) = a(us) = a(sv) = -a(s) = 0$ and thus $a = 0$.

(ii) Using the same argument as above, we can say that if $s \notin \mathrm{R}_T\mathrm{C}_T$, then $a(s) = 0$. Instead, let $s = pq, \ p \in \mathrm{R}_T, q \in \mathrm{C}_T$. Then $a(pq) = \epsilon_q a(1)$, hence $a = a(1)c_T$. □

Before we conclude let us recall some simple facts about algebras and group algebras.

If R is a finite-dimensional algebra over a field F, we can consider any element $r \in R$ as a linear operator on R (as vector space) by right or left action. Let us define $\mathrm{tr}(r)$ to be the trace of the operator $x \mapsto xr$.[72] Clearly $\mathrm{tr}(1) = \dim_F R$. For a group algebra $F[G]$ of a finite group G, an element $g \in G, \ g \neq 1$, gives rise to a permutation $x \to xg, \ x \in G$ of the basis elements without fixed points. Hence, $\mathrm{tr}(1) = |G|$, $\mathrm{tr}(g) = 0$ if $g \neq 0$.

We are now ready to conclude. For $R = \mathbb{Q}[S_n]$ the theorems that we aim at are:

Theorem 1.

(i) $c_T R c_T = c_T R a_T = s_T R c_T = s_T R a_T = \mathbb{Q} c_T$.
(ii) $c_T^2 = p(\lambda) c_T$ *with* $p(\lambda) \neq 0$ *a positive integer.*
(iii) $\dim_{\mathbb{Q}} R c_T = \frac{n!}{p(\lambda)}$.
(iv) If U, V are two tableaux of shapes $\lambda > \mu$, then $s_U R a_V = a_V R s_U = 0$.
(v) If U, V are tableaux of different shapes λ, μ, we have $c_U R c_V = 0 = s_U R a_V$.

Proof. (i) We cannot have $c_T R c_T = 0$ since R is semisimple. Hence it is enough to prove $s_T R a_T = \mathbb{Q} c_T$. We apply the previous proposition and get that every element of $s_T R a_T$ satisfies (ii) of that proposition, hence $s_T R a_T = \mathbb{Q} c_T$.

(ii) In particular we have $c_T^2 = p(\lambda) c_T$. Now compute the trace of c_T for the right regular representation. From the previous discussion we have $\mathrm{tr}(c_T) = n!$, hence

[72] One can prove in fact that the operator $x \to rx$ has the same trace.

$c_T^2 \neq 0$. Since $c_T^2 = p(\lambda)c_T$ we have that $p(\lambda) \neq 0$. Since $p(\lambda)$ is the coefficient of 1 in the product c_T^2, it is clear that it is an integer.

(iii) $e_T := \frac{c_T}{p(\lambda)}$ is idempotent and $\frac{n!}{p(\lambda)} = \frac{\mathrm{tr}(c_T)}{p(\lambda)} = \mathrm{tr}(e_T)$. The trace of an idempotent operator is the dimension of its image. In our case $Re_T = Rc_T$, hence $\frac{n!}{p(\lambda)} = \dim_{\mathbb{Q}} Rc_T$. In particular this shows that $p(\lambda)$ is positive.

(iv) If $\lambda > \mu$ we have, by part (i), $s_U Ra_V = 0$.

(v) If $\lambda > \mu$ we have, by (iv), $c_U Rc_V = s_U a_U Rs_V a_V \subset s_U Ra_V = 0$. Otherwise $c_V Rc_U = 0$, which, since R has no nilpotent ideals, implies $c_U Rc_V = 0$ (Chapter 6, §3.1).

From the general discussion performed in 2.1 we finally obtain

Theorem 2. *(i) The elements $e_T := \frac{c_T}{p(\lambda)}$ are primitive idempotents in $R = \mathbb{Q}[S_n]$.*

(ii) The left ideals Re_T give all the irreducible representations of S_n explicitly indexed by partitions.

(iii) These representations are defined over $\mathbb{Q}$.

We will indicate by M_λ the irreducible representation associated to a (row) partition λ.

Remark. The Young symmetrizer a priori does not depend only on the partition λ but also on the labeling of the diagram. Two different labelings give rise to conjugate Young symmetrizers which therefore correspond to isomorphic irreducible representations.

We could have used, instead of the product $s_T a_T$, the product $a_T s_T$ in reverse order. We claim that also in this way we obtain a primitive idempotent $\frac{a_T s_T}{p(\lambda)}$, relative to the same irreducible representation.

The same proof could be applied, but we can also argue by applying the antiautomorphism $a \to \overline{a}$ of the group algebra which sends a permutation σ to σ^{-1}. Clearly,

$$\overline{a}_T = a_T, \ \overline{s}_T = s_T, \ \overline{s_T a_T} = a_T s_T.$$

Thus $\frac{1}{p(T)} a_T s_T = \overline{e}_T$ is a primitive idempotent.

Since clearly $c_T a_T s_T = s_T a_T a_T s_T$ is nonzero (a_T^2 is a nonzero multiple of a_T and so $(c_T a_T s_T) a_T$ is a nonzero multiple of c_T^2) we get that e_T and $\overline{e}_T$ are primitive idempotents relative to the same irreducible representation, and the claim is proved.

We will need two more remarks in the computation of the characters of the symmetric group.

Consider the two left ideals Rs_T, Ra_T. We have given a first description of their structure as representations in §2.2. They contain respectively $a_T Rs_T$, $s_T Ra_T$ which are both 1 dimensional. Thus we have

Lemma. *M_λ appears in its isotypic component in Rs_T (resp. Ra_T) with multiplicity 1. If M_μ appears in Rs_T, then $\mu \leq \lambda$, and if it appears in Ra_T, then $\mu \geq \lambda$.*[73]

[73] We shall prove a more precise theorem later.

Proof. To see the multiplicity with which M_μ appears in a representation V it suffices to compute the dimension of $c_T V$ or of $\overline{c}_T V$ where T is a tableau of shape μ. Therefore the statement follows from the previous results. □

In particular we see that the only irreducible representation which appears in both Rs_T, Ra_T is M_λ.

The reader should apply to the idempotents that we have discussed the following fact:

Exercise. Given two idempotents e, f in a ring R we can identify

$$\hom_R(Re, Rf) = eRf.$$

2.5 Duality

There are several deeper results on the representation theory of the symmetric group which we will describe.

A first remark is about an obvious duality between diagrams. Given a tableau T relative to a partition λ, we can exchange its rows and columns obtaining a new tableau $\tilde{T}$ relative to the partition $\tilde{\lambda}$, which in general is different from λ. It is thus natural to ask in which way the two representations are tied.

Let $\mathbb{Q}(\epsilon)$ denote the sign representation.

Proposition. $M_{\tilde{\lambda}} = M_\lambda \otimes \mathbb{Q}(\epsilon)$.

Proof. Consider the automorphism τ of the group algebra defined on the group elements by $\tau(\sigma) := \epsilon_\sigma \sigma$.

Clearly, given a representation ϱ, the composition $\varrho\tau$ is equal to the tensor product with the sign representation; thus, if we apply τ to a primitive idempotent associated to M_λ, we obtain a primitive idempotent for $M_{\tilde{\lambda}}$.

Let us therefore use a tableau T of shape λ and construct the symmetrizer. We have

$$\tau(c_T) = \sum_{p \in \mathrm{R}_T,\, q \in \mathrm{C}_T} \epsilon_p \tau(pq) = \Big(\sum_{p \in \mathrm{R}_T} \epsilon_p p\Big)\Big(\sum_{q \in \mathrm{C}_T} q\Big).$$

We remark now that since $\tilde{\lambda}$ is obtained from λ by exchanging rows and columns we have

$$\mathrm{R}_T = \mathrm{C}_{\tilde{T}}, \; \mathrm{C}_T = \mathrm{R}_{\tilde{T}}.$$

Thus $\tau(c_T) = a_{\tilde{T}} s_{\tilde{T}} = \overline{c}_{\tilde{T}}$, hence $\tau(e_T) = \overline{e}_{\tilde{T}}$. □

Remark. From the previous result it also follows that $p(\lambda) = p(\tilde{\lambda})$.

3 The Irreducible Representations of the Linear Group 1

3.1 Representations of the Linear Groups

We now apply the theory of symmetrizers to the linear group.

Let M be a representation of a semisimple algebra A and B its centralizer. By the structure theorem (Chapter 6) $M = \oplus N_i \otimes_{\Delta_i} P_i$ with N_i and P_i irreducible representations, respectively of A, B. If $e \in B$ is a primitive idempotent, then the subspace $eP_i \neq 0$ for a unique index i_0 and $eM = N_{i_0} \otimes eP_i \cong N_i$ is irreducible as a representation of A (associated to the irreducible representation of B relative to e).

Thus, from Theorem 2 of §2.4, to get a list of the irreducible representations of the linear group $Gl(V)$ appearing in $V^{\otimes n}$, we may apply the Young symmetrizers e_T to the tensor space and see when $e_T V^{\otimes n} \neq 0$.

Assume we have t columns of length $n_1, n_2, \dots, n_t$, and decompose the column preserving group C_T as a product $\prod_{i=1}^t S_{n_i}$ of the symmetric groups of all columns.

By definition we get $a_T = \prod a_{n_i}$, the product of the antisymmetrizers relative to the various symmetric groups of the columns.

Let us assume, for simplicity of notation, that the first n_1 indices appear in the first column in increasing order, the next n_2 indices in the second column, and so on, so that

$$V^{\otimes n} = V^{\otimes n_1} \otimes V^{\otimes n_2} \otimes \cdots \otimes V^{\otimes n_t},$$

$$a_T V^{\otimes n} = a_{n_1} V^{\otimes n_1} \otimes a_{n_2} V^{\otimes n_2} \otimes \cdots \otimes a_{n_t} V^{\otimes n_t} = \bigwedge^{n_1} V \otimes \bigwedge^{n_2} V \otimes \cdots \otimes \bigwedge^{n_t} V.$$

Therefore we have that if there is a column of length $> \dim(V)$, then $a_T V^{\otimes n} = 0$.

Otherwise, we have $n_i \leq \dim(V), \forall i$, and we prove the equivalent statement that $a_T s_T V^{\otimes n} \neq 0$. Let $e_1, e_2, \dots, e_m$ be a basis of V and use the corresponding basis of decomposable tensors for $V^{\otimes n}$; let us consider the tensor

(3.1.1)
$$U = (e_1 \otimes e_2 \otimes \cdots \otimes e_{n_1}) \otimes (e_1 \otimes e_2 \otimes \cdots \otimes e_{n_2}) \otimes \cdots \otimes (e_1 \otimes e_2 \otimes \cdots \otimes e_{n_t}).$$

This is the decomposable tensor having e_i in the positions corresponding to the indices of the i^{th} row. By construction it is symmetric with respect to the group R_T of row preserving permutations, hence $s_T U = pU$, $p = |R_T| \neq 0$.

Finally,

(3.1.2) $a_T U =$
$$(e_1 \wedge e_2 \wedge \cdots \wedge e_{n_1}) \otimes (e_1 \wedge e_2 \wedge \cdots \wedge e_{n_2}) \otimes \cdots \otimes (e_1 \wedge e_2 \wedge \cdots \wedge e_{n_t}) \neq 0.$$

Recall that the length of the first column of a partition λ (equal to the number of its rows) is called the *height* of λ and indicated by $ht(\lambda)$. We have thus proved:

Proposition. *If T is a tableau of shape λ, then $e_T V^{\otimes n} = 0$ if and only if $ht(\lambda) > \dim(V)$.*

For a tableau T of shape λ, define

(3.1.3) $$S_\lambda(V) := e_T V^{\otimes n}, \qquad \text{the Schur functor associated to } \lambda.$$

We are implicitly using the fact that for two different tableaux T and T' of the same shape we have a unique permutation σ with $\sigma(T) = T'$. Hence we have a canonical ismorphism between the two spaces $e_T V^{\otimes n}$, $e_{T'} V^{\otimes n}$.

Remark. We shall justify the word *functor* in 7.1.

As a consequence, we thus have a description of $V^{\otimes n}$ as a representation of $S_n \times GL(V)$.

Theorem.

(3.1.4) $$V^{\otimes n} = \bigoplus_{ht(\lambda)\leq \dim(V)} M_\lambda \otimes S_\lambda(V).$$

Proof. We know that the two algebras A and B, spanned by the linear and the symmetric group, are semisimple and each the centralizer of the other. By the structure theorem we thus have $V^{\otimes n} = \bigoplus_i M_i \otimes S_i$ where the M_i are the irreducible representations of S_n which appear. We have proved that the ones which appear are the M_λ, $ht(\lambda) \leq \dim(V)$ and that $S_\lambda(V)$ is the corresponding irreducible representation of the linear group.

4 Characters of the Symmetric Group

As one can easily imagine, the character theory of the symmetric and general linear group are intimately tied together. There are basically two approaches: a combinatorial approach due to Frobenius, which first computes the characters of the symmetric group and then deduces those of the linear group, and an analytic approach based on Weyl's character formula, which proceeds in the reverse order. It is instructive to see both. There is in fact also a more recent algebraic approach to Weyl's character formula which we will not discuss (cf. [Hu1]).

4.1 Character Table

Up to now we have been able to explicitly parameterize both the conjugacy classes and the irreducible representations of S_n by partitions of n. 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

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

Let us write

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

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

We need to interpret the number $n(\mu)$ in terms of the conjugacy class $C(\mu)$:

Proposition. *If $s \in C(\mu)$, then $n(\mu)$ 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 indices 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. □

The computation of the character table of S_n consists, given two partitions λ, μ, of computing the value of the character of an element of the conjugacy class $C(\mu)$ on the irreducible representation M_λ. Let us denote this value by $\chi_\lambda(\mu)$.

The final result of this analysis is expressed in compact form through symmetric functions. Recall that we denote $\psi_k(x) = \sum_{i=1}^n x_i^k$. For a partition $\mu \vdash n := k_1, k_2, \dots, k_n$, set

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

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

$$\psi_\mu(x) = \sum_\lambda c_\lambda(\mu)S_\lambda(x). \tag{4.1.4}$$

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

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

and we have

Theorem (Frobenius). *For all partitions $\lambda, \mu \vdash n$ we have*

$$\chi_\lambda(\mu) = c_\lambda(\mu). \tag{4.1.5}$$

The proof of this theorem is quite elaborate, and we divide it into five steps.

Step 1 First we transform the Cauchy formula into a new identity.
Step 2 Next we prove that the class functions c_λ are orthonormal.
Step 3 To each partition we associate a permutation character β_λ.
Step 4 We prove that the matrix expressing the functions β_λ in terms of the c_μ is triangular with 1 on the diagonal.
Step 5 We formulate the Theorem of Frobenius in a more precise way and prove it.

Step 1 In order to follow the 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 Chapter 2, §4.1. We change its right-hand side as follows. Compute

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

$$(4.1.6) \qquad = \sum_{h=1}^{\infty} \frac{\psi_h(x)\psi_h(y)}{h}.$$

Taking the exponential we get the following expression:

$$(4.1.7) \qquad \exp\left(\sum_{h=1}^{\infty} \frac{\psi_h(x)\psi_h(y)}{h}\right) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{h=1}^{\infty} \frac{\psi_h(x)\psi_h(y)}{h}\right)^k$$

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

$$(4.1.8) \qquad \times \frac{\psi_3(x)^{k_3}\psi_3(y)^{k_3}}{3^{k_3}} \cdots.$$

Then from 4.1.3 we deduce

$$(4.1.9) \qquad \sum_\mu \frac{1}{n(\mu)} \psi_\mu(x)\psi_\mu(y) = \sum_\lambda S_\lambda(x) S_\lambda(y).$$

Step 2 Consider two class functions a and b as functions on partitions. Their Hermitian product is

$$\sum_{\mu \vdash n} \frac{1}{n!} \sum_{g \in C(\mu)} a(g)\bar{b}(g) = \sum_{\mu \vdash n} \frac{1}{n!} |C(\mu)| a(\mu)\bar{b}(\mu) = \sum_{\mu \vdash n} \frac{1}{n(\mu)} a(\mu)\bar{b}(\mu).$$

Let us now substitute in the identity 4.1.9 the expression $\psi_\mu = \sum_\lambda c_\lambda(\mu) S_\lambda$, and get

$$(4.1.10) \qquad \sum_{\mu \vdash n} \frac{1}{n(\mu)} c_{\lambda_1}(\mu) c_{\lambda_2}(\mu) = \begin{cases} 0 & \text{if} \quad \lambda_1 \neq \lambda_2 \\ 1 & \text{if} \quad \lambda_1 = \lambda_2. \end{cases}$$

We thus have that the class functions c_λ are an orthonormal basis, completing Step 2.

Step 3 We consider now some permutation characters.

Take a partition $\lambda := h_1, h_2, \dots, h_k$ of n. Consider the subgroup $S_\lambda := S_{h_1} \times S_{h_2} \times \dots \times S_{h_k}$ and the permutation representation on:

(4.1.11) $$S_n/S_{h_1} \times S_{h_2} \times \cdots \times S_{h_k}.$$

We will indicate the corresponding character by β_λ.

A permutation character is given by the formula $\chi(g) = \sum_i \frac{|G(g)|}{|H(g_i)|}$ (§1.4.3 of Chapter 8). Let us apply it to the case $G/H = S_n/S_{h_1} \times S_{h_2} \times \cdots \times S_{h_k}$, 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 $S_{h_1} \times S_{h_2} \times \cdots \times S_{h_k}$ is given by k partitions $\mu_i \vdash h_i$ of the numbers $h_1, h_2, \dots, h_k$. The conjugacy class of type μ, intersected with $S_{h_1} \times S_{h_2} \times \cdots \times S_{h_k}$, gives all possible k tuples of partitions $\mu_1, \mu_2, \dots, \mu_k$ of type

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

and

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

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

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

and remark that, with the notations of 4.1.2, $b(\lambda \oplus \mu) = b(\lambda)b(\mu)$.

When we decompose $\mu = \bigoplus_{i=1}^k \mu_i$, we have $b(\mu) = \prod b(\mu_i)$.

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

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

Now

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

So we get

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

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

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

This sum is manifestly the coefficient of $x_1^{h_1}x_2^{h_2}\dots x_k^{h_k}$ in the symmetric function $\psi_\mu(x)$. In fact when we expand

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

for each factor $\psi_k(x) = \sum_{i=1}^n x_i^k$, one selects the index of the variable chosen and constructs a corresponding 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 $\prod_i \binom{p_i}{p_{i1}p_{i2}\dots p_{ik}}$ such choices and they contribute to the monomial $x_1^{h_1}x_2^{h_2}\dots x_k^{h_k}$ if and only if $\sum_i ip_{ij} = h_j$.

Step 4 If m_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

$$\psi_\mu(x) = \sum_\lambda \beta_\lambda(\mu)m_\lambda(x). \tag{4.1.12}$$

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

$$m_\lambda(x) = \sum_\mu p_{\lambda,\mu}S_\mu(x),\ S_\lambda(x) = \sum_\mu k_{\lambda,\mu}m_\mu(x). \tag{4.1.13}$$

In order to make 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 m_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 now compute the leading monomial of the Schur function S_μ:

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

We deduce that

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

This computation has the following immediate consequence:

Corollary. *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 m_μ plus a linear combination of the m_λ, $\lambda < \mu$. This proves the claim for the matrix Q. The matrix P is the inverse of Q and the claim follows. $\square$

Step 5 We can now conclude a refinement of the computation of Frobenius:

Theorem 2. *(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) $\chi_\lambda = c_\lambda$.

Proof. From the various definitions we get

$$c_\lambda = \sum_\phi p_{\phi,\lambda} b_\phi, \ \beta_\lambda = \sum_\phi k_{\phi,\lambda} c_\phi. \tag{4.1.14}$$

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.[74]

(iii) Now we prove the equality $\chi_\lambda = c_\lambda$ by decreasing induction. If $\lambda = n$ is one row, then the module M_λ is the trivial representation as well as the permutation representation on S_n/S_n.

Assume $\chi_\mu = c_\mu$ for all $\mu > \lambda$. We may use Lemma 2.4 and we know that M_λ appears in its isotypic component in $R s_T$ with multiplicity 1 and does not appear in $R s_U$ for any tableau of shape $\mu > \lambda$.

We have remarked that $R s_T$ is the permutation representation of character β_λ in which, by assumption, the representation M_λ appears for the first time (with respect to the ordering of the λ). Thus the contribution of M_λ to its character must be given by the term c_λ. $\square$

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

$$\psi_\mu(x) V(x) = \sum_\lambda c_\lambda(\mu) A_{\lambda+\varrho}(x). \tag{4.1.15}$$

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.

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

4.2 Frobenius Character

There is a nice interpretation of the theorem of Frobenius.

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 $\chi \mapsto F(\chi)$.

Lemma. *The Frobenius character can be computed by the formula*

$$F(\chi) = \frac{1}{n!}\sum_{\sigma\in S_n}\chi(\sigma)\psi_{\mu(\sigma)}(x) = \sum_{\mu\vdash n}\frac{\chi(\mu)}{n(\mu)}\psi_\mu(x). \tag{4.2.1}$$

Proof. By linearity it is enough to prove it for $\chi = \chi_\lambda$. From 4.1.4 and 4.1.10 we have

$$F(\chi_\lambda) = \sum_{\mu\vdash n}\frac{c_\lambda(\mu)}{n(\mu)}\psi_\mu(x) = \sum_{\mu\vdash n}\frac{c_\lambda(\mu)}{n(\mu)}\sum_\gamma c_\gamma(\mu)S_\gamma(x)$$
$$= \sum_\gamma\sum_\mu \frac{c_\lambda(\mu)c_\gamma(\mu)}{n(\mu)}S_\gamma(x) = S_\lambda(x). \qquad \square$$

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

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

$$F(\mathrm{Ind}_{S_n\times S_m}^{S_{n+m}}(V\otimes W)) = F(V)F(W). \tag{4.2.2}$$

Proof. Let us denote by χ the character of $\mathrm{Ind}_{S_m\times S_n}^{S_{m+n}}(V\otimes W)$. Recall the discussion of induced characters in Chapter 8. There we proved (formula 1.4.2) $\chi(g) = \sum_i \frac{|G(g)|}{|H(g_i)|}\chi_V(g_i)$. Where $|G(g)|$ is the order of the centralizer of g in G, the elements g_i run over representatives of the conjugacy classes O_i in H, decomposing the intersection of the conjugacy class of g in G with H.

In our case we deduce that $\chi(\sigma) = 0$ unless σ is conjugate to an element (a, b) of $S_n \times S_m$. In terms 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 partitions the previous formula 1.4.2 becomes

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

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

$$F(\chi) = \sum_{\nu\vdash m+n}\frac{\chi(\nu)\psi_\nu}{n(\nu)} = \sum_\nu \frac{\psi_\nu}{n(\nu)}\sum_{\nu=\lambda+\mu}\frac{n(\lambda+\mu)}{n(\lambda)n(\mu)}\chi_V(\lambda)\chi_W(\mu)$$
$$= \sum_{\lambda\vdash m,\mu\vdash n}\frac{\chi_V(\lambda)\chi_W(\mu)}{n(\lambda)n(\mu)}\psi_\lambda\psi_\mu = F(\lambda)F(\mu). \qquad \square$$

4.3 Molien's Formula

We discuss 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 Chapter 2, §1.1. With this in mind we think of the identities of §4 as identities in infinitely many variables.

First, 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:[75]

$$\chi_U(t) := \sum_i \chi_i q^i, \tag{4.3.1}$$

where χ_i is the character of the representation U_i.

Definition. The expression 4.3.1 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

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

For the graded characters we clearly have

$$\chi_{U \oplus V}(q) = \chi_U(q) + \chi_V(q), \ \chi_{U \otimes V}(q) = \chi_U(q)\chi_V(q). \tag{4.3.2}$$

Let us consider a simple example.[76]

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:*

$$\sum_{i=0}^{\infty} \operatorname{tr}(S^i(A))q^i = \frac{1}{\det(1 - qA)}. \tag{4.3.3}$$

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

$$\sum_{i=0}^{\dim U} \operatorname{tr}(\wedge^i(A))q^i = \det(1 + qA). \tag{4.3.4}$$

[75] It is now quite customary to use q as a variable since it often appears to come from computations on finite fields where $q = p^r$ or as a quantum deformation parameter.

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

Proof. For every symmetric power $S^k(U)$ the character of the operator induced by A is a polynomial in A. Thus 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

$$S(U) = S(u_1) \otimes S(u_2) \otimes \cdots \otimes S(u_n) \quad \text{and} \quad S(u_i) = \sum_{h=0}^{\infty} F u_i^h.$$

The graded character of $S(u_i)$ is $\sum_{h=0}^{\infty} \lambda_i^h q^h = \frac{1}{1-\lambda_i q}$, hence

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

Similarly, $\bigwedge U = \wedge[u_1] \otimes \wedge[u_2] \otimes \cdots \otimes \wedge[u_n]$ and $\wedge[u_i] = F \oplus F u_i$, hence

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

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

$$\det(1 - q\sigma) = \prod_{i=1}^{k} \prod_{j=1}^{m_i} \left(1 - e^{j2\pi\sqrt{-1}/m_i} q\right) = \prod_{i=1}^{k}(1 - q^{m_i}).$$

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

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

To summarize

Theorem 1. *The graded character of S_n acting on the polynomial ring is*

$$\sum_{\lambda \vdash n} \chi_\lambda S_\lambda(1, q, q^2, \dots, q^k, \dots). \tag{4.3.5}$$

Exercise. Prove this formula directly.

(*Hint.*) $\mathbb{C}[x_1, \dots, x_n] = \mathbb{C}[x]^{\otimes n} = \bigoplus_\lambda M_\lambda \otimes S_\lambda(\mathbb{C}[x])$.

We have a corollary of this formula. If $\lambda = h_1 \geq h_2 \dots \geq h_n$, 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.[77]

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 $R_{\underline{a}} := \mathbb{C}[x_1, x_2, \dots, x_n]/\langle \sigma_i - a_i \rangle$ 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.

Proof. 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 $\mathbb{C}[x_1, x_2, \dots, x_n]/\langle \sigma_i - a_i \rangle$ is the coordinate ring of the set of the $n!$ distinct points $\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(n)}$, $\sigma \in S_n$. This is clearly the regular representation.

We know that the condition for a polynomial to have distinct roots is given by the condition that the discriminant is not zero (Chapter 1). This condition defines a dense open set.

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. $\square$

It is of particular interest (combinatorial and geometric) to analyze the special case $\underline{a} = 0$ and the ring $R := \mathbb{C}[x_1, x_2, \dots, x_n]/\langle \sigma_i \rangle$ 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

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

and thus an identity of graded characters.

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

Theorem 2.

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

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

---

[77] It appears in the Springer representation, for instance, cf. [DP2].

Exercise. Prove that the Specht module has nonzero image in the quotient ring $R := \mathbb{C}[x_1, x_2, \dots, x_n]/\langle\sigma_i\rangle$.

The ring $R := \mathbb{Z}[x_1, x_2, \dots, x_n]/\langle\sigma_i\rangle$ has an interesting geometric interpretation as the *cohomology algebra* of the flag variety. This variety 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 induced by the topological action permuting the summands of the decomposition (Chapter 10, §6.5).

5 The Hook Formula

5.1 Dimension of M_λ

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

From 4.1.15 applied to the partition 1^n, corresponding to the conjugacy class of the identity, we obtain

$$\left(\sum_{i=1}^n x_i\right)^n V(x) = \sum_\lambda d(\lambda) A_{\lambda+\varrho}(x). \tag{5.1.1}$$

Write the expansion 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

$$\sum_{k_1+\cdots+k_n=n} \binom{n}{k_1\, k_2 \cdots k_n} \prod_{i=1}^n x_i^{k_i} \sum_{\sigma\in S_n} \epsilon_\sigma \prod_{i=1}^n x_i^{\sigma(n-i+1)-1}.$$

Thus a term $\epsilon_\sigma \binom{n}{k_1\, k_2 \cdots k_n} \prod_{i=1}^n x_i^{\sigma(n-i+1)-1+k_i}$ 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

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

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

Thus

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

$\bar{d}(\lambda)$ 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:

$$(5.1.2) \qquad d(\lambda) = \frac{n!}{\prod_{i=1}^{n} \ell_i!} \prod_{i<j} (\ell_i - \ell_j) = n! \prod_{j=1}^{n} \frac{\prod_{i<j} (\ell_i - \ell_j)}{\ell_j!}.$$

5.2 Hook Formula

We want to give a combinatorial interpretation of 5.1.2. Notice that, fixing j, in $\frac{\prod_{i<j}(\ell_i - \ell_j)}{\ell_j!}$ the $j - 1$ factors of the numerator cancel the corresponding factors in the denominator, leaving $\ell_j - j + 1$ factors. In all $\sum_j \ell_j - \sum_{j=1}^{n} (j - 1) = n$ factors are left. These factors can be interpreted as 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. For example, we mark the hooks of $1, 2;\ 2, 1;\ 2, 2$ in 4,3,1,1:

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

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

The Frobenius formula for the dimension $d(\lambda)$ can be reformulated in the settings of the *hook formula.*

Theorem. *Denote by $B(\lambda)$ the set of boxes of a diagram of shape λ. Then*

$$(5.2.1) \qquad d(\lambda) = \frac{n!}{\prod_{x \in B(\lambda)} h_x}, \qquad \textit{hook formula.}$$

Proof. It is enough to show that the factors in the factorial $\ell_i!$, which are not canceled by the factors of the numerator, are the hook lengths of the boxes in the i^{th} row. This will prove the formula.

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 $\ell_i - \ell_k = h_i - h_k + k - i$.

Observe that $h_{k-1} - h_k$ is the number of cases in the i^{th} row for which the hook ends vertically on the $k - 1$ row. 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 $k - i + h_i - h_{k-1}$, the previously considered numbers. □

6 Characters of the Linear Group

6.1 Tensor Character

We plan to deducethe character theory of the linear group from previous computations. For this we need to perform another character computation. Given a permutation $s \in S_n$ and a matrix $X \in GL(V)$ consider the product sX as an operator in $V^{\otimes n}$. We want to compute its trace.

Let $\mu = h_1, h_2, \ldots, h_k$ denote the cycle partition of s; introduce the obvious notation:

(6.1.1) $$\Psi_\mu(X) = \prod_i \operatorname{tr}(X^{h_i}).$$

Clearly $\Psi_\mu(X) = \psi_\mu(x)$, where by x, we denote the eigenvalues of X.

Proposition. *The trace of sX as an operator in $V^{\otimes n}$ is $\Psi_\mu(X)$.*

We shall deduce this proposition as a special case of a more general formula. Given n matrices $X_1, X_2, \ldots, X_n$ and $s \in S_n$ we will compute the trace of $s \circ X_1 \otimes X_2 \otimes \cdots \otimes X_n$ (an operator in $V^{\otimes n}$).

Decompose s into cycles $s = c_1 c_2 \ldots c_k$ and, for a cycle $c := (i_p\ i_{p-1}\ \ldots\ i_1)$, define the function of the n matrix variables $X_1, X_2, \ldots, X_n$:

(6.1.2) $$\phi_c(X) = \phi_c(X_1, X_2, \ldots, X_n) := \operatorname{tr}(X_{i_1} X_{i_2} \ldots X_{i_p}).$$

The previous proposition then follows from the following:

Theorem.

(6.1.3) $$\operatorname{tr}(s \circ X_1 \otimes X_2 \otimes \cdots \otimes X_n) = \prod_{j=1}^{k} \phi_{c_j}(X).$$

Proof. We first remark that for fixed s, both sides of 6.1.3 are multilinear functions of the matrix variables X_i. Therefore in order to prove this formula it is enough to do it when $X_i = u_i \otimes \psi_i$ is decomposable.

Let us apply in this case the operator $s \circ X_1 \otimes X_2 \otimes \cdots \otimes X_n$ to a decomposable tensor $v_1 \otimes v_2 \cdots \otimes v_n$. We have

(6.1.4)

$$s \circ X_1 \otimes X_2 \otimes \cdots \otimes X_n(v_1 \otimes v_2 \cdots \otimes v_n) = \prod_{i=1}^{n} \langle \psi_i | v_i \rangle u_{s^{-1}1} \otimes u_{s^{-1}2} \ldots \otimes u_{s^{-1}n}.$$

This formula shows that

(6.1.5)

$$s \circ X_1 \otimes X_2 \otimes \cdots \otimes X_n = (u_{s^{-1}1} \otimes \psi_1) \otimes (u_{s^{-1}2} \otimes \psi_2) \ldots \otimes (u_{s^{-1}n} \otimes \psi_n),$$

so that

$$\text{(6.1.6)} \qquad \operatorname{tr}(s \circ X_1 \otimes X_2 \otimes \cdots \otimes X_n) = \prod_{i=1}^{n} \langle \psi_i | u_{s^{-1}i} \rangle = \prod_{i=1}^{n} \langle \psi_{s(i)} | u_i \rangle.$$

Now let us compute for a cycle $c := (i_p \; i_{p-1} \; \ldots \; i_1)$ the function

$$\phi_c(X) = \operatorname{tr}(X_{i_1} X_{i_2} \ldots \; X_{i_p}).$$

We get

$$\begin{aligned} &\operatorname{tr}(u_{i_1} \otimes \psi_{i_1} \circ u_{i_2} \otimes \psi_{i_2} \circ \cdots \circ u_{i_p} \otimes \psi_{i_p}) \\ &= \operatorname{tr}(u_{i_1} \otimes \langle \psi_{i_1} | u_{i_2} \rangle \langle \psi_{i_2} | u_{i_3} \rangle \cdots \langle \psi_{i_{p-1}} | u_{i_p} \rangle \psi_{i_p}) \\ \text{(6.1.7)} \qquad &= \langle \psi_{i_1} | u_{i_2} \rangle \langle \psi_{i_2} | u_{i_3} \rangle \cdots \langle \psi_{i_{p-1}} | u_{i_p} \rangle \langle \psi_{i_p} | u_{i_1} \rangle = \prod_{j=1}^{p} \langle \psi_{c(i_j)} | u_{i_j} \rangle. \end{aligned}$$

Formulas 6.1.6 and 6.1.7 imply the claim. □

6.2 Character of $S_\lambda(V)$

According to Theorem 3.2 of Chapter 2, the formal ring of symmetric functions in infinitely many variables has as basis all Schur functions S_λ. The restriction to symmetric functions in m-variables sets to 0 all S_λ with height $> m$.

We are ready to complete our work. Let $m = \dim V$. For a matrix $X \in GL(V)$ and a partition $\lambda \vdash n$ of height $\leq m$, let us denote by $S_\lambda(X) := S_\lambda(x)$ the Schur function evaluated at $x = (x_1, \ldots, x_m)$, the eigenvalues of X.

Theorem. *Denote $\rho_\lambda(X)$ to be the character of the representation $S_\lambda(V)$ of $GL(V)$, paired with the representation M_λ of S_n in $V^{\otimes n}$. We have $\rho_\lambda(X) = S_\lambda(X)$.*

Proof. If $s \in S_n$, $X \in GL(V)$, we have seen that the trace of $s \circ X^{\otimes n}$ on $V^{\otimes n}$ is computed by $\psi_\mu(X) = \sum_\lambda c_\lambda(\mu) S_\lambda(X)$ (definition of the c_λ).

If $m = \dim V < n$, only the partitions of height $\leq m$ contribute to the sum. On the other hand, $V^{\otimes n} = \bigoplus_{ht(\lambda)\leq \dim(V)} M_\lambda \otimes S_\lambda(V)$; thus,

$$\psi_\mu(X) = \mathrm{tr}(s \circ X^{\otimes n}) = \sum_{\lambda \vdash n, ht(\lambda)\leq m} \mathrm{tr}(s \mid M_\lambda)\,\mathrm{tr}(X^{\otimes n} \mid S_\lambda(V))$$
$$= \sum_{\lambda \vdash n, ht(\lambda)\leq m} c_\lambda(\mu)\rho_\lambda(X) = \sum_{\lambda \vdash n, ht(\lambda)\leq m} c_\lambda(\mu) S_\lambda(X).$$

If $m < n$, the $\psi_\mu(X)$ with parts of length $\leq m$ (i.e., $ht(\tilde{\mu}) \leq m$) are a basis of symmetric functions in m variables; hence we can invert the system of linear equations and get $S_\lambda(X) = \rho_\lambda(X)$. □

The eigenvalues of $X^{\otimes n}$ are monomials in the variables x_i, and thus we obtain:

Corollary. *$S_\lambda(x)$ is a sum of monomials with positive coefficients.*

We will see in Chapter 13 that one can index combinatorially the monomials which appear by *semistandard tableaux*.

We can also deduce a dimension formula for the space $S_\lambda(V)$, $\dim V = n$. Of course its value is $S_\lambda(1, 1, \dots, 1)$ which we want to compute from the determinantal formulas giving $S_\lambda(x) = A_{\lambda+\varrho}(x)/V(x)$.

Let as usual $\lambda := h_1, h_2, \dots, h_n$ and $l_i := h_i + n - i$. Of course we cannot substitute directly the number 1 for the x_i, or we get $0/0$. Thus we first substitute to $x_i \to x^{i-1}$ and then take the limit as $x \to 1$. Under the previous substitution we see that $A_{\lambda+\varrho}$ becomes the Vandermonde determinant of the elements x^{l_i}, hence

$$S_\lambda(1, x, x^2, \dots, x^{n-1}) = \prod_{1\leq i<j\leq n} \frac{(x^{l_i} - x^{l_j})}{(x^{n-i} - x^{n-j})}.$$

If $a > b$, we have $x^a - x^b = x^b(x-1)(x^{a-b-1} + x^{a-b-2} + \dots + 1)$, hence we deduce that

$$\dim S_\lambda(V) = S_\lambda(1, 1, 1, \dots, 1) = \prod_{1\leq i<j\leq n} \frac{(l_i - l_j)}{(j-i)} = \prod_{1\leq i<j\leq n} \frac{(h_i - h_j + j - i)}{(j-i)}.$$

6.3 Cauchy Formula as Representations

We want to now give an interpretation, in the language of representations, of the Cauchy formula.

Suppose we are given a vector space U over which a torus T acts with a basis of weight vectors u_i with weight χ_i.

The graded character of the action of T on the symmetric and exterior algebras are given by Molien's formula, §4.5 and are, respectively,

(6.3.1)
$$\frac{1}{\prod 1 - \chi_i q}, \qquad \prod 1 + \chi_i q.$$

As an example consider two vector spaces U, V with bases $u_1, \dots, u_m$; $v_1, \dots, v_n$, respectively. We may assume $m \leq n$.

The maximal tori of diagonal matrices have eigenvalues $x_1, \dots, x_m$; $y_1, \dots, y_n$ respectively. On the tensor product we have the action of the product torus, and the basis $u_i \otimes v_j$ has eigenvalues $x_i y_j$. Therefore the graded character on the symmetric algebra $S(U \otimes V)$ is $\prod_{i=1}^{n} \prod_{j=1}^{m} \frac{1}{1 - x_i y_j q}$.

By Cauchy's formula we deduce that the character of the n^{th} symmetric power $S^n(U \otimes V)$ equals $\sum_{\lambda \vdash n,\ ht(\lambda) \leq m} S_\lambda(x) S_\lambda(y)$.

We know that the rational representations of $GL(U) \times GL(V)$ are completely reducible and their characters can be computed by restricting to diagonal matrices. Thus we have the description:

(6.3.2)
$$S^n(U \otimes V) = \bigoplus_{\lambda \vdash n,\ ht(\lambda) \leq m} S_\lambda(U) \otimes S_\lambda(V).$$

This is also referred to as *Cauchy's formula*.

Observe that if $W \subset V$ is the subspace which is formed by the first k basis vectors, then the intersection of $S_\lambda(U) \otimes S_\lambda(V)$ with $S(U \otimes W)$ has as basis the part of the basis of weight vectors of $S_\lambda(U) \otimes S_\lambda(V)$ corresponding to weights in which the variables y_j, $j > k$ do not appear. Thus its character is obtained by setting to 0 these variables in $S_\lambda(y_1, y_2, \dots, y_m)$; thus we clearly get that

(6.3.3)
$$S_\lambda(U) \otimes S_\lambda(V) \cap S(U \otimes W) = S_\lambda(U) \otimes S_\lambda(W).$$

Similarly it is clear, from Definition 3.1.3: $S_\lambda(V) := e_T V^{\otimes n}$, that:

Proposition. *If $U \subset V$ is a subspace, then $S_\lambda(U) = S_\lambda(V) \cap U^{\otimes n}$.*

6.4 Multilinear Elements

Consider a rational representation $\rho : GL(n, \mathbb{C}) \to GL(W)$ for which the matrix coefficients are polynomials in the coordinates $x_{i,j}$, and thus do not contain the determinant at the denominator.

Such a representation is called a *polynomial representation*, the map ρ extends to a multiplicative map $\rho : M(n, \mathbb{C}) \to \mathrm{End}(W)$ on all matrices.

Polynomial representations are closed under taking direct sums, tensor products, subrepresentations and quotients. A typical polynomial representation of $GL(V)$ is $V^{\otimes n}$ and all its subrepresentations, for instance the $S_\lambda(V)$.

One should stress the strict connection between the two formulas, 6.3.2 and 3.1.4.

(6.3.2)
$$S^n(U \otimes V) = \bigoplus_{\lambda \vdash n,\ ht(\lambda) \leq m} S_\lambda(U) \otimes S_\lambda(V),$$

(3.1.4)
$$V^{\otimes n} = \bigoplus_{ht(\lambda) \leq \dim(V)} M_\lambda \otimes S_\lambda(V).$$

This is clearly explained when we assume that $U = \mathbb{C}^n$ with canonical basis e_i and we consider the diagonal torus T acting by matrices $Xe_i = x_i e_i$.

Let us go back to formula 6.3.2, and apply it when $\dim V = n$, $W = \mathbb{C}^n$.

Consider the subspace T_n of $S(\mathbb{C}^n \otimes V)$ formed by the elements $\prod_{i=1}^n e_i \otimes v_i$, $v_i \in V$. T_n is stable under the subgroup $S_n \times GL(V) \subset GL(n, \mathbb{C}) \times GL(V)$, where S_n is the group of permutation matrices. We have a mapping $i : V^{\otimes n} \to T_n$ defined by

$$i : v_1 \otimes v_2 \otimes \cdots \otimes v_n \mapsto \prod_{i=1}^n e_i \otimes v_i. \tag{6.4.1}$$

Proposition. *(i) T_n is the weight space in $S(\mathbb{C}^n \otimes V)$, of weight $\chi(X) = \prod_i x_i$ for the torus T.*

(ii) The map i is an $S_n \times GL(V)$ linear isomorphism between $V^{\otimes n}$ and T_n.

Proof. The verification is immediate and left to the reader. □

Remark. The character $\chi := \prod_{i=1}^n x_i$ is invariant under the symmetric group (and generates the group of these characters). We call it *the multilinear character.*

As usual, when we have a representation W of a torus, we denote by W^χ the weight space of character χ.

Now for every partition λ consider

$$S_\lambda(\mathbb{C}^n)^\chi := \left\{ u \in S_\lambda(\mathbb{C}^n) | Xu = \prod_i x_i u, \ \forall X \in T \right\}, \tag{6.4.2}$$

the weight space of $S_\lambda(\mathbb{C}^n)$ formed by the elements which are formally *multilinear.*

Since the character $\prod_i x_i$ is left invariant by conjugation by permutation matrices it follows that the symmetric group $S_n \subset GL(n, \mathbb{C})$ of permutation matrices acts on $S_\lambda(\mathbb{C}^n)^\chi$. We claim that:

Proposition. $S_\lambda(\mathbb{C}^n)^\chi = 0$ *unless $\lambda \vdash n$ and in this case $S_\lambda((\mathbb{C}^n)^*)^\chi$ is identified with the irreducible representation M_λ of S_n.*

Proof. In fact assume $Xu = \prod_i x_i u$. Clearly u is in a polynomial representation of degree n. On the other hand

$$S^n(\mathbb{C}^n \otimes V) = \bigoplus_{\lambda \vdash n} S_\lambda((\mathbb{C}^n)^*) \otimes S_\lambda(V),$$

hence

(6.4.3)
$$V^{\otimes n} := S^n(\mathbb{C}^n \otimes V)^\chi = \bigoplus_{\lambda \vdash n} S_\lambda((\mathbb{C}^n)^*)^\chi \otimes S_\lambda(V) = \bigoplus_{\lambda \vdash n} M_\lambda \otimes S_\lambda(V)$$

and we get the required identification. □

Therefore, given a polynomial representation P of $GL(n, \mathbb{C})$, if it is homogeneous of degree n, in order to determine its decomposition $P = \bigoplus_{\lambda \vdash n} m_\lambda S_\lambda(\mathbb{C}^n)$ we can equivalently restrict to $M := \{p \in P \mid X \cdot p = \prod_i x_i p\}$, the multilinear weight space (for X diagonal with entries x_i) and see how it decomposes as a representation of S_n since

$$P = \bigoplus_{\lambda \vdash n} m_\lambda S_\lambda(\mathbb{C}^n) \iff M = \bigoplus_{\lambda \vdash n} m_\lambda M_\lambda. \tag{6.4.4}$$

7 Polynomial Functors

7.1 Schur Functors

Consider two vector spaces V, W, the space $\hom(V, W) = W \otimes V^*$, and the ring of polynomial functions on $\hom(V, W)$ decomposed as

$$(7.1.1) \qquad \mathcal{P}[\hom(V, W)] = S(W^* \otimes V) = \bigoplus\nolimits_\lambda S_\lambda(W^*) \otimes S_\lambda(V).$$

A way to explicitly identify the spaces $S_\lambda(W^*) \otimes S_\lambda(V)$ as spaces of functions is obtained by a variation of the method of matrix coefficients.

We start by stressing the fact that the construction of the representation $S_\lambda(V)$ from V is in a sense *natural*, in the language of categories.

Recall that a map between vector spaces is called a *polynomial map* if in coordinates it is given by polynomials.

Definition. A functor F from the category of vector spaces to itself is called a *polynomial* functor if, given two vector spaces V, W, the map $A \to F(A)$ from the vector space $\hom(V, W)$ to the vector space *hom*$(F(V), F(W))$ is a polynomial map.

We say that F is homogeneous of degree k if, for all vector spaces V, W, the map $\hom(V, W) \xrightarrow{F(-)} \hom(F(V), F(W))$ is homogeneous of degree k.

The functor $F : V \to V^{\otimes n}$ is clearly a polynomial functor, homogeneous of degree n. When $A : V \to W$ the map $F(A)$ is $A^{\otimes n}$.

We can now justify the word *Schur functor* in the Definition 3.1.3, $S_\lambda(V) := e_T V^{\otimes n}$, where e_T is a Young symmetrizer, associated to a partition λ.

As V varies, $V \mapsto S_\lambda(V)$ can be considered as a functor. In fact it is a subfunctor of the tensor power, since clearly, if $A : V \to W$ is a linear map, $A^{\otimes n}$ commutes with a_T. Thus $A^{\otimes n}(e_T V^{\otimes n}) \subset e_T W^{\otimes n}$ and we define

$$(7.1.2) \qquad S_\lambda(A) : S_\lambda(V) \longrightarrow V^{\otimes n} \xrightarrow{A^{\otimes n}} W^{\otimes n} \xrightarrow{e_T} S_\lambda(W).$$

Summarizing:

Proposition 1. *Given any partition $\mu \vdash n$, $V \mapsto S_\mu(V)$ is a homogeneous polynomial functor on vector spaces of degree n, called a* ***Schur functor****.*

Remark. This functor is independent of T but depends only on the partition λ. The choice of T determines an embedding of $S_\lambda(V)$ as subfunctor of $V^{\otimes n}$.

Remark. The exterior and symmetric power $\bigwedge^k V$ and $S^k(V)$ are examples of Schur functors.

Since the map $S_\mu : \hom(V, W) \to \hom(S_\mu(V), S_\mu(W))$ defined by $S_\mu : X \to S_\mu(X)$ is a homogeneous polynomial map of degree n, the dual map $S_\mu^* : \hom(S_\mu(V), S_\mu(W))^* \to \mathcal{P}[\hom(V, W)]$ defined by

$$S_\mu^*(\phi)(X) := \langle \phi | S_\mu(X) \rangle, \ \ \phi \in \hom(S_\mu(V), S_\mu(W))^*, \ \ X \in \hom(V, W)$$

is a $GL(V) \times GL(W)$-equivariant map into the homogeneous polynomials of degree n.

By the irreducibility of $\hom(S_\mu(V), S_\mu(W))^* = S_\mu(V) \otimes S_\mu(W)^*$, S_μ^* must be a linear isomorphism to an irreducible submodule of $\mathcal{P}[\hom(V, W)]$ uniquely determined by Cauchy's formula. By comparing the isotypic component of type $S_\mu(V)$ we deduce:

Proposition 2. $\mathcal{P}[\hom(V, W)] = \bigoplus_\mu \hom(S_\mu(V), S_\mu(W))^*$ *and we have the isomorphism* $S_\mu(W^*) = S_\mu(W)^*$.

Let us apply the previous discussion to $\hom(\bigwedge^i V, \bigwedge^i W)$.

Choose bases e_i, $i = 1, \dots, h$, f_j, $j = 1, \dots, k$ for V, W respectively, and identify the space $\hom(V, W)$ with the space of $k \times h$ matrices. Thus the ring $\mathcal{P}[\hom(V, W)]$ is the polynomial ring $\mathbb{C}[x_{ij}], i = 1, \dots, h, \ j = 1, \dots, k$ where x_{ij} are the matrix entries.

Given a matrix X the entries of $\bigwedge^i X$ are the determinants of all the minors of order i extracted from X, and:

Corollary. $\bigwedge^i V \otimes (\bigwedge^i W)^*$ *can be identified with the space of polynomials spanned by the determinants of all the minors of order* i *extracted from* X, *which is thus irreducible as a representation of* $GL(V) \times GL(W)$.

7.2 Homogeneous Functors

We want to prove that any polynomial functor is equivalent to a direct sum of Schur functors. We start with:

Proposition 1. *A polynomial functor is a direct sum of homogeneous functors.*

Proof. The scalar multiplications by $\alpha \in \mathbb{C}^*$ on a space V induce, by functoriality, a polynomial representation of $\mathbb{C}^*$ on $F(V)$ which then decomposes as $F(V) = \bigoplus_k F_k(V)$, with $F_k(V)$ the subspace of weight α^k. Clearly $F_k(V)$ is a subfunctor and $F = \bigoplus_k F_k(V)$. Moreover $F_k(V)$ is a homogeneous functor of degree k.

We can polarize a homogeneous functor of degree k as follows. Consider, for a k-tuple $V_1, \dots, V_k$ of vector spaces, their direct sum $\bigoplus_i V_i$ together with the action of a k-dimensional torus T with the scalar multiplication x_i on each summand V_i. T acts in a polynomial way on $F(\bigoplus_i V_i)$ and we can decompose by weights

$$F(\bigoplus\nolimits_i V_i) = \bigoplus\nolimits_\lambda F_\lambda(V_1, \dots, V_k), \quad \lambda = x_1^{h_1} x_2^{h_2} \dots x_k^{h_k}, \ \sum h_i = k.$$

One easily verifies that the inclusion $V_i \to \oplus V_j$ induces an isomorphism between $F(V_i)$ and $F_{x_i^k}(V_1, \dots, V_k)$. $\square$

Let us now consider a polynomial functor $V \mapsto F(V)$, homogeneous of degree k. We start by performing the following constructions.

(1) First consider the functor

$$T : V \mapsto S^k(\hom(\mathbb{C}^k, V)) \otimes F(\mathbb{C}^k)$$

And the natural transformation:

$$\pi_V : S^k(\hom(\mathbb{C}^k, V)) \otimes F(\mathbb{C}^k) \to F(V)$$

defined by the formula

$$\pi_V(f^k \otimes u) := F(f)(u), \ f \in \hom(\mathbb{C}^k, V), \ u \in F(\mathbb{C}^k)$$

This formula makes sense, since $F(f)$ is a homogeneous polynomial map of degree k in f by hypothesis.

The fact that π_V is natural depends on the fact that, if $h : V \to W$ we have that $T(h)(f^k \otimes u) = (hf)^k \otimes u$ so that

$$F(h)\pi_V(f^k \otimes u) = F(hf)(u) = \pi_W((hf)^k \otimes u = \pi_W(T(h)(f^k \otimes u).$$

(2) The linear group $GL(k, \mathbb{C})$ acts by natural isomorphisms on the functor $T(V)$ by the formula

$$(f \circ g^{-1})^k \otimes F(g)u,$$

Lemma 1. *The natural transformation π_V is $GL(k, \mathbb{C})$ invariant.*

Proof. We have $F(f \circ g^{-1}) = F(f) \circ F(g)^{-1}$ and $\pi_V((f \circ g^{-1})^k \otimes F(g)u) = F(f \circ g^{-1})(F(g)u) = F(f)u.$ □

These invariance properties mean that if we decompose $T(V)$, as a representation of $GL(n, \mathbb{C})$ into the invariant space $T(V)^{GL(n,\mathbb{C})}$ and the other isotypic components, the sum of the nontrivial irreducible representations $T(V)_{GL(n,\mathbb{C})}$, we have $\pi_V = 0$ on $T(V)_{GL(n,\mathbb{C})}$.

Our goal is to prove

Theorem. *The map π_V restricted to the $GL(n, \mathbb{C})$ invariants:*

$$\pi_V : [S^k(\hom(\mathbb{C}^k, V)) \otimes F(\mathbb{C}^k)]^{GL(k,\mathbb{C})} \to F(V), \quad \pi_V(f \otimes u) := F(f)(u),$$

is a functorial isomorphism.

In order to prove this theorem we need a simple general criterion:

Proposition 2. *Let $\eta : F \to G$ be a natural transformation of polynomial functors, each of degree k. Then η is an isomorphism if and only if $\eta_{\mathbb{C}^k} : F(\mathbb{C}^k) \to G(\mathbb{C}^k)$ is an isomorphism.*

Proof. Since any vector space is isomorphic to $\mathbb{C}^m$ for some m we have to prove an isomorphism for these spaces.

The diagonal torus acts on $\mathbb{C}^m$, which by functoriality acts also on $F(\mathbb{C}^m)$ and $G(\mathbb{C}^m)$. By naturality $\eta_{\mathbb{C}^m} : F(\mathbb{C}^m) \to G(\mathbb{C}^m)$ must preserve weight spaces with respect to the diagonal matrices. Now each weight involves at most k indices and so it can be deduced from the corresponding weight space for $\mathbb{C}^k$. For these weight spaces the isomorphism is guaranteed by the hypotheses. $\square$

In order to apply this criterion to π_V we have to understand the map:

$$\pi_{\mathbb{C}^k} : [S^k(\hom(\mathbb{C}^k, \mathbb{C}^k)) \otimes F(\mathbb{C}^k)]^{GL(k,\mathbb{C})} \to F(\mathbb{C}^k).$$

The invariants are taken with respect to the diagonal action on $S^k(\hom(\mathbb{C}^k, \mathbb{C}^k))$ by acting on the source of the homomorphisms and on $F(\mathbb{C}^k)$.

Lemma 2. *$\pi_{\mathbb{C}^k}$ is an isomorphism.*

Proof. The definition of this map depends just on the fact that $F(\mathbb{C}^k)$ is a polynomial representation of $GL(k, \mathbb{C})$ which is homogeneous of degree k. It is clear that, if this map is an isomorphism for two different representations $F_1(\mathbb{C}^k)$, $F_2(\mathbb{C}^k)$ it is also an isomorphism for their direct sum. Thus we are reduced to study the case in which $F(\mathbb{C}^k) = S_\lambda(\mathbb{C}^k)$ for some partition $\lambda \vdash k$.

Identifying

$$S^k(\hom(\mathbb{C}^k, \mathbb{C}^k)) \otimes S_\lambda(\mathbb{C}^k) = \bigoplus_{\mu \vdash k} S_\mu(\mathbb{C}^k) \otimes S_\mu(\mathbb{C}^k)^* \otimes S_\lambda(\mathbb{C}^k)$$

the invariants are by definition

$$\bigoplus_{\mu \vdash k} S_\mu(\mathbb{C}^k) \otimes [S_\mu(\mathbb{C}^k)^* \otimes S_\lambda(\mathbb{C}^k)]^{GL(k,\mathbb{C})} = S_\lambda(\mathbb{C}^k) \otimes \mathbb{C}1_{S_\lambda(\mathbb{C}^k)}.$$

By the irreducibility of the representations $S_\lambda(\mathbb{C}^k)$.

Since clearly $\pi_{\mathbb{C}^k}(u \otimes 1_{S_\lambda(\mathbb{C}^k)}) = u$ we have proved the claim. $\square$

By the classification of polynomial representations, we have that

$$F(\mathbb{C}^k) = \bigoplus_{\lambda \vdash k} m_\lambda S_\lambda(\mathbb{C}^k)$$

for some nonnegative integers m_i. We deduce:

Corollary. *A polynomial functor F of degree k is of the form*

$$F(V) = \bigoplus_{\lambda \vdash k} m_\lambda S_\lambda(V).$$

Proof.

$$\begin{aligned}[S^k(\hom(\mathbb{C}^k, V)) \otimes S_\lambda(\mathbb{C}^k)]^{GL(k,\mathbb{C})} &= \bigoplus_{\mu \vdash k} S_\mu(V) \otimes [S_\mu(\mathbb{C}^k)^* \otimes S_\lambda(\mathbb{C}^k)]^{GL(k,\mathbb{C})} \\ &= S_\lambda(V) \otimes \mathbb{C}1_{S_\lambda(\mathbb{C}^k)}.\end{aligned}$$

$\square$

Polynomial functors can be summed (direct sum) multiplied (tensor product) and composed. All these operations can be extended to a ring whose elements are purely formal differences of functors (a Grothendieck type of ring). In analogy with the theory of characters an element of this ring is called a *virtual functor.*

Proposition 3. *The ring of virtual functors is canonically isomorphic to the ring of infinite symmetric functions.*

Proof. We identify the functor S_λ with the symmetric function $S_\lambda(x)$. $\square$

Exercise. Given two polynomial functors F, G of degree k prove that we have an isomorphism between the space $\mathrm{Nat}(F, G)$ of natural transformations between the two functors, and the space $\hom_{GL(k,\mathbb{C})}(F(\mathbb{C}^k), G(\mathbb{C}^k))$.

Discuss the case $F = G$ the tensor power $V^{\otimes k}$.

7.3 Plethysm

The composition of functors becomes the *Plethysm operation* on symmetric functions. In general it is quite difficult to compute such compositions, even such simple ones as $\bigwedge^i(\bigwedge^h V))$. There are formulas for $S^k(S^2(V))$, $S^k(\bigwedge^2(V))$ and some dual ones.

In general the computation $F \circ G$ should be done according to the following:

Algorithm. Apply a polynomial functor G to the space $\mathbb{C}^m$ with its standard basis.

For the corresponding linear group and diagonal torus T, $G(\mathbb{C}^m)$ is a polynomial representation of some dimension N. It then has a basis of T-weight vectors with characters a list of monomials M_i. The character of T on $G(\mathbb{C}^m)$ is $\sum_i^N M_i$, a symmetric function $S_G(x_1, \dots, x_m)$.

If G is homogeneous of degree k, this symmetric function is determined as soon as $m \geq k$.

When we apply F to $G(\mathbb{C}^m)$ we use the basis of weight vectors to see that the symmetric function

$$S_{F\circ G}(x_1, \dots, x_m) = S_F(M_1, \dots, M_N). \qquad \square$$

Some simple remarks are in order. First, given a fixed functor G the map $F \mapsto F \circ G$ is clearly a ring homomorphism. Therefore it is determined by the value on a set of generators. One can choose as generators the exterior powers. In this case the operation $\bigwedge^i \circ F$ as transformations in F are called λ-operations and written λ^i.

These operations satisfy the basic law: $\lambda^i(a + b) = \sum_{h+k=i} \lambda^h(a)\lambda^k(b)$.

It is also convenient to use as generators the Newton functions $\psi_k = \sum_i x_i^k$ since then

$$\psi_k(S(x_1, \dots, x_m)) = S(x_1^k, \dots, x_m^k), \qquad \psi_k(\psi_h) = \psi_{kh}.$$

All of this can be formalized, giving rise to the theory of λ-rings (cf. [Knu]).

8 Representations of the Linear and Special Linear Groups

8.1 Representations of $SL(V)$, $GL(V)$

Given an n-dimensional vector space V we want to give the complete list of irreducible representations for the general and special linear groups $GL(n) = GL(V), SL(n) = SL(V)$.

From Chapter 7, Theorem 1.4 we know that all the irreducible representations of $SL(V)$ appear in the tensor powers $V^{\otimes m}$ and all the irreducible representations of $GL(V)$ appear in the tensor powers $V^{\otimes m}$ tensored with integer powers of the determinant $\bigwedge^n(V)$. For simplicity we will denote by $D := \bigwedge^n(V)$ and by convention $D^{-1} := \bigwedge^n(V)^*$. From what we have already seen the irreducible representations of $GL(V)$ which appear in the tensor powers are the modules $S_\lambda(V)$, $ht(\lambda) \leq n$. They are all distinct since they have distinct characters. Given $S_\lambda(V) \subset V^{\otimes m}$, $\lambda \vdash m$, consider $S_\lambda(V) \otimes \bigwedge^n(V) \subset V^{\otimes m+n}$. Since $\bigwedge^n(V)$ is 1-dimensional, clearly $S_\lambda(V) \otimes \bigwedge^n(V)$ is also irreducible. Its character is $S_{\lambda+1^n}(x) = (x_1x_2\dots x_n)S_\lambda(x)$, hence (cf. Chapter 2, 6.2.1):

$$S_\lambda(V) \otimes \bigwedge^n(V) = S_{\lambda+1^n}(V). \tag{8.1.1}$$

We now need a simple lemma. Let $\mathcal{P}_{n-1} := \{k_1 \geq k_2 \geq \dots \geq k_{n-1} \geq 0\}$ be the set of all partitions (of any integer) of height $\leq n-1$. Consider the polynomial ring $\mathbb{Z}[x_1, \dots, x_n, (x_1, x_2 \dots x_n)^{-1}]$ obtained by inverting $e_n = x_1x_2\dots x_n$.

Lemma. *(i) The ring of symmetric elements in $\mathbb{Z}[x_1, \dots, x_n, (x_1x_2\dots x_n)^{-1}]$ is generated by $e_1, e_2, \dots, e_{n-1}, e_n^{\pm 1}$ and it has as basis the elements*

$$S_\lambda e_n^m, \ \lambda \in \mathcal{P}_{n-1}, \ m \in \mathbb{Z}.$$

(ii) The ring $\mathbb{Z}[e_1, e_2, \dots, e_{n-1}, e_n]/(e_n - 1)$ has as basis the classes of the elements S_λ, $\lambda \in \mathcal{P}_{n-1}$.

Proof. (i) Since e_n is symmetric it is clear that a fraction f/e_n^k is symmetric if and only if f is symmetric, hence the first statement. Any element of $\mathbb{Z}[e_1, e_2, \dots, e_{n-1}, e_n^{\pm 1}]$ can be written in a unique way in the form $\sum_{k\in\mathbb{Z}} a_k e_n^k$ with $a_k \in \mathbb{Z}[e_1, e_2, \dots, e_{n-1}]$. We know that the Schur functions S_λ, $\lambda \in \mathcal{P}_{n-1}$ are a basis of $\mathbb{Z}[e_1, e_2, \dots, e_{n-1}, e_n]/(e_n)$ and the claim follows.

(ii) follows from (i). □

Theorem. *(i) The list of irreducible representations of $SL(V)$ is*

$$S_\lambda(V), \ ht(\lambda) \leq n-1. \tag{8.1.2}$$

(ii) The list of irreducible representations of $GL(V)$ is

$$S_\lambda(V) \otimes D^k, \ ht(\lambda) \leq n-1, \ k \in \mathbb{Z}. \tag{8.1.3}$$

Proof. (i) The group $GL(V)$ is generated by $SL(V)$ and the scalar matrices which commute with every element. Therefore in any irreducible representation of $GL(V)$ the scalars in $GL(V)$ also act as scalars in the representation. It follows immediately that the representation remains irreducible when restricted to $SL(V)$.

Thus we have to understand when two irreducible representations $S_\lambda(V)$, $S_\mu(V)$, with $ht(\lambda) \leq n$, $ht(\mu) \leq n$, are isomorphic once restricted to $SL(V)$.

Any λ can be uniquely written in the form $(m, m, m, \ldots, m) + (k_1, k_2, \ldots, k_{n-1}, 0)$ or $\lambda = \mu + m\, 1^n$, $ht(\mu) \leq n-1$, and so $S_\lambda(V) = S_\mu(V) \otimes D^m$. Clearly $S_\lambda(V) = S_\mu(V)$ as representations of $SL(V)$. Thus to finish we have to show that if $\lambda \neq \mu$ are two partitions of height $\leq n-1$, the two $SL(V)$ representations $S_\lambda(V)$, $S_\mu(V)$ are not isomorphic. This follows from the fact that the characters of the representations $S_\lambda(V)$, $\lambda \in \mathcal{P}_{n-1}$, are a basis of the invariant functions on $SL(V)$, by the previous lemma.

(ii) We have seen at the beginning of this section that all irreducible representations of $GL(V)$ appear in 8.1.3 above. Now if two different elements of this list in 8.1.3 were isomorphic, by multiplying by a high enough power of D we would obtain two isomorphic polynomial representations belonging to two different partitions, a contradiction. $\square$

Remark. $\bigwedge^i V$ corresponds to the partition 1^i, made of a single column of length i. Its associated Schur function S_{1^i} is the i^{th} elementary function e_i. Instead the symmetric power $S^i(V)$ corresponds to the partition made of a single row of length i, it corresponds to a symmetric function S_i which is often denoted by h_i and it is the sum of all the monomials of degree i.

8.2 The Coordinate Ring of the Linear Group

We can interpret the previous theory in terms of the coordinate ring $\mathbb{C}[GL(V)]$ of the general linear group.

Since $GL(V)$ is the open set of $\mathrm{End}(V) = V \otimes V^*$ where the determinant $d \neq 0$, its coordinate ring is the localization at d of the ring $S(V^* \otimes V)$ which, under the two actions of $GL(V)$, decomposes as $\bigoplus_{ht(\lambda)\leq n} S_\lambda(V^*) \otimes S_\lambda(V) = \bigoplus_{ht(\lambda)\leq n-1,\, k\geq 0} d^k S_\lambda(V^*) \otimes S_\lambda(V)$. It follows immediately then that

$$(8.2.1) \qquad \mathbb{C}[GL(V)] = \bigoplus_{ht(\lambda)\leq n-1,\, k\in\mathbb{Z}} d^k S_\lambda(V^*) \otimes S_\lambda(V).$$

This of course is, for the linear group, the explicit form of formula 3.1.1 of Chapter 7. From it we deduce that

$$(8.2.2) \qquad S_\lambda(V^*) = S_\lambda(V)^*, \qquad \forall \lambda,\ ht(\lambda) \leq n-1, \quad (d^* = d^{-1}).$$

Similarly,

$$(8.2.3) \qquad \mathbb{C}[SL(V)] = \bigoplus_{ht(\lambda)\leq n-1} S_\lambda(V^*) \otimes S_\lambda(V).$$

8.3 Determinantal Expressions for Schur Functions

In this section we want to discuss a determinant development for Schur functions which is often used.

Recall that $V_\lambda = e_T(V^{\otimes n})$ is a quotient of

$$a_T(V^{\otimes n}) = \bigwedge^{k_1} V \otimes \bigwedge^{k_2} V \cdots \otimes \bigwedge^{k_n} V,$$

where the k_i are the columns of a tableau T, and is contained in

$$s_T(V^{\otimes n}) = S^{h_1}(V) \otimes S^{h_2}(V) \cdots \otimes S^{h_n}(V)$$

where the h_i are the rows of T. Here one has to interpret both antisymmetrization and symmetrization as occurring respectively in the columns and row indices.[78] The composition $e_T = \frac{1}{p(\lambda)} s_T a_T$ can be viewed as the result of a map

$$\bigwedge^{k_1} V \otimes \bigwedge^{k_2} V \cdots \otimes \bigwedge^{k_n} V \to s_T(V^{\otimes n}) = S^{h_1}(V) \otimes S^{h_2}(V) \cdots \otimes S^{h_n}(V).$$

As representations $\bigwedge^{k_1} V \otimes \bigwedge^{k_2} V \cdots \otimes \bigwedge^{k_n} V$ and $S^{h_1}(V) \otimes S^{h_2}(V) \cdots \otimes S^{h_n}(V)$ decompose in the direct sum of a copy of V_λ and other irreducible representations.

The character of the exterior power $\bigwedge^i(V)$ is the elementary symmetric function $e_i(x)$. The one of $S^i(V)$ is the function $h_i(x)$ sum of all monomials of degree i.

In the formal ring of symmetric functions there is a formal duality between the elements e_i and the h_j. From the definition of the e_i and from Molien's formula:

$$\sum_{i=0}^{\infty} h_i(x) q^i = \frac{1}{\prod(1 - x_i q)}, \quad \sum_{i=0}^{\infty} (-1)^i e_i(x) q^i = \prod(1 - x_i q),$$

$$1 = \left(\sum_{i=0}^{\infty} (-1)^i e_i(x) q^i\right)\left(\sum_{i=0}^{\infty} h_i(x) q^i\right).$$

Hence for $m > 0$ we have $\sum_{i+j=m} (-1)^i e_i(x) h_j(x) = 0$. These identities tell us that

$$\mathbb{Z}[e_1, e_2, \dots, e_i, \dots] = \mathbb{Z}[h_1, h_2, \dots, h_i, \dots]$$

and also that we can present the ring of infinite symmetric functions with generators e_i, h_j and the previous relations:

$$(8.3.1) \qquad \mathbb{Z}[e_1, e_2, \dots, e_i, \dots; h_1, h_2, \dots, h_i, \dots] \Big/ \Big(\sum_{i+j=m} (-1)^i e_i h_j\Big).$$

The mapping $\tau : e_i(x) \mapsto h_i(x),\ h_i(x) \mapsto e_i(x)$ preserves these relations and gives an involutory automorphism in the ring of symmetric functions. Take the Cauchy identity

[78] It is awkward to denote symmetrization on non-consecutive indices as we did. More correctly, one should compose with the appropriate permutation which places the indices in the correct positions.

(8.3.2)
$$\sum S_\lambda(x)S_\lambda(y) = \prod_{i,j} \frac{1}{1 - x_i y_j} = \prod_j \sum_{k=0}^{\infty} h_k(x) y_j^k$$

and multiply it by the Vandermonde determinant $V(y)$, getting

$$\sum S_\lambda(x) A_{\lambda+\varrho}(y) = \prod_j \sum_{k=0}^{\infty} h_k(x) y_j^k V(y).$$

For a given $\lambda = a_1, \dots, a_n$ we see that $S_\lambda(x)$ is the coefficient of the monomial $y_1^{a_1+n-1} y_2^{a_2+n-2} \dots y_n^{a_n}$, and we easily see that this is

(8.3.3)
$$\sum_{\sigma \in S_n} \epsilon_\sigma \prod_{i=1}^{n} h_{\sigma(i)-i+a_i};$$

thus

Proposition. *The Schur function S_λ is the determinant of the $n \times n$ matrix, which in the position i, j, has the element h_{j-i+a_i} with the convention that $h_k = 0,\ \forall k < 0$.*

8.4 Skew Cauchy Formula

We want to complete this discussion with an interesting variation of the Cauchy formula (which is used in the computation of the cohomology of the linear group, cf. [AD]).

Given two vector spaces V, W we want to describe $\bigwedge(V \otimes W)$ as a representation of $GL(V) \times GL(W)$.

Theorem.

(8.4.1)
$$\bigwedge(V \otimes W) = \sum_\lambda S_\lambda(V) \otimes S_{\tilde{\lambda}}(W).$$

Proof. $\tilde{\lambda}$ denotes, as in Chapter 1, §1.1, the dual partition.

We argue in the following way. For very k, $\bigwedge^k(V \otimes W)$ is a polynomial representation of degree k of both groups; hence by the general theory $\bigwedge^k(V \otimes W) = \bigoplus_{\lambda \vdash k} S_\lambda(V) \otimes P_\mu(W)$ for some representations $P_\mu(W)$ to be determined. To do it in the stable case, where $\dim W = k$, we use Proposition 6.4 and formula 6.4.4, and we compute the multilinear elements of $P_\mu(W)$ as representations of S_n.

For this, as in 6.4, identify $W = \mathbb{C}^k$ with basis e_i and

$$\bigwedge^k(V \otimes W) = \bigwedge(\bigoplus_{i=1}^k V \otimes e_i) = \otimes_{i=1}^k \bigwedge(V \otimes e_i).$$

When we restrict to the multilinear elements we have $V \otimes e_1 \wedge \cdots \wedge V \otimes e_k$ which, as a representation of $GL(V)$, can be identified with $V^{\otimes n}$ except that the natural representation of the symmetric group $S_n \subset GL(n, \mathbb{C})$ is the canonical action on $V^{\otimes n}$ tensored by the sign representation.

Thus we deduce that if χ is the multilinear weight,

$$\bigoplus_{\lambda \vdash k} S_\lambda(V) \otimes M_{\tilde{\lambda}} = \left(\bigwedge^k V \otimes \mathbb{C}^k\right)^\chi.$$

This implies $P_\lambda(W) = M_{\tilde{\lambda}}$, hence $P_\lambda(W) = S_{\tilde{\lambda}}(W)$ from which the claim follows. □

Remark. In terms of characters, formula 8.4.1 is equivalent to the Cauchy formula Chapter 2, §4.1:

$$(8.4.2) \qquad \prod_{i=1,\ j=1}^{n,\ m} (1 + x_i y_j) = \sum_\lambda S_\lambda(x) S_{\tilde{\lambda}}(y).$$

There is a simple determinantal formula corollary of this identity, as in §8.3.

Here we remark that $\prod_{i=1}^n (1 + x_i y) = \sum_{j=0}^n e_j(x) y^j$ where the e_j are the elementary symmetric functions.

The same reasoning as in §8.3 then gives the formula

$$(8.4.3) \qquad S_\lambda(x) = \sum_{\sigma \in S_n} \epsilon_\sigma \prod_{i=1}^n e_{\sigma(i)-i+k_i}$$

where k_i are the columns of $\tilde{\lambda}$, i.e., the rows of λ.

Proposition. *The Schur function S_λ is the determinant of the $n \times n$ matrix which in the position i, j has the element e_{j-i+k_i}. The k_i are the rows of λ, with the convention that $e_k = 0,\ \forall k < 0$.*

From the two determinantal formulas found we deduce:

Corollary. *Under the involutive map $\tau : e_i \mapsto h_i$ we have $\tau : S_\lambda \mapsto S_{\tilde{\lambda}}$.*

Proof. In fact when we apply τ to the first determinantal formula for S_λ, we find the second determinantal formula for $S_{\tilde{\lambda}}$. □

9 Branching Rules for S_n, Standard Diagrams

9.1 Mumaghan's Rule

We wish to describe now a fairly simple recursive algorithm, due to Mumaghan, to compute the numbers $c_\lambda(\mu)$. 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 $\psi_k(x) S_\lambda(x) V(x) = \psi_k(x) A_{\lambda+\varrho}(x)$:

$$(9.1.1)\qquad \psi_k(x)A_{\lambda+\varrho}(x) = \left(\sum_{i=1}^{n} x_i^k\right)\left(\sum_{s\in S_n} \epsilon_s x_{s1}^{h_1+n-1} x_{s2}^{h_1+n-2} \dots x_{sn}^{h_n}\right).$$

Write $k_i = h_i + n - i$. We inspect the monomials appearing in the alternating function which is at the right of 9.1.1. 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 equal numbers 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 $\lambda' : h_1', \dots, h_i', \dots, h_n'$ associated to this sequence is

$$\begin{aligned} &h_t' = h_t, && \text{if} \quad t \leq i \quad \text{or} \quad t > j, \\ &h_t' = h_{t-1} + 1 && \text{if} \quad i + 2 \leq t \leq j, \ h_{i+1}' = h_j + k - j + i + 1. \end{aligned}$$

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

To understand the λ' which appear let us define the *rim* or boundary of a diagram λ as the set of points $(i, j) \in \lambda$ for which there is no point $(h, k) \in \lambda$ with $i < h$, $j < k$.

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

Notice that we have modified $j - i$ consecutive rows, adding a total of k new boxes. Each row of this set, except for the bottom row, has been replaced by the row immediately below it plus one extra box. We add the remaining boxes to the bottom row.

This property appears to be saying that the new diagram λ' is any diagram which contains the diagram λ and such that their difference is connected, made of k boxes of the rim of λ'. Intuitively it is like a *slinky*.[79] 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 is $+1$ if the number of rows occupied is odd, -1 otherwise. More formally we have:

Mumaghan's rule. *$\psi_k(x)S_\lambda(x) = \sum \pm S_{\lambda'}$, where λ' runs over all diagrams, such that by removing a connected set of k boxes of the rim of λ' we have λ.*

The sign to attribute to λ' is $+1$ if the number of rows modified from λ is odd, -1 otherwise.

For instance we can visualize $\psi_3 S_{321} = S_{321^4} - S_{32^3} - S_{3^3} - S_{4^21} + S_{621}$ as

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

[79] This was explained to me by A. Garsia and refers to a spring toy sold in novelty shops.

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

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 to the right. The sign of the slinky is -1 if it occupies an even number of rows, and $+1$ otherwise.

Next, one defines a *striped tableau* of type $\mu := k_1, k_2, \dots, k_t$ to be a tableau filled, for each $i = 1, \dots, t$, 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. For example, 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 a sign: the product of the signs of all its slinkies. In our case it is the sign pattern $- - + + - + ++$ for a total $-$ sign.

Mumaghan's rule can be formulated as:

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

Notice that when $\mu = 1^n$ the slinky is one box. The condition is that the diagram is filled with all the distinct numbers $1, \dots, n$. The filling is increasing from left to right and from the bottom to the top. Let us formalize:

Definition. A *standard tableau* of shape $\lambda \vdash n$ is a filling of a Young diagram with n-boxes of shape λ, with all the distinct numbers $1, \dots, n$. The filling is strictly increasing from left to right on each row and from the bottom to the top on each column.

From the previous discussion we have:

Theorem. *$d(\lambda)$ equals the number of standard tableaux of shape λ.*

Example. Standard tableaux of shape $3, 2$ (compute $d(\lambda) = 5$):

$$
\begin{array}{lll} 2 & 4 & \\ 1 & 3 & 5 \end{array}, \quad
\begin{array}{lll} 2 & 5 & \\ 1 & 3 & 4 \end{array}, \quad
\begin{array}{lll} 4 & 5 & \\ 1 & 2 & 3 \end{array}, \quad
\begin{array}{lll} 3 & 4 & \\ 1 & 2 & 5 \end{array}, \quad
\begin{array}{lll} 3 & 5 & \\ 1 & 2 & 4 \end{array}.
$$

9.2 Branching Rule for S_n

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

For a given partition $\mu \vdash n$, consider the module M_μ for S_n and the subgroup $S_{n-1} \subset S_n$ permuting the first $n-1$ numbers. We want to analyze M_μ as a representation of the subgroup S_{n-1}. For this we perform a character computation.

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

If $\mu \subset \lambda$ and $n = m + 1$ we will also say that μ, λ are *adjacent*,[80] in this case clearly λ is obtained from μ by removing a box lying in a corner.

With these remarks we notice a special case of Theorem 9.1:

$$\psi_1 S_\mu = \sum_{\lambda \vdash |\mu|+1, \mu \subset \lambda} S_\lambda. \tag{9.2.1}$$

Now consider an element of S_{n-1} to which is associated a partition ν. The same element, considered as a permutation in S_n, has associated the partition $\nu 1$. Computing characters we have

$$\begin{aligned}\sum_{\lambda \vdash n} c_\lambda(\nu 1) S_\lambda = \psi_{\nu 1} = \psi_1 \psi_\nu &= \sum_{\tau \vdash (n-1)} c_\tau(\nu) \psi_1 S_\tau \\ &= \sum_{\tau \vdash (n-1)} c_\tau(\nu) \sum_{\mu \vdash n,\, \tau \subset \mu} S_\mu.\end{aligned} \tag{9.2.2}$$

In other words

$$c_\lambda(\nu 1) = \sum_{\mu \vdash (n-1),\, \mu \subset \lambda} c_\mu(\nu), \tag{9.2.3}$$

This identity between characters becomes in module notation:

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

$$M_\lambda = \bigoplus\nolimits_{\mu \vdash (n-1),\, \mu \subset \lambda} M_\mu. \tag{9.2.4}$$

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 in 9.2.4 is unique.

A convenient way to record a partition $\mu \vdash n-1$ obtained from $\lambda \vdash n$ by removing a box is given by marking this box with n. We can repeat the branching to S_{n-2} and get

[80] Adjacency is a general notion in a poset; here the order is inclusion.

$$M_\lambda = \bigoplus_{\substack{\mu_2 \vdash n-2,\\ \mu_1 \vdash n-1,\\ \mu_2 \subset \mu_1 \subset \lambda}} M_{\mu_2}. \tag{9.2.5}$$

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

Example. From 4, 2, 1, 1, branching once:

$$\begin{array}{llll} \cdot & & & \\ \cdot & & & \\ \cdot & \cdot & & \\ \cdot & \cdot & \cdot & 8 \end{array} + \begin{array}{llll} \cdot & & & \\ \cdot & & & \\ \cdot & 8 & & \\ \cdot & \cdot & \cdot & \cdot \end{array} + \begin{array}{llll} 8 & & & \\ \cdot & & & \\ \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

and twice:

$$\begin{array}{llll} \cdot & & & \\ \cdot & & & \\ \cdot & \cdot & & \\ \cdot & \cdot & 7 & 8 \end{array} + \begin{array}{llll} \cdot & & & \\ \cdot & & & \\ \cdot & 7 & & \\ \cdot & \cdot & \cdot & 8 \end{array} + \begin{array}{llll} 7 & & & \\ \cdot & & & \\ \cdot & \cdot & & \\ \cdot & \cdot & \cdot & 8 \end{array} + \begin{array}{llll} \cdot & & & \\ \cdot & & & \\ \cdot & 8 & & \\ \cdot & \cdot & \cdot & 7 \end{array}$$

$$+ \begin{array}{llll} 7 & & & \\ \cdot & & & \\ \cdot & 8 & & \\ \cdot & \cdot & \cdot & \cdot \end{array} + \begin{array}{llll} 8 & & & \\ \cdot & & & \\ \cdot & \cdot & & \\ \cdot & \cdot & \cdot & 7 \end{array} + \begin{array}{llll} 8 & & & \\ \cdot & & & \\ \cdot & 7 & & \\ \cdot & \cdot & \cdot & \cdot \end{array} + \begin{array}{llll} 8 & & & \\ 7 & & & \\ \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

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 tableau* 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.

An example of a skew tableau of shape 6, 5, 2, 2/3, 2, 1:

$$\begin{array}{llllll} 6 & 7 & & & & \\ \cdot & 2 & & & & \\ \cdot & \cdot & 2 & 3 & 4 & \\ \cdot & \cdot & \cdot & 1 & 2 & 4 \end{array}$$

Notice that we have placed some dots in the position of the partition 3, 2, 1 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 $n, n-1, \dots, n-i+1$.

When we apply the branching 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 $n-k+1, n-k+2, \dots, n-1, n$.

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

Finally 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 section.

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 a 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 above, 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 nonisomorphic representations are necessarily orthogonal). If we work over $\mathbb{R}$, we can thus select a vector of norm 1 in each summand. This still leaves 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 inequivalent 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 $(i\ i+1)$, $i = 1, \dots, n-1$). We will discuss this question when we deal in a more systematic way with standard tableaux in Chapter 13.

10 Branching Rules for the Linear Group, Semistandard Diagrams

10.1 Branching Rule

When we deal with representations of the linear group we can use the character theory which identifies the Schur functions $S_\lambda(x_1, \dots, x_n)$ as the irreducible characters of $GL(n, \mathbb{C}) = GL(V)$. In general, the strategy is to interpret the various constructions on representations by corresponding operations on characters. There are two main ones: *branching* and *tensor product*. When we branch from $GL(n, \mathbb{C})$ to $GL(n-1, \mathbb{C})$ embedded as block $\begin{vmatrix} A & 0 \\ 0 & 1 \end{vmatrix}$ matrices, we can operate on characters by just setting $x_n = 1$ so the character of the restriction of $S_\lambda(V)$ to $GL(n-1, \mathbb{C})$ is

$$(10.1.1) \qquad S_\lambda(x_1, \dots, x_{n-1}, 1) = \sum c_\mu S_\mu(x_1, \dots, x_{n-1}).$$

Similarly, when we take two irreducible representations $S_\lambda(V),\ S_\mu(V)$ and form their tensor product $S_\lambda(V) \otimes S_\mu(V)$, its character is given by the symmetric function

$$(10.1.2) \qquad S_\lambda(x_1, \dots, x_n) S_\mu(x_1, \dots, x_n) = \sum_\nu c^\nu_{\lambda,\mu} S_\nu(x_1, \dots, x_n).$$

The coefficients in both formulas can be made explicit but, while in 10.1.1 the answer is fairly simple, 10.1.2 has a rather complicated answer given by the *Littlewood–Richardson rule* (discussed in Chapter 12, §5).

The reason why 10.1.1 is rather simple is that all the μ which appear actually appear with coefficient $c_\mu = 1$, so it is only necessary to explain which partitions appear. It is best to describe them geometrically by the diagrams.

WARNING For the linear group we will use *English notation*, for reasons that will be clearer in Chapter 13. Also assume that if $\lambda = h_1, h_2, \dots, h_r$, these numbers represent the *lengths of the columns*,[81] and hence r must be at most n (we assume $h_r > 0$). In 10.3 we will show that the conditions for μ to appear are the following.

1. $\mu = k_1, \dots, k_s$ is a diagram contained in λ, i.e., $s \le r,\ k_i \le h_i,\ \forall i \le s$.
2. $s \le n - 1$.
3. μ is obtained from λ by removing at most one box from each row.

The last condition means that we can remove only boxes at the end of each row, which form the *rim* of the diagram. It is convenient to mark the removed boxes by n.

For instance take $\lambda = 4, 2, 2, n = 5$ (we mark the rim). The possible 9 branchings are:

[81] Unfortunately the notation for Young diagrams is not coherent because in the literature they have arisen in different contexts, each having its notational needs.

$$\begin{array}{ccc} \cdot & \cdot & \bullet \\ \cdot & \cdot & \bullet \\ \bullet & & \\ \bullet & & \end{array} \implies$$

$$\begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & & \\ \cdot & & \end{array} \quad \begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & & \\ 5 & & \end{array} \quad \begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 5 & & \\ 5 & & \end{array} \quad \begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & 5 \\ \cdot & & \\ \cdot & & \end{array} \quad \begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & 5 \\ \cdot & & \\ 5 & & \end{array} \quad \begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & 5 \\ 5 & & \\ 5 & & \end{array}$$

$$\begin{array}{ccc} \cdot & \cdot & 5 \\ \cdot & \cdot & 5 \\ \cdot & & \\ \cdot & & \end{array} \quad \begin{array}{ccc} \cdot & \cdot & 5 \\ \cdot & \cdot & 5 \\ \cdot & & \\ 5 & & \end{array} \quad \begin{array}{ccc} \cdot & \cdot & 5 \\ \cdot & \cdot & 5 \\ 5 & & \\ 5 & & \end{array}$$

If we repeat these branchings and markings, we see that a sequence of branchings produces a *semistandard tableau* (cf. Chapter 12 §1.1 for a formal definition) like:

$$\begin{array}{ccc} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 5 & & \\ 5 & & \end{array}$$

As in the case of the symmetric group we can deduce a basis of the representation indexed by semistandard tableaux. Conversely, we shall see that one can start from such a basis and deduce a stronger branching theorem which is valid over the integers (Chapter 13, 5.4).

10.2 Pieri's Formula

Although we shall discuss the general Littlewood–Richardson rule in Chapter 12, we start with an example, the study of $S_\lambda(V) \otimes \bigwedge^i(V)$. By previous analysis this can be computed by computing the product $S_\lambda(x)e_i(x)$, where $e_i(x) = S_{1^i}(x)$ is the character of $\bigwedge^i(V)$. For this set $\lambda = h_1, h_2, \dots, h_r$ and $\{\lambda\}_i := \{\mu \mid \mu \supset \lambda,\ |\mu| = |\lambda| + i$ and each column k_i of μ satisfies $h_i \le k_i \le h_i + 1\}$.

Theorem (Pieri's formula).

$$S_\lambda(x)e_i(x) = \sum_{\mu \in \{\lambda\}_i} S_\mu(x), \quad S_\lambda(V) \otimes \bigwedge^i(V) = \bigoplus\nolimits_{\mu \in \{\lambda\}_i} S_\mu(V). \tag{10.2.1}$$

Proof. Let $\lambda = h_1, h_2, \dots, h_n$ where we take n sufficiently large and allow some h_i to be 0, and multiply $S_\lambda(x)e_i(x)V(x) = A_{\lambda+\rho}(x)e_i(x)$. We must decompose the alternating function $A_{\lambda+\rho}(x)e_i(x)$ in terms of functions $A_{\mu+\rho}(x)$. Let $l_i = h_i + n - i$. The only way to obtain in $A_{\lambda+\rho}(x)e_i(x)$ a monomial $x_1^{m_1}x_2^{m_2}\dots x_n^{m_n}$ with $m_1 > m_2 > \dots > m_n$ is possibly by multiplying $x_1^{l_1}x_2^{l_2}\dots x_n^{l_n}x_{j_1}x_{j_2}\dots x_{j_i}$. This monomial has strictly decreasing exponents for the variables $x_1, \dots, x_n$ if and only if the following condition is satisfied. Set $k_a = h_a$ if a does not appear in the indices $j_1, j_2, \dots, j_i$, and $k_a = h_a + 1$ otherwise. We must have that $k_1 \geq k_2 \cdots \geq k_n$, in other words $\mu := k_1 \geq k_2 \cdots \geq k_n$ is a diagram in $\{\lambda\}_i$. The coefficient of such a monomial is 1, hence we deduce the claim

$$A_{\lambda+\rho}(x)e_i(x) = \sum_{\mu \in \{\lambda\}_i} A_{\mu+\rho}(x). \qquad \square$$

We may now deduce also by duality, using the involutory map $\tau : e_i \to h_i$ (cf. 8.3) and the fact that $h_i(x) = S_i(x)$ is the character of $S^i(V)$, the formula

$$(10.2.2) \qquad S_\lambda(x)h_i(x) = \sum_{\tilde{\mu} \in \{\tilde{\lambda}\}_i} S_\mu(x), \quad S_\lambda(V) \otimes S^i(V) = \bigoplus\nolimits_{\tilde{\mu} \in \{\tilde{\lambda}\}_i} S_\mu(V).$$

In other words, when we perform $S_\lambda(V) \otimes \bigwedge^i(V)$ we get a sum of $S_\mu(V)$ where μ runs over all diagrams obtained from λ by adding i boxes and at most one box in each column, while when we perform $S_\lambda(V) \otimes S^i(V)$ we get a sum of $S_\mu(V)$ where μ runs over all diagrams obtained from λ by adding i boxes and at most one box in each row.[82]

Recall that, for the linear group, we have exchanged rows and columns in our conventions.

10.3 Proof of the Rule

We can now discuss the branching rule from $GL(n, \mathbb{C})$ to $GL(n-1, \mathbb{C})$. From the point of view of characters it is clear that if $f(x_1, \dots, x_n)$ is the character of a representation of $GL(n, \mathbb{C})$, $f(x_1, \dots, x_{n-1}, 1)$ is the character of the restriction of the representation to $GL(n-1, \mathbb{C})$. We thus want to compute $S_\lambda(x_1, \dots, x_{n-1}, 1)$. For this we use Cauchy's formula, getting

$$\begin{aligned}
&\sum_\lambda S_\lambda(x_1, \dots, x_{n-1}, 1) S_\lambda(y_1, \dots, y_{n-1}, y_n) \\
&\quad = \prod_{i=1,\ j=1}^{n-1,\ n} \frac{1}{1 - x_i y_j} \prod_{j=1}^{n} \frac{1}{1 - y_j} \\
&\quad = \sum_\mu S_\mu(x_1, \dots, x_{n-1}) S_\mu(y_1, \dots, y_{n-1}, y_n) \sum_{j=0}^{\infty} h_j(y_1, \dots, y_{n-1}, y_n).
\end{aligned}$$

Use 10.2.2 to get

$$\begin{aligned}
&\sum_\lambda S_\lambda(x_1, \dots, x_{n-1}, 1) S_\lambda(y_1, \dots, y_{n-1}, y_n) \\
&\quad = \sum_\mu S_\mu(x_1, \dots, x_{n-1}) \sum_{j=0}^{\infty} \sum_{\tilde{\lambda} \in \{\tilde{\mu}\}_j} S_\lambda(y_1, \dots, y_{n-1}, y_n).
\end{aligned}$$

[82] These two rules are sometimes referred to as Pieri's rule.

Comparing the coefficients of $S_\lambda(y_1, \dots, y_{n-1}, y_n)$, we obtain

$$S_\lambda(x_1, \dots, x_{n-1}, 1) = \sum_{\mu \mid \tilde{\lambda} \in \{\tilde{\mu}\}_j} S_\mu(x_1, \dots, x_{n-1}).$$

In other words, let $\{\lambda\}^j$ be the set of diagrams which are obtained from λ by removing j boxes and at most one box in each row. Then the branching of $S_\lambda(\mathbb{C}^n)$ to $GL(n-1)$ is $\bigoplus_{\mu \in \{\lambda\}^j} S_\mu(\mathbb{C}^{n-1})$. In particular we have the property that the irreducible representations which appear come with multiplicity 1.

Since $S_\lambda(x_1, \dots, x_{n-1}, x_n)$ is homogeneous of degree $|\lambda|$ while $\mu \in \{\lambda\}^j$ is homogeneous of degree $|\lambda| - j$, we must have

$$S_\lambda(x_1, \dots, x_{n-1}, x_n) = \sum_{j=0}^{|\lambda|} x_n^j \sum_{\mu \in \{\lambda\}^j} S_\mu(x_1, \dots, x_{n-1}). \qquad \square$$

We may iterate the branching. At each step the branching to $GL(n-i)$ is a direct sum of representations $S_\mu(\mathbb{C}^{n-1})$ with indexing a sequence of diagrams $\mu = \mu_i \subset \mu_{i-1} \subset \cdots \subset \mu_0 = \lambda$ where each μ_j is obtained from μ_{j-1} by removing u_j boxes and at most one box in each row. Furthermore we must have $ht(\mu_j) \leq n - j$. Correspondingly,

$$S_\lambda(x_1, \dots, x_{n-1}, x_n) = \sum_{\mu=\mu_i \subset \mu_{i-1} \subset \cdots \subset \mu_0 = \lambda} x_n^{u_1} x_{n-1}^{u_2} \dots x_{n-i+1}^{u_i} S_\mu(x_1, \dots, x_{n-j}).$$

If we continue the branching all the way to 1, we decompose the space $S_\lambda(V)$ into 1-dimensional subspaces which are weight vectors for the diagonal matrices. Each such weight vector is indexed by a complete *flag* of subdiagrams

$$\emptyset = \mu_n \subset \mu_1 \subset \dots \mu_i \subset \cdots \subset \mu_0 = \lambda$$

and weight $\prod_{i=1}^n x_i^{u_{n-i+1}}$.

A convenient way to encode such flags of subdiagrams is by filling the diagram λ as a semistandard tableau, placing $n - i$ in all the boxes of μ_i not in μ_{i-1}. The restriction we have placed implies that all the rows are strictly increasing, since we remove at most one box from each row, while the columns are weakly increasing, since we may remove more than one box at each step but we fill the columns with a strictly decreasing sequence of numbers. Thus we get a semistandard tableau T of (column-) shape λ filled with the numbers $1, 2, \dots, n$. Conversely, such a semistandard tableau corresponds to an allowed sequence of subdiagrams $\emptyset = \mu_n \subset \mu_1 \subset \dots \mu_i \subset \cdots \subset \mu_0 = \lambda$. Then the monomial $\prod_{i=1}^n x_i^{u_{n-i+1}}$ is deduced directly from T, since u_{n-i+1} is the number of appearances of i in the tableau.

We set $x^T := \prod_{i=1}^n x_i^{u_{n-i+1}}$ and call it the *weight of the tableau* T. Finally we have:

Theorem.

$$(10.3.1) \qquad S_\lambda(x_1, \dots, x_{n-1}, x_n) = \sum_{T \text{ semistandard of shape } \lambda} x^T.$$

Of course the set of semistandard tableaux depends on the set of numbers $1, \dots, n$. Since the rows have to be filled by strictly increasing numbers we must have a restriction on height. The rows have at most n-elements.

Example. $S_{3,2,2}(x_1, x_2, x_3)$ is obtained from the tableaux:

$$\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 3 & \\ 2 & 3 & \end{array} \quad \begin{array}{ccc} 1 & 2 & 3 \\ 1 & 3 & \\ 2 & 3 & \end{array} \quad \begin{array}{ccc} 1 & 2 & 3 \\ 1 & 2 & \\ 2 & 3 & \end{array} \quad \begin{array}{ccc} 1 & 2 & 3 \\ 1 & 3 & \\ 1 & 3 & \end{array} \quad \begin{array}{ccc} 1 & 2 & 3 \\ 1 & 2 & \\ 1 & 3 & \end{array} \quad \begin{array}{ccc} 1 & 2 & 3 \\ 1 & 2 & \\ 1 & 2 & \end{array}$$

$$S_{3,2,2}(x_1, x_2, x_3) = x_1 x_2^3 x_3^3 + x_1^2 x_2^2 x_3^3 + x_1^2 x_2^3 x_3^2 + x_1^3 x_2 x_3^3 + x_1^3 x_2^2 x_3^2 + x_1^3 x_2^3 x_3.$$

Of course if we increase the number of variables, then also the number and types of monomials will increase.

We may apply at the same time the branching rule for the symmetric and the linear group. We take an n-dimensional vector space V and consider

$$V^{\otimes m} = \bigoplus\nolimits_{\lambda \vdash m} M_\lambda \otimes S_\lambda(V).$$

When we branch on both sides we decompose $V^{\otimes m}$ into a direct sum of 1-dimensional weight spaces indexed by pairs $T_1 \mid T_2$ where T_1 is a standard diagram of shape $\lambda \vdash m$ and T_2 is a semistandard diagram of shape λ filled with $1, 2, \dots, n$. We will see, in Chapter 12, §1, that this construction of a basis has a purely combinatorial counterpart, the Robinson–Schensted correspondence.

Note that from Theorem 10.3.1 it is not evident that the function $S_\lambda(x_1, \dots, x_{n-1}, x_n)$ is even symmetric. Nevertheless there is a purely combinatorial approach to Schur functions which takes Theorem 10.3.1 as definition. In this approach the proof of the symmetry of the formula is done by a simple *marriage* argument.