III TENSOR SYMMETRY

November 1995

Contents

§22. Symmetry in tensor spaces

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

22.1 We have seen in 1.1.6 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 2 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 $U \times V^*$.

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

$$< f|u \otimes \varphi >=< \varphi|f(u) > .$$

We have thus a correspondence between intertwiners and invariants.

We will find it particularly useful, according to the Arhonold method, to use it 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:

(22.1.1) $$End(A^{\otimes m}) = End(A)^{\otimes m} = A^{*\otimes m} \otimes A^{\otimes m} = (A^{*\otimes m} \otimes A^{\otimes m})^*.$$

Thus in this case

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 2 group actions, one given by the linear group $GL(V)$ by the formula:

(22.1.2) $$g(v_1 \otimes v_2 \otimes \ldots \otimes v_n) := gv_1 \otimes gv_2 \otimes \ldots \otimes gv_n,$$

and the other of the symmetric group S_n given by:

(22.1.3) $$\sigma(v_1 \otimes v_2 \otimes \ldots \otimes v_n) = v_{\sigma^{-1}1} \otimes v_{\sigma^{-1}2} \otimes \ldots \otimes v_{\sigma^{-1}n}.$$

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

Before we make any further analysis of these actions we want to discuss symmetric tensors as described in 11.8. Fix a basis $e_1, e_2, \ldots, e_m$ in V.

Definition. A tensor $u \in V^{\otimes n}$ is *symmetric* if $\sigma(u) = u, \ \forall \sigma \in S_n$.

Given a vector $v \in V$ the tensor $v^n = v \otimes v \otimes v \ldots \otimes v$ is clearly symmetric.

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

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

If $\underline{h} := h_1, h_2, \ldots, h_m$ is such a sequence we will denote by $e_{\underline{h}}$ be the sum of elements in the corresponding orbit.

Notice that the image of the symmetric tensor $e_{\underline{h}}$ in the symmetric algebra is

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

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

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

Consider a linear function ϕ on the space of symmetric tensors defined by $< \phi|e_{\underline{h}} >= a_{\underline{h}}$ and compute it on the tensors v^n:

$$< \phi|(\sum_k x_k e_k)^n >= \sum_{h_1+h_2+\ldots+h_m=n} x_1^{h_1} x_2^{h_2} \ldots x_m^{h_m} a_{\underline{h}}.$$

Thus we see that the dual of the of the space of symmetric tensors is identified with the space of homogeneous polynomials.

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 polynomials on X. Since X is Zariski dense this polynomial vanishes if and only if the form is 0 hence the given vectors span the space of symmetric tensors. $\square$

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

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

Our aim is to prove:

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

Proof. We start by identifying:

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

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

$$A_1 \otimes A_2 \otimes \ldots \otimes A_n(v_1 \otimes v_2 \otimes \ldots \otimes v_n) = A_1 v_1 \otimes A_2 v_2 \otimes \ldots \otimes A_n v_n.$$

Thus, if $g \in GL(V)$, the corresponding operator in $V^{\otimes n}$ is $g \otimes g \otimes \ldots \otimes g$. From the Lemma 22.1 it follows that the algebra A coincides with the symmetric tensors in $End(V)^{\otimes n}$, since $GL(V)$ is Zariski dense.

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

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

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

$$\sigma A_1 \otimes A_2 \otimes \ldots \otimes A_n \sigma^{-1}(v_1 \otimes v_2 \otimes \ldots \otimes v_n) = \sigma A_1 \otimes A_2 \otimes \ldots \otimes A_n(v_{\sigma 1} \otimes v_{\sigma 2} \otimes \ldots \otimes v_{\sigma n}) =$$
$$\sigma(A_1 v_{\sigma 1} \otimes A_2 v_{\sigma 2} \otimes \ldots \otimes A_n v_{\sigma n}) = A_{\sigma^{-1}1} v_1 \otimes A_{\sigma^{-1}2} v_2 \ldots A_{\sigma^{-1}n} v_n =$$
$$(A_{\sigma^{-1}1} \otimes A_{\sigma^{-1}2} \ldots A_{\sigma^{-1}n})(v_1 \otimes v_2 \otimes \ldots \otimes v_n).$$

This computation shows that the conjugation action is in fact the symmetry action and finishes the proof. $\square$

We draw now 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 Matcshke's theorem, therefore, by the theory of semisimple algebras all statements follow from the previous Theorem which states that A is the centralizer of B. $\square$

22.3 We want to formulate the same theorem 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 ($A \in hom(V, W)$, $\alpha \in W^*$, $v \in V$):

$$< \alpha | Av > . \tag{22.3.1}$$

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 $< \alpha | Av >$ 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 to the space of multilinear functions of n *covector and* n *vector variables* $f(\alpha_1, \alpha_2, \ldots, \alpha_n, v_1, v_2, \ldots, v_n)$ 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 22.3.1 evaluated on decomposable tensors we get

$$f_\sigma(\alpha_1, \alpha_2, \ldots, \alpha_n, v_1, v_2, \ldots, v_n) =< \alpha_1 \otimes \alpha_2 \otimes \ldots \otimes \alpha_n | \sigma(v_1 \otimes v_2 \otimes \ldots \otimes v_n) >=$$
$$< \alpha_1 \otimes \alpha_2 \otimes \ldots \otimes \alpha_n | v_{\sigma^{-1}1} \otimes v_{\sigma^{-1}2} \otimes \ldots \otimes v_{\sigma^{-1}n} >= \prod_{i=1}^n < \alpha_i | v_{\sigma^{-1}i} >= \prod_{i=1}^n < \alpha_{\sigma i} | v_i >$$

We can thus deduce:

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

$$f_\sigma(\alpha_1, \alpha_2, \ldots, \alpha_n, v_1, v_2, \ldots, v_n) := \prod_{i=1}^n < \alpha_{\sigma i} | v_i > . \tag{22.3.2}$$

22.4 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 the next chapter.

We want to drop now the restriction of being multilinear in the invariants.

Take the space $(V^*)^p \times V^q$ of p covector and q vector variables as 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 $< \alpha_i | v_j >$ which are clearly $GL(V)$ invariant, we prove:

First fundamental theorem for the linear group.
The ring of polynomial functions on $V^{*p} \times V^q$ which are $GL(V)$ invariant is generated by the functions $< \alpha_i | v_j >$.

Before starting to prove this theorem we want to make some remarks on 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^* to the space of m dimensional row vectors.

The group $GL(V)$ is then identified to 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 $< \alpha_i | v_j >$ is then identified to 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 to 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 to the space of pairs:

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

The action of the matrix group is by:

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

Consider the multiplication map:

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

the entries of the matrix XY are the basic invariants $< \alpha_i | v_j >$ thus the theorem can also be formulated as:

Theorem. *The ring of polynomial functions on $M_{p,m} \times M_{m,q}$ which 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 $g(\alpha_1, \alpha_2, \ldots, \alpha_p, v_1, v_2, \ldots, v_q)$ 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 $\overline{g}(\alpha_1, \alpha_2, \ldots, \alpha_N, v_1, v_2, \ldots, v_M)$ where N, 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 in λv and α in $\lambda^{-1}\alpha$ 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 22.3 and deduce that $\overline{g}$ is a linear combination of functions of type $\prod_{i=1}^{N} < \alpha_{\sigma i} | v_i >$.

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

The study of the relations among invariants will be the topic of the **second fundamental theorem**. At this moment let us only remark that, by elementary linear algebra, the multiplication map f has, as image, the subvariety 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 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 $M_{p,q}$.

It will be the content of the second fundamental theorem to prove that these determinants generate a prime ideal which is thus the full ideal of relations among the invariants $< \alpha_i | v_j >$.

§23 Young symmetrizers

23.1 Primitive idempotents We return to the representation theory of the symmetric group and present Young theory of the symmetrizers.

First a generality on algebras. An idempotent e, in an algebra R, is an element such that $e^2 = e$, two idempotents e, f are *orthogonal* if $ef = fe = 0$ in this case $e + f$ is also an idempotent.

An idempotent is called *primitive* if it cannot be decomposed as a sum $e = e_1 + e_2$ of two non zero orthogonal idempotents.

If $R = M_k(F)$ is a matrix algebra over a field F it is easily verified that a primitive idempotent $e \in R$ is just an idempotent matrix of rank 1, which in a suitable basis, can be identified with the elementary matrix $e_{1,1}$.

In this case, the left ideal Re is formed by all matrices with 0 on the columns different from the first one. As an R module it is irreducible and isomorphic to F^k.

If an algebra $R = R_1 \oplus R_2$ is the direct sum of two algebras every idempotent in R is the sum $e_1 + e_2$ of two orthogonal idempotents $e_i \in R_i$. In particular the primitive idempotents in R are the primitive idempotents in R_1 and in R_2.

Thus, if $R = \sum_i M_{n_i}(F)$ is semisimple a primitive idempotent $e \in R$ is just a primitive idempotent in one of the summands $M_{n_i}(F)$.

$M_{n_i}(F) = ReR$ while Re is irreducible as R module (and isomorphic to the module F^{n_i} for the summand $M_{n_i}(F)$).

Conversely, let $R = \oplus R_i$ be a general semisimple algebra over a field F with R_i simple and isomorphic to the algebra of matrices over some division algebra D_i, if an idempotent $e \in R$ is such that $dim_F eRe = 1$ then it is easily seen that $e \in R_i$ for some index i and the corresponding division algebra D_i reduces to F.

Lemma. *A sufficient condition that $e \in R$ is primitive is that $dim_F eRe = 1$. In this case ReR is a matrix algebra over F.*

Proof. From the previous discussion. □

23.2 Young diagrams and symmetrizers We discuss now the symmetric group. The theory of cycles (cf. 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 $[1, 2, \ldots, n]$ and these are parametrized by partitions of n.

We shall express that $\mu := k_1, k_2, \ldots, k_n$ is a partition of n by the symbol $\mu \vdash n$.

We shall denote by C_μ the conjugacy class in S_n formed by the permutations decomposed in cycles of length $k_1, k_2, \ldots, k_n$.

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 flavour. We will exhibit a decomposition as a sum of matrix algebras:

(23.2.1)
$$R = \mathbb{Q}[S_n] := \sum_{\mu \vdash n} M_{n(\mu)}(\mathbb{Q}),$$

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

From the theory of group characters we know 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 $R = \oplus_{\lambda \vdash n} Re_\lambda R$ and the left ideal Re_λ as an R module is an irreducible representation.

In this way we will exhaust all irreducible representations.

By the above remarks it will be enough to construct idempotents e_λ so that $dim_{\mathbb{Q}} e_\lambda Re_\lambda = 1$ and $e_\lambda Re_\mu = 0$ if $\lambda \neq \mu$. This will also prove 23.2.1 by an easy argument which we leave to the reader.[1]

[1] In fact the construction of the elements e_λ is not completely canonical and, as we shall see, we will subject it to some partly arbitrary choices.

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.

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

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

Thre 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 as points in the fourth:

Example: partition 4311:

$$\text{French}\quad \begin{array}{llll} \cdot & & & \\ \cdot & & & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot \end{array} \qquad\qquad \text{English}\quad \begin{array}{llll} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \\ \cdot & & & \\ \cdot & & & \end{array}$$

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.

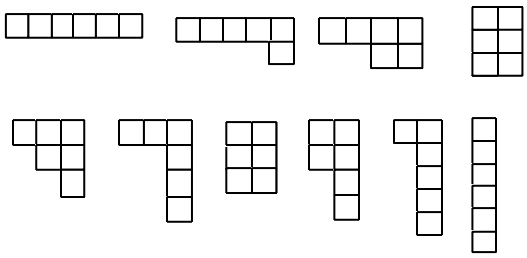
Given a Young diagram D (in French form) the set $C_i := \{(i, j) \in D\}$, i fixed, will be called the i^{th} column, the set $R_j := \{(i, j) \in D,\ j\}$ fixed, will be called the j^{th} row.

The lenghts $k_1, k_2, k_3, \ldots$ of the rows are a decreasing sequence of numbers which determine completely 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 is dual is an involutory map which geometrically can be visualized as flipping the Ferrer's 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, english display):



Definition. 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: partition 4311[2]:

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

The symmetric group S_n acts on 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* defined by:

i, j are in the same part if they appear in the same row of T.

Similarly for the *column partition*.

To a partition π of $(1, 2, 3, \dots, n-1, n)$[3] 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 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.

[2] the reader will notice the peculiar properties of the right tableau which we will encounter over and over in the future

[3] 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, whle a partition of a set is in fact a decomposition of the set into dijoint parts.

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

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

We define

(23.2.1)
$$s_T = \sum_{\sigma \in \mathrm{R}_T} \sigma \qquad \text{the } \textit{symmetrizer on the rows} \text{ and}$$
$$a_T = \sum_{\sigma \in \mathrm{C}_T} \epsilon_\sigma \sigma \quad \text{the } \textit{antisymmetrizer on the columns.}$$

It is trivial to verify the following two identities:

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

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

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

(23.2.2)
$$ps_T = s_T = s_T p, \; \forall p \in \mathrm{R}_T; \; qa_T = a_T q = \epsilon_q a_T, \; \forall q \in \mathrm{C}_T.$$

It is then an easy exercise to check

Proposition. *The left ideal Rs_T has as basis the elements gs_T as g runs over a set representatives of the cosets $g\mathrm{R}_T$ and it equals, as representation, the permutation representation on such cosets.*

The left ideal Ra_T has as basis the elements ga_T as g runs over a set representatives of the cosets $g\mathrm{C}_T$ and it equals, as representation, the representation induced to S_n by the sign representation of C_T.

Now the remarkable fact comes, consider the product:

(23.2.3)
$$c_T := s_T a_T = \sum_{p \in \mathrm{R}_T, \; q \in \mathrm{C}_T} \epsilon_q pq$$

we will show that there exists a positive integer $p(T)$ such that the element $e_T := \frac{c_T}{p(T)}$ is an idempotent.

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

Remark

(23.2.4)
$$c_{sT} = sc_T s^{-1}.$$

We have thus, for a given $\lambda \vdash n$ several conjugate idempotents which we will show to be primitive and which will generate an irreducible module associated to λ.

For the moment let us remark that, from 23.2.4 follows that the integer $p(T)$ depends only on the shape λ of T and thus we will denote it by $p(T) = p(\lambda)$.

23.3 The main Lemma The main property of the element c_T which we will exploit is the following, clear from its definition and 23.2.2.

(23.3.1) $$pc_T = c_T,\ \forall p \in \mathrm{R}_T;\ c_T q = \epsilon_q c_T,\ \forall q \in \mathrm{C}_T.$$

We need a fundamental combinatorial lemma, consider the partitions of n as decreasing sequences of integers (including 0) and order them lexicographically.[4]

E.g. the partitions of 6 in increasing lexicographic order:

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

Lemma. *Let S, 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 hold:

i) There are two numbers i, j which are on 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 pigeon hole 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).

Remark 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, T and procede 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, this is impossible if $pS = qT$. $\square$

[4]We often drop 0 in the display.

Corollary. *i) Given $\lambda > \mu$ partitions, S, 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 $\mathrm{R}_T\mathrm{C}_T$ then there exists a transposition $u \in \mathrm{R}_T$ and a transposition $v \in \mathrm{C}_T$ such that $us = sv$.

Proof. i) From the previous lemma there are two numbers i, j in the same row for S and in the same column for sT. If $u = (i, j)$ is the corresponding transposition we set $v := s^{-1}us \in \mathrm{C}_T$ (from 23.2.3) these two transpositions satisfy the desired conditions.

ii) The proof is similar, we consider again a tableau T, construct sT and apply the Lemma to sT, T. $\square$

23.4 We draw now the conclusions relative to Young symmetrizers.

Proposition. *i) Let S, T be two tableaux of shapes $\lambda > \mu$.*

If an element a in the group algebra is such that:

$$pa = a, \ \forall p \in \mathrm{R}_S, \ \textit{and } aq = \epsilon_q a, \ \forall q \in \mathrm{C}_T$$

then $a = 0$.

ii) Given a tableau T and an element a in the group algebra such that:

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

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) Same type of proof. First we can say that, if $s \notin \mathrm{R}_T\mathrm{C}_T$ then $a(s) = 0$. Otherwise let $s = pq, \ p \in \mathrm{R}_T, q \in \mathrm{C}_T$ then $a(pq) = \epsilon_q a(1)$ hence $a = a(1)c_T$. $\square$

Before we conclude let us recall some simple facts on 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 by right (or left) action. Let us define $tr(r)$ to be the trace of the operator $x \to xr$. Clearly $tr(1) = dim_F R$. For a group algebra $F[G]$ of a finite group G the elements $g \in G$ give rise to permutations $x \to xg, \ x \in G$ of the basis elements in G and clearly then: $tr(1) = |G|, \ tr(g) = 0$ if $g \neq 0$.

We are now ready to conclude, the theorem which we aim at is:

Theorem. *i) $c_T^2 = p(\lambda)c_T$ with $p(T) \neq 0$ a positive integer.*
ii) $dim_{\mathbb{Q}} c_T R c_T = 1 = dim_{\mathbb{Q}} s_T R a_T$.
iii) If U, V are tableaux of different shapes λ, μ we have $c_U R c_V = 0 = s_U R a_V$.

iv) $dim_{\mathbb{Q}} Rc_T = \frac{n!}{p(\lambda)}$.

Proof. We apply the previous proposition and get that every element of $c_T R c_T$ satisfies ii) of that proposition hence $c_T R c_T = \mathbb{Q} c_T$ in particular we have $c_T^2 = p(\lambda) c_T$.

Now compute the trace of c_T, from the formula 23.2.1 we have $tr(c_T) = n!$ but since $c_T^2 = p(\lambda) c_T$ we have (by elementary matrix theory) that $tr(c_T) = p(\lambda) dim_{\mathbb{Q}} Rc_T$ and so not only we have $p(\lambda) \neq 0$ but also iv).

Next we prove iii), if $\lambda > \mu$ we have by part i) of the same proposition $s_U R a_V = 0$.

As for $c_U R c_V$ we make the following remark:

For a semisimple algebra R, if $aRb = 0$ then $bRa = 0$ in fact from $aRb = 0$ we deduce $(RbRaR)^2 = 0$.

Since $RbRaR$ is an ideal of R and we know that any non zero ideal in R is generated by an idempotent, we must have $RbRaR = 0$ hence the claim. □

Corollary. *The elements* $e_T := \frac{c_T}{p(\lambda)}$ *are primitive idempotents in* $R = \mathbb{Q}[S_n]$. *The left ideals* Re_T *give all the irreducible representations of* S_n *explicitely indexed by partitions. These representations are defined over* $\mathbb{Q}$.

We will indicate by M_λ the irreducible representation thus 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, but to two different labelings we obtain conjugate Young symmetrizers which therefore correspond to the same irreducible representation.

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 also we can argue 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 $\frac{1}{p(T)} a_T s_T$ is a primitive idempotent.

Since clearly $c_T a_T s_T = s_T a_T a_T s_T$ is non zero (a_T^2 is a non zero multiple of a_T and so $(c_T a_T s_T) a_T$ is a non zero multiple of c_T^2) we get that e_T and $\tau(e_T)$ are primitive idempotents relative to the same irreducible representation and the claim is proved.

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

Consider the two left ideals Rs_T, Ra_T, we have given a first description of their structure as representations in 23.2. They contain respectively $a_T R s_T$, $s_T R a_T$ whch are both 1 dimensional.

Thus we have

Lemma. *M_λ appears in its isotypic component in Rs_T, resp Ra_T with multiplicity 1. M_μ appears in Rs_T if and only if $\mu \geq \lambda$ and it appears in Ra_T if and only if $\mu \leq \lambda$.*

Proof. To see the multiplicity with which $M_\mu$appears in a representation V it suffices to compute the dimension of $c_T V$or of $\overline{c}_T V$. Therefore the statement follows from the previous results. $\square$

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

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 partititon $\tilde{\lambda}$ which in general is different from λ. It is thus natural to ask in which way are the two representations 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 $\varrho\tau$ equals 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 thus use a tableau T of shape λ and construct the symmetrizer, we have

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

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

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

thus $\tau(c_T) = a_{\tilde{T}} s_{\tilde{T}}$. $\square$

§24 The irreducible representations of the linear group 1

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

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

Thus, to get a list of the irreducible representations of the linear group $Gl(V)$ appearing in $V^{\otimes n}$, we may apply to tensor space the Young symmetrizers e_T.

We wish to discuss now the problem, when is $e_T V^{\otimes n} \neq 0$.

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 notations, that the first n_1 indeces appear in the first column in increasing order, the next n_2 indeces in the second column and so on, so that:

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

$$a_T V^{\otimes n} = a_{n_1} V^{\otimes n_1} \otimes a_{n_2} V^{\otimes n_2} \otimes \dots \otimes a_{n_t} V^{\otimes n_t} = \wedge^{n_1} V \otimes \wedge^{n_2} V \otimes \dots \otimes \wedge^{n_t} V.$$

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

Otherwise let $e_1, e_2, \dots, e_m$ be a basis of V and use for $V^{\otimes n}$ the corresponding basis of decomposable tensors; let us consider the tensor:

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

This is the decomposable tensor having e_i in the positions relative to the indeces of the i^{th} row. By construction it is symmetric with respect to the group R_T of row preserving permutations.

The tensor $a_T U$ is a sum of U and some other tensors in the product basis different from U.

When we apply to a tensor in the basis, different from U, any element of the group R_T we can never obtain U since this vector is fixed by R_T and the symmetric group permutes the basis.

Thus $c_T U = s_T a_T U = rU + U'$ where, U' is a sum of elements from the basis different from U and $r \neq 0$ is the order of the group R_T. In particular:

$$c_T U \neq 0.$$

Let us summarize using a definition.

Definition. *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 $ht(\lambda) > dim(V)$.*

We thus have as consequence a description of $V^{\otimes n}$ as representation of $S_n \times GL(V)$, define (up to isomorphism)

$$S_\lambda(V) := e_T V^{\otimes n}$$

for a tableau of shape λ we have:

Theorem.

$$V^{\otimes n} = \oplus_{ht(\lambda)\leq dim(V)} M_\lambda \otimes S_\lambda(V) \tag{24.1.1}$$

§25 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, computes first 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 procedes 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 (ins).

25.1 Up to now we have been able to explicitely parametrize both the conjugacy classes and the irreducible representations of S_n by partitions of n. Thus the computation of the character table consists, given two partitions λ, μ, to compute the value of the character of an element of the conjugacy class C_μ on the irreducible representation M_λ.

Let us denote by $\chi_\lambda(\mu)$ this value.

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

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

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

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

$$\psi_\mu(x) = \sum_\lambda c_\lambda(\mu) S_\lambda(x). \tag{25.1.1}$$

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

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

and have.

Theorem Frobenius. *For all partitions λ, $\mu \vdash n$ we have*

$$\chi_\lambda(\mu) = c_\lambda(\mu) \tag{25.1.2}$$

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

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

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

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

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

proved in 7.1.

We change its right hand side as follows. Compute:

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

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

Taking the exponential we get the following expression:

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

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

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

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

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

Let us indicate by

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

then 25.1.4 becomes

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

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

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

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

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

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

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

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

Let us now substitute in the identity:

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

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

(25.2.1)
$$\sum_{\mu \vdash n} \frac{1}{n(\mu)} c_{\lambda_1}(\mu) c_{\lambda_2}(\mu) = \begin{cases} 0 & \text{if} \quad \lambda_1 \neq \lambda_2 \\ 1 & \text{if} \quad \lambda_1 = \lambda_2 \end{cases}.$$

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

25.3 *Step 2* We consider now some permutation characters (cf. 23.2). Take a partition $\lambda := h_1, h_2, \ldots, h_k$ of n. Consider then the subgroup $S_{h_1} \times S_{h_2} \times \ldots \times S_{h_k}$ and the permutation representation on:

(25.3.1)
$$S_n / S_{h_1} \times S_{h_2} \times \ldots \times S_{h_k}$$

we will indicate by β_λ the corresponding character.

In order to compute this character let us make a general remark; consider a permutation representation associated to a coset space G/H, let χ denote its character and $g \in G$.

The character $\chi(g)$, i.e. the trace of the permutation induced by g equals the number of fixed points of g on G/H, i.e. the number of cosets xH such that $gxH = xH$.

Now $gxH = xH$ if and only if $x^{-1} g x \in H$, so let $X := \{x \in G | x^{-1} g x \in H\}$, we have $\chi(g) = \frac{|X|}{|H|}$.

The set X is a union of right cosets $G(g)x$ where $G(g)$ is the centralizer of g and the map $x \to x^{-1} g x$ is a bijection between the set of such cosets and the intersection of the conjugacy class C_g of g with H, we have shown:

Proposition. *The number of fixed points of g on G/H equals:*

(25.3.2)
$$\chi(g) = \frac{|C_g \cap H| |G(g)|}{|H|}.$$

Remark In practice we will compute $|C_g \cap H|$ by exhibiting the H conjugacy classes in which it splits. If then $C_g \cap H = \cup_i O_i$ fix an element $g_i \in O_i$ and let $H(g_i)$ be the centralizer of g_i in H, then $|O_i| = |H|/|H(g_i)|$ and finally.

$$\chi(g) = \sum_i \frac{|G(g)|}{|H(g_i)|}. \tag{25.3.3}$$

We compute it now for the case $G/H = S_n/S_{h_1} \times S_{h_2} \times \ldots \times S_{h_k}$ and for a permutation g relative to a partition $\mu := 1^{p_1} 2^{p_2} 3^{p_3} \ldots i^{p_i} \ldots$.

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

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

and:

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

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

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

and remark that, with the notations of 25.1.7 $b(\lambda + \mu) = b(\lambda) b(\mu)$.

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

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

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

Now

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

So we get:

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

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

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

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

In fact when we expand

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

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

For each such monomial denote by p_{ij} the number of choices of the term x_j^i in the p_i factors $\psi_i(x)$, we have $\prod_i \binom{p_i}{p_{i1}p_{i2}\dots p_{ik}}$ 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 $\sum_i ip_{ij} = h_j$.

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

$$\psi_\mu(x) = \sum_\lambda \beta_\lambda(\mu)\Sigma_\lambda(x). \tag{25.3.4}$$

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

$$\Sigma_\lambda(x) = \sum_\mu p_{\lambda,\mu}S_\mu(x),\ S_\lambda(x) = \sum_\mu k_{\lambda,\mu}\Sigma_\mu(x) \tag{25.4.1}$$

In order to explicit some information about the matrices:

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

recall that the partitions are totally ordered by lexicographic ordering.

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

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

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

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

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

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

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

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

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

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

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

we deduce that:

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

This computation has the following immediate consequence:

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

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

We can now conclude:

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

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

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

Proof. From the various definitions we get:

$$c_\lambda = \sum_\phi p_{\phi,\lambda} b_\phi,\ \beta_\lambda = \sum_\phi k_{\phi,\lambda} c_\phi, \tag{25.4.2}$$

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 repreentation into irreducible components.[5]

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 permutatin representation on S_n/S_n.

Assume $\chi_\mu = c_\mu$ for all $\mu < \lambda$, then we may use the Lemma 23.4 and know that M_λ appears in its isotypic component in Rs_T with multiplicity 1. But we have remarked that Rs_T is the permutation representation of character β_λ in which by assumption the representation M_λ appears and thus it must be given in the character y 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 determinand getting

$$\psi_\mu(x)V(x) = \sum_\lambda c_\lambda(\mu) A_\lambda(x) \tag{25.4.3}$$

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_λ.

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

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

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

It will be necessary to work formally with symmetric functions in infinitely many variables, a justification of the validity of this formalism maybe derived by the following remark.

The ring of symmetric functions in n variables x_i is the polynomial ring in the elementary symmetric functions $\sigma_i(n)$. When we set $x_n = 0$ the function $\sigma_n(n)$ goes to 0 while the remaing $\sigma_i(n)$ specialize to $\sigma_i(n-1)$, thus if we restrict to functions of degree $< n$ setting $x_n = 0$ induces an isomorphism.

In other words we can consider the symmetric functions in infinite variables as the polynomial ring in infinite variables σ_i where σ_i has degree i. If we need to establish a formal identity between such variables in a given degree $< n$ it is enough to establish it for symmetric functions in n variables. With this in mind we think of the identities 25.1.1, 25.3.4, 25.4.1 etc. as identities in infinitely many variables.

Consider S_n acting on the space C^n permuting the coordinates and thus on the polynomial ring $\mathbb{C}[x_1, x_2, \ldots, x_n]$ we want to study this representation.

Since this is a graded algebra we shall compute the character as a function χ_i of the degree and express the result compactly with a generating function $\sum_{i=0}^{\infty} \chi_i q^i$.

If σ is a permutation with cycle decomposition of lengths $\mu := m_1, m_2, \ldots m_n$ the bstandard basis of $\mathbb{C}^n$ decomposes cycles. On a subspace relative to a cycle of length k, σ acts with eigenvalues the m-roots of 1 and

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

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

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

To summarize

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

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

Now we want to discuss another related representation.

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

We claim that it is always the regular representation.

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

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

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

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

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

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

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

and thus an identity of graded characters.

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

Theorem.

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

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 sum the dimension of χ_λ.

§26 The hook formula

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

From 25.4.3 applied to the partition 1^n, corresponding to the conjugacy class of the identity, we obtain:

$$(\sum_{i=1}^{n} x_i)^n V(x) = \sum_{\lambda} d(\lambda) A_\lambda(x) \tag{26.1.1}$$

Write the development of the Vandermonde determinant as $\sum_{\sigma\in S_n}\epsilon_\sigma\prod_{i=1}^n x_i^{\sigma(n-i+1)-1}$. Letting $\lambda+\rho=\ell_1>\ell_2>\cdots>\ell_n$ the number $d(\lambda)$ is the coefficient of $\prod_i x_i^{\ell_i}$ in

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

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

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

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

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

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

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

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

$$d(\lambda)=\frac{n!\prod_{i<j}(\ell_i-\ell_j)}{\prod_{i=1}^n\ell_i!} \tag{26.1.2}$$

26.2 We want to give the combinatorial interpretation of 26.1.2. Notice that of the $\sum_i\ell_i=n+\binom{n}{2}$ factors in $\prod_{i=1}^n\ell_i!$ exactly $\binom{n}{2}$ cancel with the corresponding factors in $\prod_{i<j}(\ell_i-\ell_j)$, and n factors are left.

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

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

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

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

Now the reader can convince himself that the factors in the factorial $\ell_i!$ which are not cancelled are the hooks of the boxes in the i^{th} row.

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

$$d(\lambda) = \frac{n!}{\prod_{x \in B(\lambda)} h_x}. \tag{26.2.1}$$

§27 Characters of the linear group

27.1 We plan to deduce from the previous computations the character theory of the linear group.

For this we need to perform another character computation. Given a permutaion $s \in S_n$ 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 $\mu = h_1, h_2, \dots, h_k$ denote the cycle partition of s, introduce the obvious notation (justified by 8.1):

$$\Psi_\mu(X) = \prod_i tr(X^{h_i}). \tag{27.1.1}$$

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 $X_1, X_2, \dots, X_n$ compute the trace of $sX_1 \otimes X_2 \otimes \ldots \otimes X_n$ (an operator in $V^{\otimes n}$).

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

$$\phi_c(X) = \phi_c(X_1, X_2, \dots, X_n) := tr(X_{i_1} X_{i_2} \dots\ X_{i_p}). \tag{27.1.2}$$

The previous proposition then follows from:

Theorem.

$$tr(sX_1 \otimes X_2 \otimes \dots \otimes X_n) = \prod_{j=1}^{k} \phi_{c_j}(X). \tag{27.1.3}$$

Proof. Remark that, for fixed s, both sides of 27.1.3 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 $sX_1 \otimes X_2 \otimes \dots \otimes X_n$ to a decomposable tensor $v_1 \otimes v_2 \dots \otimes v_n$ we have:

$$sX_1 \otimes X_2 \otimes \dots \otimes X_n(v_1 \otimes v_2 \dots \otimes v_n) = \prod_{i=1}^{n} < \psi_i | v_i > u_{s^{-1}1} \otimes u_{s^{-1}2} \dots \otimes u_{s^{-1}n}. \tag{27.1.4}$$

This formula shows that:

$$sX_1 \otimes X_2 \otimes \dots \otimes X_n = (u_{s^{-1}1} \otimes \psi_1) \otimes (u_{s^{-1}2} \otimes \psi_2) \dots \otimes (u_{s^{-1}n} \otimes \psi_n) \tag{27.1.5}$$

so that:

$$tr(sX_1 \otimes X_2 \otimes \dots \otimes X_n) = \prod_{i=1}^{n} < \psi_i | u_{s^{-1}i} > = \prod_{i=1}^{n} < \psi_{si} | u_i > . \tag{27.1.6}$$

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

$$\phi_c(X) = tr(X_{i_1} X_{i_2} \dots\ X_{i_p}).$$

We get:

$$tr(u_{i_1} \otimes \psi_{i_1} u_{i_2} \otimes \psi_{i_2} \dots \otimes u_{i_p} \otimes \psi_{i_p}) = tr(u_{i_1} \otimes < \psi_{i_1} | u_{i_2} > < \psi_{i_2} | u_{i_3} > \dots < \psi_{i_{p-1}} | u_{i_p} > \psi_{i_p})$$

$$= < \psi_{i_1} | u_{i_2} > < \psi_{i_2} | u_{i_3} > \dots < \psi_{i_{p-1}} | u_{i_p} > < \psi_{i_p} | u_{i_1} > = \prod_{j=1}^{p} < \psi_{c(i_j)} | u_{i_j} > \tag{27.1.7}$$

Formulas 27.1.6 and 27.1.7 imply the claim. $\square$

27.2 In order to apply the theory of symmetric functions to the representation theory of the linear group it is necessary to complete in some simple points our discussion of symmetric functions. We let $m + 1$ be the number of variables. We want to understand, given a Schur function $S_\lambda(x_1, \dots, x_{m+1})$ the form of $S_\lambda(x_1, \dots, x_m, 0)$ as symmetric function in m variables.

Let $\lambda := h_1 \geq h_2 \geq \cdots \geq h_{m+1} \geq 0$, in 6.3 we have seen that, if $h_{m+1} > 0$ then $S_\lambda(x_1, \ldots, x_{m+1}) = \prod_{i=1}^{m+1} x_i S_{\overline{\lambda}}(x_1, \ldots, x_{m+1})$ where $\overline{\lambda} := h_1 - 1 \geq h_2 - 1 \geq \cdots \geq h_{m+1} - 1$.

In this case, clearly $S_\lambda(x_1, \ldots, x_m, 0) = 0$.

Assume now $h_{m+1} = 0$ and denote by the same symbol λ the sequence $h_1 \geq h_2 \geq \cdots \geq h_m$. Let us start from the Vandermonde determinant $V(x_1, \ldots, x_m, x_{m+1}) = \prod_{i<j\leq m+1}(x_i < x_j)$ and set $x_{m+1} = 0$ getting

$$V(x_1, \ldots, x_m, 0) = \prod_{i=1}^{m} x_i \prod_{i<j\leq m} (x_i - x_j) = \prod_{i=1}^{m} x_i V(x_1, \ldots, x_m).$$

Now consider the alternating function $A_\lambda(x_1, \ldots, x_m, x_{m+1})$.

Set $\ell_i := h_i + m + 1 - i$ so that $\ell_{m+1} = 0$ and

$$A_\lambda(x_1, \ldots, x_m, x_{m+1}) = \sum_{\sigma \in S_{m+1}} \epsilon_\sigma x_1^{\ell_{\sigma(1)}} \ldots x_{m+1}^{\ell_{\sigma(m+1)}},$$

setting $x_{m+1} = 0$ we get the sum restricted only on the terms for which $\sigma(m+1) = m+1$ or

$$A_\lambda(x_1, \ldots, x_m, 0) = \sum_{\sigma \in S_m} \epsilon_\sigma x_1^{\ell_{\sigma(1)}} \ldots x_m^{\ell_{\sigma(m)}}$$

now in $m-$ variables the partition λ corresponds to the decreasing sequence $h_i + m + 1 - i$ hence

$$A_\lambda(x_1, \ldots, x_m, 0) = \prod_{i=1}^{m} A_\lambda(x_1, \ldots, x_m), \quad S_\lambda(x_1, \ldots, x_m, 0) = S_\lambda(x_1, \ldots, x_m).$$

Thus we see that, under the evaluation of x_{m+1} to 0 the Schur functions S_λ vanish, if $heigth(\lambda) = m + 1$ otherwise they map to the corresponding Schur functions in $m-$variables.

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

From the theory of Schur functions the space S_m^n has as basis the functions $S_\lambda(x_1, \ldots, x_m)$ where $\lambda \vdash n$ has heigth $\leq m$. Under the evaluation $x_m \to 0$ we have a map $S_m^n \to S_{m-1}^n$. We have proved that this map is an isomorphism as soon as $m > n$ hence all identities which we prove for symmetric functions in n variables of degree n are valid in any number of variables.[6]

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 heigth $\leq m$ let us denote by $S_\lambda(X) := S_\lambda(x)$ the Schur function evaluated in $x = (x_1, \ldots, x_m)$ the eigenvalues $x = (x_1, \ldots, x_m)$ of X.

[6]One way of formalizing this is to pass formally to a ring of symmetric functions in infinitely many variables which has as basis all Schur functions without restriction to the heigth and is a polynomial ring in infinitely many vaiables corresponding to all possible *elementary symmetric functions*.

Theorem. *$S_\lambda(X)$ is the character ρ_λ of the representation V_λ of $GL(V)$, paired with the representation M_λ of S_n in $V^{\otimes n}$.*

Proof. If $s \in S_n, X \in GL(V)$ we have seen that the trace of sX on $V^{\otimes n}$ is $\psi_\mu(X) = \sum_\lambda c_\lambda(\mu)S_\lambda(X)$ by definition of the c_λ, if $m = \dim V < n$ only the partitions of heigth $\leq m$ contribute to the sum. On the other hand $V^{\otimes n} = \sum_\lambda V_\lambda \otimes M_\lambda$ thus:

$$\psi_\mu(X) = \sum_{\lambda \vdash n, ht(\lambda) \leq m} \chi_\lambda(\mu)\rho_\lambda(X).$$

Since the functions $c_\lambda = \chi_\lambda$ are the irreducible characters of S_n by Frobenius Theorem, and the functions $S_\lambda(X)$ are linearly independent it follows that $S_\lambda(X) = \rho_\lambda(X)$ □

Remark. The eigenvalues of $X^{\otimes n}$ are monomials in the variables x_i and thus the characters $S_\lambda(x)$ are sums of monomials with positive coefficients.

We make some comments. Recall that $V_\lambda = e_T(V^{\otimes n})$ is a quotient of $a_T(V^{\otimes n}) = \wedge^{k_1} V \otimes \wedge^{k_2} V \ldots \otimes \wedge^{k_n} V$ where the k_i are the columns of a tableau T and is also contained in $s_T(V^{\otimes n}) = S^{h_1} V \otimes S^{h_2} V \ldots \otimes S^{h_n} V$ 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 indeces.[7] The composition $e_T = \frac{1}{p(\lambda)} s_T a_T$ can be viewd as the result of a map

$$\wedge^{k_1} V \otimes \wedge^{k_2} V \ldots \otimes \wedge^{k_n} V \longrightarrow s_T(V^{\otimes n}) = S^{h_1} V \otimes S^{h_2} V \ldots \otimes S^{h_n} V.$$

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

The character of the exterior power $\wedge^i(V)$ is the i^{th} elementary symmetric function $\sigma_i(x)$ the one of the symmetric power $S^h(V)$ is the sum of all monomials of degree h.

We may want to stress the fact that the construction of the representation V_λ from V is in a sense *natural*. The modern language is that 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.*

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

Choosing a Young symmetrizer c associated to a partition λ we consider $cV^{\otimes n}$. As V varies this can be considered as a functor. In fact it is a subfunctor of the tensor power since clearly by the commuting relations we have $A^{\otimes n}(cV^{\otimes n}) \subset cW^{\otimes n}$.

It is then useful to use a functorial notation and indicate the space $cV^{\otimes n}$ by:

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

[7] it is awkward to denote symmetrization on non consecutive indeces as we did, more correctly one should compose with the appropriate permutation which places the indeces in the correct positions.

If $A : V \to W$ is a linear map we define:

$$S_\lambda(A) : S_\lambda(V) \longrightarrow V^{\otimes n} \xrightarrow{A^{\otimes n}} W^{\otimes n} \xrightarrow{e_\lambda} S_\lambda(W)\cdot$$

Remark that the exterior power $\wedge^k V$ and the symmetric power $S^k(V)$ are both examples of Schur functors.

We will need the following simple property of Schur functors which is clear from the definition

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

One can prove that any polynomial functor is equivalent to a direct sum of Schur functors (ins.).

27.3 We want to give now the complete list of irreducible representations for the general and special linear groups.

Recall first of all that, in 6.3 we have shown that, if $\lambda = (m_1, m_2, \ldots, m_n)$ and $\lambda' = (m_1+1, m_2+1, \ldots, m_n+1)$ then $S_{\lambda'} = (x_1 x_2 \ldots x_n) S_\lambda$.

In the language of representations of $GL(n) = GL(V)$, $x_1 x_2 \ldots x_n$ is the character of the 1-dimensional representation $\wedge^n(V)$ which for simplicity we will denote by D.

From the point of view of Young symmetrizers we have $\lambda \vdash m$, $\lambda' \vdash m+n$ and a tableau for λ' can be taken with first column, of n boxes, filled with the numbers $1, 2, \ldots, n$.

Let S be the tableau of shape λ obtained by removing this column.

The effect of the antisymmetrizer a of this column on $V^{\otimes n+m}$ is $\wedge^n V \otimes V^{\otimes m}$. If e_S is the Young symmetrizer for S then $ae_S V^{\otimes n+m} = \wedge^n V e_S \otimes V^{\otimes m}$ is an irreducible $GL(V)$ since $e_S \otimes V^{\otimes m}$ is irreducible and $\wedge^n V$ is 1-dimensional.

Thus if s is any permutation $se_S \otimes V^{\otimes m}$ is an irreducible in the same isotypic component. It follows that $e_T V^{\otimes n+m}$ is isomorphic to $\wedge^n V \otimes e_S V^{\otimes m}$.

Any λ can be uniquely written in the form $(m, m, m, \ldots, m) + (k_1, k_2, \ldots, k_{n-1}, 0)$ (in other words representing it as a diagram with $\leq n$ columns, we collect all rows of length exactly n and then the rest) and we obtain that the irreducible representations appearinfg in the tensor powers are isomorphic to $D^k S_\lambda(V)$ with $ht(\lambda) \leq n-1$..

Let $\mathcal{P}_{n-1} := \{(k_1 \geq k_2 \geq \ldots \geq k_{n-1} \geq 0)$. Consider the polynomial ring

$$\mathbb{Z}[x_1, \ldots, x_n, (x_1, x_2 \ldots x_n)^{-1}]$$

obtained by inverting $\sigma_n = x_1 x_2 \ldots x_n$

Lemma. *The ring of symmetric elements in $\mathbb{Z}[x_1, \ldots, x_n, (x_1 x_2 \ldots x_n)^{-1}]$ is generated by $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}, \sigma_n^{\pm 1}$ and it has as basis the elements:*

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

The ring $\mathbb{Z}[\sigma_1, \sigma_2, \ldots, \sigma_{n-1}, \sigma_n]/(\sigma_n - 1)$ has as basis the classes of the elements: $S_\lambda \sigma_n^m$, $\lambda \in \mathcal{P}_{n-1}$.

Proof. Since σ_n is symmetric it is clear that a fraction $\frac{f}{\sigma_n^k}$ is symmetric if and only if f is symmetric, hence the first statement. As for the second remark that any element of $\mathbb{Z}[\sigma_1, \sigma_2, \dots, \sigma_{n-1}, \sigma_n^{\pm 1}]$ can be written in a unique way in the form $\sum_{k\in\mathbb{Z}} a_k \sigma_n^k$ with $a_k \in \mathbb{Z}[\sigma_1, \sigma_2, \dots, \sigma_{n-1}]$. Now we know that the Schur functions S_λ, $\lambda \in \mathcal{P}_{n-1}$ are a basis of $\mathbb{Z}[\sigma_1, \sigma_2, \dots, \sigma_{n-1}, \sigma_n]/(\sigma_n)$ and the claims follow. $\square$

27.4 We need to give now an interpretation in the language of representations of the Cauchy formula.

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

$$\chi_U(t) := \sum_i \chi_i t^i. \tag{27.4.1}$$

Where χ_i is the character of the representation U_i.

Definition. *The expression 27.4.1 is called a graded character.*

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

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

For the graded characters we have clearly:

$$\chi_{U\oplus V}(t) = \chi_U(t) + \chi_V(t),\ \chi_{U\otimes V}(t) = \chi_U(t)\chi_V(t). \tag{27.4.2}$$

Let us consider a simple example.

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

$$\frac{1}{det(1 - tA)} \tag{27.4.3}$$

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

$$det(1 + tA) \tag{27.4.4}$$

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

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

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

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

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

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

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

□

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.

We want to compute the character of the action of T on the symmetric and exterior algebras, the same proof as before shows that they are respectively:

(27.4.5)
$$\frac{1}{\prod 1 - \chi_i t}, \ \prod 1 + \chi_i t.$$

As an example consider two vector spaces U, V with basis $u_1, \ldots, u_m;\ v_1, \ldots, v_n$ respectively.

The maximal torus of diagonal matrices has eigenvalues $x_1, \ldots, x_m;\ y_1, \ldots, y_n$ respectively for each of them. On the tensor product we have the action of the product group 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 $\prod_{i=1}^{n} \prod_{j=1}^{m} \frac{1}{1 - x_i y_j t}$. We can drop the variable t just recalling the degree in the variables $x_i y_j$.

We want to apply now Cauchy's formula. We need a remark, assume for instance $m \leq n$ so that $\prod_{i=1}^{n} \prod_{j=1}^{m} \frac{1}{1 - x_i y_j t}$ is obtained from $\prod_{i=1}^{n} \prod_{j=1}^{m} \frac{1}{1 - x_i y_j t}$ by setting $y_j = 0, \forall m < j \leq n$.

From Cauchy's formula and Theorem 27.2 to deduce that

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

This identity expresses the characteer in terms of irreducible characters of $GL(U) \times GL(V)$, the character of the n^{th} symmetric power $S^n(U \otimes V)$ equals $\sum_{\lambda \in \vdash (n)} S_\lambda(x) S_\lambda(y)$.

We know that the rational representations of this group are completely reducible thus we have the description:

(27.4.6)
$$S^n(U \otimes V) = \oplus_{\lambda \vdash n,\, ht(\lambda) \leq m} S_\lambda(U) \otimes S_\lambda(V).$$

Remark that, if $W \subset V$ is a subspace which we may assume 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 $S_\lambda(U) \otimes S_\lambda(V)$ relative 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 $S_\lambda(y_1, y_2, \ldots, y_m)$ thus we clearly get that

(27.4.7)
$$S_\lambda(U) \otimes S_\lambda(V) \cap S(U \otimes W) = S_\lambda(U) \otimes S_\lambda(W)$$

It is useful to discuss also a variation of the previous Theorem. Consider two vector spaces V, W the space $\hom(V,W) = W \otimes V^*$ the ring of polynomial functions

$$\mathcal{P}[\hom(V,W)] = S[W^* \otimes V] = \oplus_\lambda S_\lambda(W^*) \otimes S_\lambda(V) \tag{27.4.8}$$

A way to explicitely 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.

Given any partition μ we have seen that $A \to S_\mu(A)$ is a polynomial functor on vector spaces and thus for every $X : V \to W$ we have a map $S_\mu(X) : S_\mu(V) \to S_\mu(W)$.

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, if $\mu \vdash n$.

Thus the dual map $\quad S_\mu^* : \hom(S_\mu(V), S_\mu(W))^* \to \mathcal{P}(\hom(V,W)) \quad$ defined by

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

is a $GL(V) \times GL(W)$ equivariant map.

By the irreducibility of $\hom(S_\mu(V), S_\mu(W))^* = S_\mu(V) \otimes S_\mu(W)^*$ it 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 comonent of type $S_\mu(V)$ we also recover the isomorphism $S_\mu(W^*) = S_\mu(W)^*$.

Apply the previous discussion to $\hom(\wedge^i V, \wedge^i W)$.

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 $\mathcal{P}(\hom(V,W)) = \mathbb{C}[x_{ij}], i = 1, \dots, h,\ j = 1, \dots, k$ where x_{ij} are the matrix entries.

Given a matrix X the entries of $\wedge^i X$ are the determinants of all the minors of order i extracted from X, and $\wedge^i V \otimes (\wedge^i W)^*$ can be identified to the space of polynomials spanned by the determinants of all the minors of order i extracted from X.

One should bear in mind the strict connection between the two formulas, 27.4.5 and 24.1.1,

$$S^n(U \otimes V) = \oplus_{\lambda \vdash n,\, ht(\lambda) \le m} S_\lambda(U) \otimes S_\lambda(V) \qquad V^{\otimes n} = \oplus_{ht(\lambda) > dim(V)} M_\lambda \otimes S_\lambda(V).$$

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 $Xe_i = x_i e_i$.

If $\chi : T \to \mathbb{C}^*$ is the character $\chi(X) = \prod_i x_i$ (which we will call *the multilinear characer*), we see that

Proposition. *$V^{\otimes n}$ can be canonically identified in an S_n equivariant way to the weight space $S^n(U \otimes V)^\chi = \oplus_{\lambda \vdash n,\, ht(\lambda) \le m} S_\lambda(U)^\chi \otimes S_\lambda(V)$, under this identification $S_\lambda(U)^\chi = M_\lambda$ as representations of S_n.*

27.5 We can interpret the previous theory as part of the description of the coordinate ring $A[GL(V)]$ of the general linear group.

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

(27.5.1)
$$A[GL(V)] = \oplus_{ht(\lambda)\leq n-1,\, k\in\mathbb{Z}} d^k S_\lambda(V^*) \otimes S_\lambda(V).$$

We deduce

Theorem.

(1) *The irreducible representations of $GL(V)$ are:*

$$S_\lambda(V)D^m,\ \lambda \in \mathcal{P}_{n-1},\ m \in \mathbb{Z}.$$

(2) $[S_\lambda(V)D^m]^* = S_\lambda(V^*)D^{-m}$

(3) *For a partition $\lambda := h_1, h_2, \dots, h_n$ let $\lambda' := h'_i$ be as in 6.3 ($h_i + h'_{n-i+1} = h_1$)we have $S_\lambda(V^*) = S_{\lambda'}(V)D^{-h_1}$.*

Proof. 1) and 2) follow from 27.4.6 and theorem 15.2.1. 3) follows from the computation of 6,3 remarking that on V^* the eigenvalues of the torus $x_1, \dots, x_n$ are $x_1^{-1}, \dots, x_n^{-1}$. □

As for $SL(V)$ we remark first of all that the determinant representation D restricted to $SL(V)$ is trivial. Therefore all the representations $S_\lambda(V)D^m$, restricted to $SL(V)$, are isomorphic. Now we claim that $S_\lambda(V)$ is irreducible as representation of $SL(V)$ and that:

Theorem. *The irreducible representations of $SL(V)$ are:*

$$S_\lambda(V),\ \lambda \in \mathcal{P}_{n-1}.$$

Proof. 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 $SL(V)$. To conclude it is enough to remark 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 8.1 and the previous lemma. □

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

We go back to the Cauchy identity

(27.6.1)
$$\sum S_\lambda(x)S_\lambda(y) = \prod_{i,j} \frac{1}{1 - x_i y_j} = \prod_j \sum_{k=0}^{\infty} s_k(x) y_j^k$$

the symmetric functions $s_k(x)$ are defined by the power series development of

$$\frac{1}{\prod_i 1 - x_i t} = \sum_{k=0}^{\infty} s_k(x) t^k$$

and are by 27.4 the characters of the symmetric powers of V.

One easily sees that $s_k(x)$ is the sum of all the monomials in $x_1, \ldots, x_n$ of degree k.

Multiply 27.6.1 by the Vandermonde determinant $V(y)$ getting

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

For a given $\lambda = h_1, \ldots, h_n$ we see that $S_\lambda(x)$ is the coefficient of the monomial $y_1^{h_1+n-1} y_2^{h_2+n-2} \ldots y_n^{h_n}$ and we easily see that this is

$$\sum_{\sigma \in S_n} \epsilon_\sigma s_{\sigma(i)-i+h_i} \tag{27.6.1}$$

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 s_{j-i+h_i} with the convention that $s_k = 0,\ \forall k < 0$.*

27.7 We want to complete this discussion with an interesting variation of the Cauchy formula which we will use later in the computation of the cohomology of the linear group.

We are given again two vector spaces V, W and want to describe $\wedge(V \otimes W)$ as representation of $GL(V) \times GL(W)$.

Theorem.

$$\wedge(V \otimes W) = \sum_\lambda S_\lambda(V) \otimes S_{\tilde{\lambda}}(W) \tag{27.7.1}$$

Proof. $\tilde{\lambda}$ denotes as in 25, the dual partition.

We argue in the following way, for very k, $\wedge^k(V \otimes W)$ is a polynomial representation of degree k of both groups hence by the general theory $\wedge^k(V \otimes W) = \oplus_{\lambda \vdash k,\ \mu \vdash k} c^k_{\lambda,\mu} S_\lambda(V) \otimes S_\mu(W)$ for some multiplicities $c^k_{\lambda,\mu}$. To compute these multiplicities we do it in the stable case where $\dim W = k$ using the results of §27.

Next identify $W = \mathbb{C}^k$ with basis e_i and we restrict as in 27.4, to the multilinear elements which are $V \otimes e_1 \wedge \ldots V \otimes e_k$ which, as representation of $GL(V)$ can be identified to $V^{\otimes n}$ 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 mutilinear weight

$$\oplus_{\lambda \vdash k} S_\lambda(V) \otimes M_{\tilde{\lambda}} = (\wedge^k V \otimes \mathbb{C}^k)^\chi = \oplus_{\lambda \vdash k,\ \mu \vdash k} c^k_{\lambda,\mu} S_\lambda(V) \otimes M_\mu$$

from which the claim follows. □

Remark. In terms of characters the previous formula 27.7.1 is equivalent to

$$\prod_{i=1,\ j=1}^{n,\ m}(1+x_iy_j)=\sum_{\lambda}S_{\lambda}(x)S_{\tilde{\lambda}}(y). \tag{27.7.2}$$

There is a simple determinantal formula corollary of this identity as in 27.6.

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

The same reasoning as in 27.6 then gives the formula

$$S_{\lambda}(x)=\sum_{\sigma\in S_n}\epsilon_{\sigma}e_{\sigma(i)-i+k_i} \tag{27.7.3}$$

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 $e_k=0,\ \forall k<0$.*

§28 Branching rules, standard diagrams

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

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

Let h_i denote the rows of λ, we may as well compute:

$$\psi_k(x)A_\lambda(x) = (\sum_{i=1}^{n} x_i^k)(\sum_{s\in S_n} \epsilon_s x_{s1}^{h_1+n-1} x_{s2}^{h_1+n-2} \dots x_{sn}^{h_n}). \tag{28.1.1}$$

Indicate by $k_i = h_i + n - i$. We inspect the monomials appearing in the alternating function which is at the right of 27.4.1.

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

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

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

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

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

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

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

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

This property appears saying that the new diagram λ' is thus any diagram which contains the diagram λ and such that their difference is connected, made of k boxes and it is like a "slinky"[8] i.e. it is part of the *rim* of the diagram λ' (which are the points on its boundary or the points i, j for which there is no point h, k with $i < h, j < k$ lying in the diagram).

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

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

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

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

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

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

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

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

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

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

Murnagham's rule can be formulated as:

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

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

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

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

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

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

With these remarks we notice a special case of 28.1:

$$\psi_1 S_\lambda = \sum_{\mu \vdash |\lambda|+1, \lambda \subset \mu} S_\mu. \tag{28.2.1}$$

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

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

$$= \sum_{\lambda \in \vdash (n-1)} c_\lambda(\nu) \sum_{\mu \in \vdash (n), \lambda \subset \mu} S_\mu. \tag{28.2.1}$$

In other words:

$$c_\lambda(\nu 1) = \sum_{\mu \in \vdash (n-1), \mu \subset \lambda} c_\lambda(\nu), \tag{28.2.2}$$

In module notations we have:

Branching rule for the symmetric group :

$$M_\lambda = \oplus_{\mu \in \vdash (n-1), \mu \subset \lambda} M_\mu. \tag{28.2.3}$$

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 28.2.3 is unique.

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

$$M_\lambda = \oplus_{\mu_2 \vdash = n-2,\ \mu_1 \vdash = n-1, \mu_2 \subset \mu_1 \subset \lambda} M_{\mu_2}. \tag{28.2.4}$$

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

From 4261, branching once:

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

Banching twice:

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

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

In general we give the following definitions:

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

Example of a skew tableau of shape 6527/326:

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

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

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

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

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

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

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

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

In particular, for a given shape μ the multiplicity of M_μ in M_λ equals the number of such tableaux.

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

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

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

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

Let Z_k be the center of $\mathbb{Q}[S_k]$. The algebras $Z_k \subset \mathbb{Q}[S_n]$ generate a commutative subalgebra C.

On any irreducible representation this subalgebra, by the analysis made has a basis of common eigenvectors given by the decomposition into 1 dimensional spaces previously described. It is easy to see that the eigenvalues are distinct and so this decomposition is again unique.

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

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

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

§29 Applications to invariant theory

29.1 Assume now we have an action of a group G on a space U of dimension m and we want to compute the invariants of n copies of U. Assume first $n \geq m$.

We think of n copies of U as $U \otimes \mathbb{C}^n$ and the linear group $GL(n, \mathbb{C})$ acts on this vector space by tensor action on the second factor. As we have seen in Chapter 1 the Lie algebra of $GL(n, \mathbb{C})$ acts by polarization operators. The ring of G invariants is stable under these actions.

From 27.4.5 we have that the ring of polynomial functions on $U \otimes \mathbb{C}^n$ equals

$$S(U^* \otimes \mathbb{C}^n) = \oplus_{\lambda \vdash n,\, ht(\lambda) \leq m} S_\lambda(U^*) \otimes S_\lambda(\mathbb{C}^n).$$

For the invariants under G we clearly have that

$$S(U^* \otimes \mathbb{C}^n)^G = \oplus_{\lambda \vdash n,\, ht(\lambda) \leq m} S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^n). \tag{29.1.1}$$

This formula describes the ring of invariants of G acting on the polynomial ring of U^n as representation of $GL(n, \mathbb{C})$. In particular we see that the multiplicity of $S_\lambda(\mathbb{C}^n)$ in the ring $\mathcal{P}(U^n)^G$ equals the dimension of the space of invariants $S_\lambda(U^*)^G$.

The restriction on the heigth implies that, when we restrict to m copies we again have the same formula and hence we deduce comparing isotypic components and by 27.4.6.

$$S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^n) \cap S(U^* \otimes \mathbb{C}^m) = S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^n)$$

We deduce

Theorem. *If* $\dim U = m$*, the ring of invariants of* $S(U^* \otimes \mathbb{C}^n)$ *is generated, under polarization, by the invariants of* m *copies of* U*.*

Proof. Each isotypic component (under $GL(n, \mathbb{C})$), $S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^n)$ is generated under this group by $S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^m)$ since $ht(\lambda) \leq m$. $\square$

There is a useful refinement of this theorem. Assume that the group G is contained in the special linear group. In this case an invariant in m variabes is the *determinant* of $m-$vectors, i.e. a generator of the 1-dimensional space $\wedge^m U \otimes \wedge^m \mathbb{C}^m$.

If λ is a partition of heigth exactly m we know that $S_\lambda(U \otimes \mathbb{C}^m) = (\wedge^m U \otimes \wedge^m \mathbb{C}^n)^k S_\mu(U \otimes \mathbb{C}^m)$ where $ht(\mu) \leq m - 1$. Thus we obtain by the same reasoning.

Theorem. *If* $\dim U = m$ *and* $G \subset SL(U)$*, the ring of* G *invariants of* $S(U^* \otimes \mathbb{C}^n)$ *is generated, under polarization, by the determinant and invariants of* $m - 1$ *copies of* U*.*

Alternatively we could use the theory of highest weights (cf. §33) and remark that the $G \times U^+$ invariants are $\oplus S_\lambda(U^*)^G \otimes S_\lambda(\mathbb{C}^n)^{U^+}$. They are contained in $S(U^* \otimes C^{m-1})[d]$ (the determinant d of the first m vectors u_i) and generate under polarization all the invariants.[9]

29.2 We shall discuss now the first fundamental Theorem of the special linear group $SL(V),\ V = \mathbb{C}^n$ acting on m copies of the fundamental representation.

We identify V^m with the space of $n \times m$ matrices with $SL(n, \mathbb{C})$ acting by left multiplication.

We consider the polynomial ring on V^m as the ring $\mathbb{C}[x_{ij}],\ x := (x_{ij})$ of polynomials in the entries of these matrices.

Given n indeces $i_1, \ldots, i_n$ between $1, \ldots, m$ we shall denote by $[i_1, \ldots, i_n]$ the determinant of the maximal minor of X extracted from the corresponding columns.

[9] In the same way any subspace $W \subset S(U^* \otimes \ {}^n)$ stable under polaryzation is generated by $W \cap S(U^* \otimes \ {}^m)$, while a subspace $W \subset S(U^* \otimes \ {}^n)$ stable under polaryzation and multiplication by the determinant d of the first m vectors u_i is generated by $W \cap S(U^* \otimes \ {}^{m-1})$ under polarizations and multiplication by d.

Theorem. *The ring of invariants* $\mathbb{C}[x_{ij}]^{SL(n,\mathbb{C})}$ *coincides with the ring* $\mathbb{C}[[i_1,\dots,i_n]]$ *generated by the* $\binom{m}{n}$ *elements* $[i_1,\dots,i_n],\ 1 \le i_1 < \cdots < i_n \le m$.

Proof. We can apply the previous theorem using the fact that the proposed ring is certainly made of invariants and closed under polarization, thus it suffices to show that it coincides with the ring of invariants for $n-1$ copies of V.

Now it is clear that, given any $n-1$ linearly independent vectors of V they can be completed to a basis with determinant 1 hence these set of $n-1$ tuples, which is open, forms a unique orbit under $SL(V)$, therefore the invariants of $n-1$ copies are just the constants and the theorem is proved. $\square$

Let us understand this ring of invariants as representation of $GL(m,\mathbb{C})$ using the general formula 29.1.1 and 27.4 we see that $S_\lambda(U)$ is the trivial representation of $SL(U)$ if and only if $\lambda := k^n$ thus we have the result that

Corollary. *The space of polynomials of degree* k *in the elements* $[i_1,\dots,i_n]$ *as representation of* $GL(m,\mathbb{C})$ *equals* $S_{k^n}(\mathbb{C}^m)$.

One can in fact combine this Theorem with the first fundamental Theorem of 22.3 to get a theorem of invariants for $SL(U)$ acting on $U^m \oplus (U^*)^p$, we describe it as pairs of matrices X, Y and we have invariants

$$< \phi_i | u_j >,\ [u_{i_1},\dots,u_{i_n}],\ [\phi_{j_1},\dots,\phi_{j_n}]$$

where the $[u_{i_1},\dots,u_{i_n}]$ are the determinants of the maximal minors of X while $[\phi_{j_1},\dots,\phi_{j_n}]$ are the determinants of the maximal minors of Y.

Theorem. *The ring of invariants of* $\mathcal{P}(U^m \oplus (U^*)^p)^{SL(U)}$ *is the ring generated by*

$$< \phi_i | u_j >,\ [u_{i_1},\dots,u_{i_n}],\ [\phi_{j_1},\dots,\phi_{j_n}].$$

Proof. We apply the methods of 29.1 to both U, U^* and thus reduce to $n-1$ copies of both U, U^*.

In this case the invariants described are only the $< \phi_i | u_j >$, one restricts to the open set in which the $n-1$ vectors are linearly independent and thus up to $SL(U)$ action can be fixed to be $e_1,\dots,e_{n-1}$ the linear forms are $\phi_1,\dots,\phi_{n-1}$. This open set is stable under the subgroup

$$H := \left\{ \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & a_1 \\ 0 & 1 & 0 & \dots & 0 & a_2 \\ & \dots & \dots & \dots & & \dots \\ 0 & 0 & 0 & \dots & 1 & a_{n-1} \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \right\}$$

If we write each $\phi_i = (\psi_i, y_i)$ where ψ_i is a $n-1$ row vector and think of the a_i as the entries of an $n-1$ column vector u the action of this group on the covectors is by

$$\psi_i, y_i \to \psi_i, < \psi_i | u > + y_i$$

one can now restrict to the open set where the ψ_i are linearly independent and where then one can determine u so to make the $y_i = 0$.

It follows then that a function in the coordinates of the ϕ_i is H invariant if and only if it is independent of the y_i.

Next one remarks that the coordinates of the ψ_i are the scalar products $< \phi_i | e_j >$, $j \leq n-1$ and this implies that any $SL(U)$ invariant coincides with a polynomial in the elements $< \phi_i | u_j >$ on a dense open set which implies the Theorem. $\square$

§30 The analytic approach to Weyl's character formula

30.1 The analytic approach is based on the idea of applying the orthogonality relations for class functions in the compact case.

Let us illustrate $U(n, \mathbb{C})$.

Let T be the torus of diagonal matrices $(y_1, \ldots, y_n),\ |y_i| = 1$.

Consider $U(n, \mathbb{C})/T$, this is a compact differentiable manifolds with $U(n, \mathbb{C})$ acting on the left. Locally it can be parametrized by the classes $e^A T$ where A is an antihermitian matrix with 0 in the diagonal. Let H_0 denote this space, it has basis the elements $i(e_{hj} + e_{jh}),\ e_{hj} - e_{jh}$ with $h < j$.

We claim that $U(n, \mathbb{C})/T$ has a left invariant measure. To see this it is enough to check that T (the stabilizer of $\overline{1} = T$ acts with elements of determinant 1 on the tangent space. The action is $t(1 + A) = t(1 + A)t^{-1} = 1 + tAt^{-1}$. Let us compute the determinant of this action we can complexify the space H_0 (which is like complexifying the homogeneous space to $GL(n, \mathbb{C})/T_{\mathbb{C}}$) and obtain as basis the elements $e_{ij},\ i \neq j$ then in local coordinates in $GL(n, \mathbb{C})/T_{\mathbb{C}}$ we have $t(1 + \sum x_{ij} e_{ij}) = t(1 + \sum x_{ij} e_{ij})t^{-1} = (1 + \sum x_{ij} t_i t_j^{-1} e_{ij})$ hence for the determinant we have $\prod_{i \neq j} t_i t_j^{-1} = 1$.

Next we need to prove Weyl's integration formula for class functions.

Let $d\mu$ denote the Haar measure on $U(n, \mathbb{C})$ and $d\nu$, $d\tau$ the normalized Haar measures on $U(n, \mathbb{C})/T$ and T respectively. We have

Theorem Weyl's integration formula. *For a class function f on $U(n, \mathbb{C})$ we have*

$$\int_{U(n,\mathbb{C})} f(g) d\mu = \frac{1}{n!} \int_T f(t) V(t) \overline{V}(t) d\tau \tag{30.1.1}$$

where $V(t_1, \ldots, t_n) := \prod_{i<j} (t_i - t_j)$ is the Vandermonde determinant.

Proof. First of all we consider the map $\pi : U(n, \mathbb{C})/T \times T \to U(n, \mathbb{C})$ given by $\pi(gT, y) := gyg^{-1}$.

Since every unitary matrix has an orthonormal basis of eigenvectors the map π is surjective.

Consider the open sets in $T^0, U(n,\mathbb{C})^0$ made of elements with distinct eigenvalues in $T, U(n,\mathbb{C})$ respectively. It is easily seen that the complements of these sets have measure 0 and that we have again a mapping

$$\pi : U(n,\mathbb{C})/T \times T^0 \to U(n,\mathbb{C})^0$$

this last mapping is a covering with exactly $n!$ sheets, in fact the centralizer of a given matrix X in T^0 is T itself and the conjugacy class of X intersects T in $n!$ elements obtained by permuting the diagonal entries.

Let $\pi^*(d\mu)$ be the form induced by the form defining the Haar measure then,

$$\int_{U(n,\mathbb{C})} f(g)d\mu = \frac{1}{n!}\int_{U(n,\mathbb{C})/T\times T^0} f(gyg^{-1})\pi^*(d\mu).$$

The theorem will be proved since f is constant on conjugacy classes and using Fubini's Theorem, if one shows that $\pi^*(d\mu) = V(t)\overline{V}(t)d\nu d\tau$. This last identity follows from the computation of a Jacobian.

Let us denote $\pi^*(d\mu) = F(g,t)d\nu d\tau$.

In fact let $\omega_1, \omega_2, \omega_3$ represent respectively the value of the form defining the normalized measure in the class of 1 for $U(n,\mathbb{C})/T, T, U(n,\mathbb{C})$ respectively.

By construction $\omega_1 \wedge \omega_2$ is the pull-back under the map $\ell_{g,y} : (hT, z) \to (ghT, yz)$ of the value of $d\nu d\tau$ in the point (gT, y); similarly for $U(n,\mathbb{C})$, the value of $d\mu$ in an element h is the pull-back of ω_3 under the map $r_h : x \to xh^{-1}$.

Thus if we compose the maps

$$\psi := r_{gyg^{-1}}\pi\ell_{g,y}$$

we get that the pull-back of ω_3 under this map ψ is $F(g,t)\omega_1 \wedge \omega_2$.

In order to compute the function $F(g,t)$ we fix a basis of the tangent spaces of $U(n,\mathbb{C})/T, T, U(n,\mathbb{C})$ in 1 and compute the Jacobian of $r_{gyg^{-1}}\pi\ell_{g,y}$ in this basis, this is $F(g,t)$ up to some constant which measures the difference between the determinants of the given bases and the normalized invariant form.

At the end we will compute the constant by comparing the integrals of the constant function 1.

Take as local cordinates in $U(n,\mathbb{C})/T \times T$ the parameters $(1+A)T, 1+D$ where $A \in H_0$ and D is diagonal.

We compute

$$\psi((1+A)T, 1+D) = (g(1+A)y(1+D)(g(1+A))^{-1}gy^{-1}g^{-1} = 1+g[D+A-yAy^{-1}]g^{-1}+\mathcal{O}$$

where we are writing everything in power series and $\mathcal{O}$ are the terms of order > 1.

The required Jacobian is the determinant of the linear map $(D,A) \to g[D + A - yAy^{-1}]g^{-1}$. Since conjugating by g has determinant 1 we are reduced to the map $(D,A) \to$

$(D, A - yAy^{-1})$ and this is a block matrix with one block the identity and we are further reduced to $\gamma : A \to A - yAy^{-1}$.

Again to compute this determinant we complexify and see that $\gamma(e_{i,j}) = (1 - y_i y_j^{-1})e_{i,j}$ and finally the determinant is

$$\prod_{i \neq j}(1 - y_i y_j^{-1}) = \prod_{i<j}(1 - y_i y_j^{-1})(1 - y_j y_i^{-1}) = \prod_{i<j}(y_j - y_i)(y_j^{-1} - y_i^{-1}) = V(y)\overline{V}(y)$$

since y is unitary.

Finally we see that the formula 30.1.1 is true maybe with some multiplicative constant.

The constant is 1 since if we take $f = 1$ the left hand side of 30.1.1 is 1 as for the right remember that the monomials in the y_i (coordinates of T) are the irreducible characters and so they are orthonormal, it follows since $V(y) = \sum_{\sigma \in S_n} \epsilon_\sigma y_{\sigma(1)}^{n-1} y_{\sigma(2)}^{n-2} \cdots y_{\sigma(n-1)}$ that $\frac{1}{n!}\int_T V(y)\overline{V}(y)d\tau = 1$ and the proof is complete. □

Corollary. *The Schur functions $S_\lambda(y)$ are irreducible characters.*

Proof. Now we compute

$$\frac{1}{n!}\int_T S_\lambda(y)\overline{S}_\mu(y)V(y)\overline{V}(y)d\tau = \frac{1}{n!}\int_T A_\lambda(y)\overline{A}_\mu(y)d\tau = \delta_{\lambda,\mu}$$

for the same reason as before. It follows that the class functions on $U(n, \mathbb{C})$ which restricted to T give the Schur functions are orthonormal with respect to Haar measure.

Since the irreducible characters restricted to T are symmetric functions sums of monomials with positive integer coefficients it follows that the Schur functions are $\pm$ irreducible characters.

The sign must be plus since their leading coefficient is 1.

By the description of the ring of symmetric functions with the function $e_n = \prod y_i$ inverted follows immediately that the characters $S_\lambda e_n^k,\ ht(\lambda) < n,\ k \in \mathbb{Z}$ are a basis of this ring and so they exahust all possible irreducible characters. □

§31 The classical groups

31.1 We want to start here the description of the representation theory of other classical groups, in particular the orthogonal and the symplectic group; again we can by the same methods used for the linear group relate invariant theory with representation theory.

We fix a vector space V with a non degenerate invariant bilinear form, let us indicate by (u, v) the case of a symmetric bilinear form and by $[u, v]$ the antisymmetric form defining the two groups of linear transformations fixing the form.

$O(V)$ the orthogonal group and $Sp(V)$ the symplectic group.

We start remarking that for the fundamental representation of either one of these two groups we have a non degenerate invariant bilinear form which identifies the fundamental representation with its dual, thus for the first fundamental theorem it suffices to analyze the invariants of several copies of the fundamental representation.

The symplectic group $Sp(V)$ on a $2n$ dimensional vector space is formed by unimodular matrices.

Fix a symplectic basis for which the matrix of the skew form is $J = \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}$. Given $2n$ vectors u_i which we write as column vectors of a $2n \times 2n$ matrix A the determinant of A equals the Pfaffian of the matrix $A^t JA$ which has as entries the skew products $[u_i, u_j]$ (by Theorem 12.9) hence.

Lemma. *The determinant $|u_1, \ldots, u_{2n}|$ equals the Pfaffian of the matrix with entries $[u_i, u_j]$.*

The orthogonal group instead contains a subgroup of index 2, the special linear group $SO(V)$ formed by the orthogonal matrices of determinant 1.

When we discuss either the symplectic or the special orthogonal group we suppose to have chosen a trivialization $\wedge^{\dim V} V = \mathbb{C}$ of the top exterior power $\wedge^{\dim V} V = \mathbb{C}$ of V.

If $m = \dim V$ and $v_1, \ldots, v_m$ are m vector variables, the element $v_1 \wedge \cdots \wedge v_m$ is to be understood as a function on $V^{\oplus m}$ invariant under the special linear group.

Given a direct sum $V^{\oplus k}$ of the fundamental representation we indicate by $u_1, \ldots, u_k$ a typical element of this space.

Theorem.

i) *The ring of invariants of several copies of the fundamental representation of $SO(V)$ is generated by the scalar products (u_i, u_j) and by the determinants $u_{i_1} \wedge u_{i_2} \wedge \ldots u_{i_m}$.*

ii) *The ring of invariants of several copies of the fundamental representation of $O(V)$ is generated by the scalar products (u_i, u_j)*

iii) *The ring of invariants of several copies of the fundamental representation of $Sp(V)$ is generated by the functions $[u_i, u_j]$.*

Before we prove this theorem we formulate it in matrix language.

Consider the group $O(n, \mathbb{C})$ of $n \times n$ matrices X with $X^t X = XX^t = 1$ and consider the space of $n \times m$ matrices Y with the action of $O(n, \mathbb{C})$ given by multiplication XY then the mapping $Y \to Y^t Y$ from the space of $n \times m$ matrices to the symmetric $m \times m$ matrices of rank $\leq n$ is a quotient under the orthogonal group.

Similarly let

$$J_n := \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}$$

be the standard $2n \times 2n$ skew symmetric matrix and $Sp(2n, \mathbb{C})$ the standard symplectic group of matrices such that $X^t J_n X = J_n$. Then consider the space of $2n \times m$ matrices Y with the action of $Sp(2n, \mathbb{C})$ given by multiplication XY then the mapping $Y \to Y^t J_n Y$ from the space of $2n \times m$ matrices to the antisymmetric $m \times m$ matrices of rank $\leq 2n$ is a quotient under the symplectic group.

Proof of Theorem for $SO(V)$, $O(V)$. We prove first that the theorem for $SO(V)$ implies the theorem for $O(V)$.

One should remark then, that since $SO(V)$ is a normal subgroup of index 2 in $O(V)$ we have a natural action of the group $O(V)/SO(V) \cong \mathbb{Z}/(2)$ on the ring of $SO(V)$ invariants.

Let τ be the element of $\mathbb{Z}/(2)$ corresponding to the orthogonal transformations of determinant -1 (*improper transformations*). The elements (u_i, u_j) are invariants of this action while $\tau(u_{i_1} \wedge u_{i_2} \wedge \ldots u_{i_m}) = -u_{i_1} \wedge u_{i_2} \wedge \ldots u_{i_m}$.

It follows that the orthogonal invariants are polynomials in the special orthogonal invariants in which every monomial contains a product of an even number of elements of type $u_{i_1} \wedge u_{i_2} \wedge \ldots u_{i_m}$.

Thus it is enough to verify the following identity

$$(u_{i_1} \wedge u_{i_2} \wedge \ldots u_{i_m})(u_{j_1} \wedge u_{j_2} \wedge \ldots u_{j_m}) = \det \begin{pmatrix} (u_{i_1}, u_{j_1}) & (u_{i_1}, u_{j_2}) & \ldots & (u_{i_1}, u_{j_m}) \\ (u_{i_1}, u_{j_1}) & (u_{i_1}, u_{j_2}) & \ldots & (u_{i_1}, u_{j_m}) \\ \vdots & \vdots & \vdots & \vdots \\ (u_{i_1}, u_{j_1}) & (u_{i_1}, u_{j_2}) & \ldots & (u_{i_1}, u_{j_m}) \end{pmatrix}$$

This is easily verified since in an orthonormal basis the matrix having as rows the coordinates of the vectors u_i times the matrix having as columns the coordinates of the vectors u_j yields the matrix of scalar products.

Now we discuss $SO(V)$.

Let A be the proposed ring of invariants, from the definition, this ring contains the determinants and is closed under polarization operators.

From 29.1 we deduce that it is enough to prove the Theorem for $m-1$ copies of the fundamental representation.

We work by induction on m and can assume $m > 1$.

We have to use one of the two possible reductions.

First We first prove the theorem for the case of real invariants on a real vector space $V := \mathbb{R}^m$ with the standard euclidean norm. Let e_i denote the canonical basis of this space.

This is enough since any invariant under the compact special orthogonal group is an invariant under the complex orthogonal group and, given a polynomial on the complexified space this is a sum of real and imaginary part on the real space to which we can apply the theorem in the compact case.

In the real case consider $\overline{V} := \mathbb{R}^{m-1}$ formed by the vectors with the last coordinate 0 (and spanned by $e_i,\ i < m$).

We claim that any special orthogonal invariant G on V^{m-1} restricted to $\overline{V}^{m-1}$ is an invariant under the orthogonal group of $\overline{V}$: in fact it is clear that every orthogonal transformation of $\overline{V}$ can be extended to a special orthogonal transformation of V.

By induction therefore we have a polynomial $F((u_i, u_j))$ which, restricted to $\overline{V}^{m-1}$ coincides with G.

We claim that F, G coincide everywhere; let then us choose $m-1$ vectors $u_1, \dots, u_{m-1}$, there is a vector u of norm 1 and orthogonal to these vectors.

There is a special orthogonal transformation which brings this vector u into e_m and thus the vectors u_i into the space $\mathbb{R}^{m-1}$, since both F, G are invariant and they coincide on $\mathbb{R}^{m-1}$ the claim follows.

Second If one does not like the reduction to the real case one can argue as follows, prove first that the set of $m-1$ tuples of vectors which span a non degenerate subspace in V are a dense open set of V^{m-1} and then argue as before. $\square$

Proof of Theorem for $Sp(V)$.
Again let A be the proposed ring of invariants, from the remark on the Pfaffian in the $Sp(V),\ \dim V = 2m$ case we see that also A contains the determinants and is closed under polarization operators.

From 29.1 we deduce that it is enough to prove the Theorem for $2m-1$ copies of the fundamental representation.

We work by induction on $2m$ and can assume $m > 1$.

Assume we have chosen a symplectic basis $e_i, f_i,\ i = 1, \dots, m$ and consider the space of vectors $\overline{V}$ having coordinate 0 in e_1.

On this space the symplectic form is degenerate with kernel spanned by f_1 and it is again non degenerate on the subspace W where both the coordinates of e_1, f_1 vanish.

First we want to prove that a symplectic invariant $F(u_1, \dots, u_{2m-1})$ when computed on elements $u_i \in \overline{V}$ is a function which depends only on the coordinates in $e_i, f_i,\ i > 1$ and thus by induction it is a polynomial in the skew products $[u_i, u_j]$.

In fact consider the symplectic transformations $e_1 \to te_1,\ f_1 \to t^{-1} f_1$ and identity on W.

This induces multiplication by t on the coordinate of f_1 and fixes the other coordinates.

If a polynomial is invariant under this group of transformations it must be independent of this coordinate.

Now given any symplectic invariant G there exists a polynomial $F([u_i,u_j])$ which coincides with G on $\overline{V}^{2m-1}$; we claim that it coincides everywhere.

It is enough by continuity to show this on the set of $2m-1$ vectors which are linearly independent. In this case such a set of vectors generates a subspace where the symplectic form is degennerate with kernel 1 dimensional, by the theory of symplectic forms there is a symplectic transformation which brings this subspace to $\overline{V}$ and the claim follows. $\square$

We can now apply this theory to representation theory.

§32 The classical groups (representations)

32.1 We start from a vector space with a non degenerate symmetric or skew form, before we do any further computations we need to establish some basic dictionary deduced from the identification $V = V^*$ induced by the form.

We want first of all study the identification $End(V) = V \otimes V^* = V \otimes V$. We explicit some formulas in the 2 cases (easy verification):

(32.1.1)
$$(u \otimes v)(w) = u(v,w),\ (u \otimes v) \circ (w \otimes z) = u \otimes (v,w)z,\ tr(u \otimes v) = (u,v)$$
$$(u \otimes v)(w) = u[v,w],\quad (u \otimes v) \circ (w \otimes z) = u \otimes [v,w]z,\quad tr(u \otimes v) = -[u,v]$$

Now we enter in the more interesting point, we want to study the tensor powers $V^{\otimes n}$ under the action of $O(V)$ or of $Sp(V)$.

We already know that these groups are linearly reductive (15.2) and in particular all the tensor powers are completely reducible and we want to study these decompositions.

Let us denote by G one of the two previous groups. Let us for a while use the notations of the symmetric case but the discussion is completely formal and it applies also to the skew case.

First we have to study $\hom_G(V^{\otimes h}, V^{\otimes k})$, from the basic principle of 22.1 we identify

(32.1.2)
$$\hom_G(V^{\otimes h}, V^{\otimes k}) = [V^{\otimes h} \otimes (V^*)^{\otimes k}]^G = [(V^{\otimes h+k})^*]^G$$

thus the space of intertwiners between $V^{\otimes h}, V^{\otimes k}$ can be identified with the space of multilinear invariants in $h+k$ vector variables.

Explicitely on each $V^{\otimes p}$ we have the scalar product $(w_1 \otimes \cdots \otimes w_p, z_1 \otimes \cdots \otimes z_p) := \prod_i (w_i, z_i)$. Next we identify $A \in \hom_G(V^{\otimes h}, V^{\otimes k})$ with

(32.1.3)
$$\psi_A(X \otimes Y) := (A(X), Y),\ X \in V^{\otimes h},\ Y \in V^{\otimes k}.$$

It is convenient to denote the variables as $(u_1, \dots, u_h, v_1, \dots, v_k)$. Theorem 26 implies that these invariants are spanned by suitable monomials (the multilinear ones) in the scalar or skew products between these vectors.

In particular there are non trivial intertwiners if and only if $h + k = 2n$ is even.

It is necessary to identify some special intertwiners.

Contraction The map $V \otimes V \to \mathbb{C}$ given by $u \otimes v \to (u, v)$, is called an elementary contraction.

Extension By duality in the space $V \otimes V$ the space of G invariants is one dimensional; a generator can be exhibited by choosing a pair of dual bases $(e_i, f_j) = \delta_{ij}$ and setting $I := \sum_i e_i \otimes f_i$.

The map $\mathbb{C} \to V \otimes V$ given by $a \to aI$ is an elementary extension.

Remark that, using the scalar product on $V^{\otimes 2}$ since clearly $u := \sum_i (u, f_i) e_i = \sum_i (u, e_i) f_i$[10] we have

(32.1.4) $$(u_1 \otimes u_2, I) = (u_1, u_2).$$

One can easily extend these maps to general tensor powers and consider the contractions and extensions

(32.1.5) $$c_{ij} : V^{\otimes k} \to V^{\otimes k-2}, \ e_{ij} : V^{\otimes k-2} \to V^{\otimes k}$$

given by contracting in the indeces ij or inserting in the indeces i, j (e.g. $e_{13} : V \to V^{\otimes 3}$ is $v \to \sum_i e_i \otimes v \otimes f_i$.)

Notice in particular that

Remark. c_{ij} is surjective and e_{ij} is injective.

$c_{ij} = c_{ji}, \ e_{ij} = e_{ji}$ in the symmetric case while $c_{ij} = -c_{ji}, \ e_{ij} = -e_{ji}$ in the skew symmetric case.

In order to keep some order in these maps it is useful to consider the two symmetric groups $S_h, \ S_k$ which act on the two tensor powers commuting with the group G and as orthogonal transformations.

Thus $S_h \times S_k$ acts on $\hom_G(V^{\otimes h}, V^{\otimes k})$ with $(\sigma, \tau)A := \tau A \sigma^{-1}$.

We also have an action of $S_h \times S_k$ on the space $[(V^{\otimes h+k})^*]G$ of multilinear invariants by including $S_h \times S_k \subset S_{h+k}$.

We need to show that the identification $\psi : \hom_G(V^{\otimes h}, V^{\otimes k}) = [(V^{\otimes h+k})^*]G$ is $S_h \times S_k$ equivariant.

The formula $\psi_A(X \otimes Y) := (A(X), Y)$ gives

(32.1.6)
$$((\sigma, \tau)\psi_A)(X \otimes Y) := \psi_A(\sigma^{-1} \otimes \tau^{-1} X \otimes Y) = (A(\sigma^{-1} X), \tau^{-1} Y) = (\tau A(\sigma^{-1} X), Y)$$

as required.

[10]In the skew case $u := \sum_i (u, f_i) e_i = -\sum_i (u, e_i) f_i$

Consider now a multilinear monomial in the elements (u_i, u_j), (v_h, v_k), (u_p, v_q). In this monomials the $h + k = 2n$ elements $(u_1, \dots, u_h, v_1, \dots, v_k)$ appear each once and the monomial itself is described by the combinatorics of the n pairings.

Suppose we have exactly a pairings of type (u_i, v_j)[11] Then $h - a$, $k - a$ are both even and the remaining pairings are all *homosexual*.

It is clear that, under the action of the group $S_h \times S_k$, this invariant can be brought to the following canonical form

$$(32.1.7) \quad J_a := (u_1, v_1)(u_2, v_2) \dots (u_a, v_a) \prod_{i=1}^{(h-a)/2} (u_{a+2i-1}, u_{a+2i}) \prod_{j=1}^{(k-a)/2} (v_{a+2j-1}, v_{a+2j}).$$

Lemma. *The invariant J_a correspondes to the intertwiner*

$$(32.1.8) \qquad C_a : u_1 \otimes u_2 \otimes \cdots \otimes u_h \to \prod_{i=1}^{(h-a)/2} (u_{a+2i-1}, u_{a+2i}) u_1 \otimes \cdots \otimes u_a \otimes I^{\otimes (k-a)/2}$$

Proof. We compute explicitely

$$\Big(\prod_{i=1}^{(h-a)/2} (u_{a+2i-1}, u_{a+2i}) u_1 \otimes \cdots \otimes u_a \otimes I^{\otimes (k-a)/2}, v_1 \otimes v_2 \otimes \cdots \otimes v_k \Big) = J_a$$

from 32.1.4. $\square$

We have seen that there are no intertwiners between $V^{\otimes h}$, $V^{\otimes k}$ if $h+k$ is odd but there are injective G equivariant maps $V^{\otimes h-2a} \to V^h$ for all $a \leq h/2$.

In particular we can define the subspace $T^0(V^h)$ as the sum of all the irreducible representations of G which do not appear in the lower tensor powers $V^{\otimes h-2a}$ we claim.

Theorem. *The space $T^0(V^h)$ is the intersection of all the kernels of the maps c_{ij}; it is called the space of* **traceless tensors**.

Proof. If an irreducible representation M of G appears both in $V^{\otimes h-2a}$, $a > 0$ and in V^h by semisimplicity an isomorphism between these two submodules can be extended to a non 0 intertwiner between $V^h, V^{\otimes h-2a}$.

From the general formula of these intertwiners they all vanish on $T^0(V^h)$ and so all the irreducible representations in $T^0(V^h)$ do not appear in $V^{\otimes h-2a}$, $a > 0$. The converse is also clear: if a contraction does not vanish on an irreducible submodule N of $V^{\otimes h}$ then the image of N is an isomorphic submodule of $V^{\otimes h-2}$.

Thus we may say that $T^0(V^h)$ contains all the *new* representations of G in $V^{\otimes h}$. $\square$

[11] Weyl calls these pairings heterosexual and homosexual the others.

Proposition. *$T^0(V^h)$ is stable under the action of the symmetric group S_h which spans the centralizer of G in $T^0(V^h)$.*

Proof. Since clearly $\sigma c_{ij}\sigma^{-1} = c_{\sigma(i)\sigma(j)}, \forall i, j, \sigma \in S_h$ the first claim is clear.

$T^0(V^h)$ is a sum of isotypic components and we may study the restriction of the centralizer of G in $V^{\otimes h}$ to $T^0(V^h)$.

Then remark that the operators C_a (32.1.8) vanish on $T^0(V^h)$ as soon as $a > 0$, since up to multiplication by the symmetric group every intertwiner can be brought to this form the claim follows. □

We have thus again a situation similar to the one for the liner group except that the space $T^0(V^h)$ is a more complicated object to describe.

Our next task is to decompose

$$T^0(V^h) = \oplus_{\lambda\vdash h} U_\lambda \otimes M_\lambda$$

where the M_λ are the irreducible representations which are given by the theory of Young symmetrizers and the U_λ are the corresponding new representations of G, we have thus to discover which λ appear. In order to do this we will have to work out the second fundamental theorem.

§33 Highest weight theory.

33.1 The theory we are referring at is part of the theory of representations of complex semisimple Lie algebras and we shall give a short survey of its main features and illusatrate it for classical groups.[12]

Recall that, given a Lie algebra L and $x \in L$ one defines the linear operator $ad(x) : L \to L$ by $ad(x)(y) := [x, y]$.

An ideal I of L is a linear subspace stable under all the operators $ad(x)$, the quotient L/I is naturally a Lie algebra and the usual homomorphism theorems hold.

If L is finite dimensional one can compute $tr(ad(x)ad(y))$, the trace of the product of the two operators $ad(x)$, $ad(y)$; this constructs a symmetric bilinear form on L, defined by $(x, y) := tr(ad(x)ad(y))$ and called the Killing form. It has the following *invariance or associativity* property

$$([x, y], z) = (x, [y, z]).$$

Over the complex numbers there are several equivalent definitions of semisimple Lie algebra.

[12]There are many more general results over fields of characteristic 0 or just over the real numbers but they do not play a specific role in the theory we shall discuss

Definition. A complex Lie algebra L is simple if it has no non trivial ideals and it is not abelian.

For a Lie algebra L the following are equivalent (cf. [H]).

1 L is a direct sum of simple Lie algebras.
2 The Killing form $(x, y) := tr(ad(x)ad(y))$ is non degenerate.
3 L has no abelian ideals.

A Lie algebra satisfying the previous equivalent conditions is called *semisimple*.

Simple Lie algebras are completely classified and now very well understood, the key of the classification of Killing Cartan is the theory of roots.

It is a quite remarkable fact that associated to a continuous groups like a compact or a reductive algebraic group there is a finite group of Euclidean reflections and that the Theory of the continuous group can be largely analyzed in terms of the combinatorics of these reflections. We start now to give a brief sketch of this theory making it explicit for classical groups.

We have already discussed the role of maximal tori in algebraic groups. Each complex semisimple Lie algebra L is in fact the Lie algebra of a linearly reductive group G with finite center; in L we have special subalgebras called Cartan subalgebras, which are the Lie algebras of maximal tori in G.

For the classical Lie algebras we have:

(1) $sl(n, \mathbb{C})$ is the Lie algebra of the special Linear group $SL(n, \mathbb{C})$.

$sl(n, \mathbb{C})$ is the set of $n \times n$ complex matrices with trace 0; if $n > 1$ it is a simple algebra called *of type* A_{n-1}.

(2) $so(2n, \mathbb{C})$ is the Lie algebra of the special orthogonal group $SO(2n, \mathbb{C})$.

In order to describe it in matrix form it is convenient to choose an hyperbolic basis $e_1, \ldots, e_n, f_1, \ldots, f_n$ where the matrix of the form is $I_{2n} := \begin{pmatrix} 0 & 1_n \\ 1_n & 0 \end{pmatrix}$ and $so(2n, \mathbb{C}) := \{A \in M_{2n}(\mathbb{C}) | A^t I_{2n} = -I_{2n} A\}$. If $n > 3$ it is a simple algebra called *of type* D_n (for $n = 3$ we have the special isomorphism $so(6, \mathbb{C}) = sl(4, \mathbb{C})$).

(3) $so(2n+1, \mathbb{C})$ is the Lie algebra of the special orthogonal group $SO(2n+1, \mathbb{C})$.

In order to describe it in matrix form it is convenient to choose an hyperbolic basis $e_1, \ldots, e_n, f_1, \ldots, f_n, u$ where the matrix of the form is $I_{2n+1} := \begin{pmatrix} 0 & 1_n & 0 \\ 1_n & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $so(2n+1, \mathbb{C}) := \{A \in M_{2n+1}(\mathbb{C}) | A^t I_{2n+1} = -I_{2n+1} A\}$. If $n > 0$ it is a simple algebra called *of type* B_n *for* $n > 1$ (for $n = 1$ we have the special isomorphism $so(3, \mathbb{C}) = sl(2, \mathbb{C})$).

(4) $sp(2n, \mathbb{C})$ is the Lie algebra of the symplectic group $Sp(2n, \mathbb{C})$.

In order to describe it in matrix form it is convenient to choose a symplectic basis $e_1, \ldots, e_n, f_1, \ldots, f_n$ where the matrix of the form is $J := \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}$ and

$sp(2n,\mathbb{C}) := \{A \in M_{2n}(\mathbb{C}) | A^t J = -JA\}$. If $n > 0$ it is a simple algebra called *of type C_n for $n > 1$* (for $n = 1$ we have the special isomorphism $sp(2,\mathbb{C}) = sl(2,\mathbb{C})$, for $n = 2$ the isomorphism $sp(4,\mathbb{C}) = so(5,\mathbb{C})$ hence $B_2 = C_2$).

No further isomorphisms arise between these algebras.

The list of all complex simple Lie algebras is completed by adding the 5 exceptional types, called $G_2,\ F_4,\ E_6,\ E_7,\ E_8$.

The reason to choose these special bases is that in these bases it is easy to describe a Cartan subalgebra and the corresponding theory of roots.

In a Lie algebra L an element x is called *semisimple* if $ad(x)$ is a semisimple (i.e. diagonalizable) operator.

A toral subalgebra $\mathfrak{t}$ of L is a Lie subalgebra made only of semisimple elements. It turns out that toral subalgebras are abelian and thus the semisimple operators $ad(x),\ x \in \mathfrak{t}$ are simultaneously diagonalizable.

In a semisimple Lie algebra L a maximal toral subalgebra $\mathfrak{t}$ is called a *Cartan subalgebra.*

Given a Cartan subalgebra $\mathfrak{t}$ of a semisimple Lie algebra L, this has always a positive dimension r called the *rank* of L. When we decompose L into eigenspaces with respect to $\mathfrak{t}$ we find a finite set $\Phi \subset \mathfrak{t}^* - \{0\}$ of non 0 linear forms called *roots.*

For each $\alpha \in \Phi$ we have a 1-dimensional subspace $L_\alpha \subset L$ such that

$$L_\alpha := \{x \in L | \, [h,x] = \alpha(h)x,\ \forall h \in \mathfrak{t}\}, \quad L = \mathfrak{t} \oplus_{\alpha\in\Phi} L_\alpha$$

Moreover we have $\mathfrak{t} := L_0,\ = \{x \in L | \, [h,x] = 0,\ \forall h \in \mathfrak{t}\}$.

The non zero elements of L_α are called *root vectors* relative to α.

The invariance of the Killing form and the fact that it is non degenerate implies immediately that:

The Killing form restricted to $\mathfrak{t}$ is non degenerate. L_α, L_β are orthogonal unless $\alpha+\beta = 0$.

One usually identifies $\mathfrak{t}$ with its dual $\mathfrak{t}^*$ using the Killing form and then one can transport the Killing form to $\mathfrak{t}^*$. The next fact is that the Killing form computed on a pair of roots is always a rational number; the rational subspace $V := \sum_{\alpha\in\Phi} \mathbb{Q}\alpha$ is such that $\mathfrak{t}^* = V \otimes_{\mathbb{Q}} \mathbb{C}$ and moreover the Killing form restricted to V is positive definite.

One basic reduction step, in the Theory of semisimple algebras, is the introduction of $sl(2,\mathbb{C})$ triples, the fact is that, for each $\alpha \in \Phi$ we have that $[L_\alpha, L_{-\alpha}]$ is non 0 and so 1-dimensional, moreover one can choose non zero elements $e_\alpha \in L_\alpha,\ f_\alpha \in L_{-\alpha}$ such that, setting $h_\alpha := [e_\alpha, f_\alpha]$ we have the canonical commutation relations of $sl(2,\mathbb{C})$:

$$h_\alpha := [e_\alpha, f_\alpha],\ [h_\alpha, e_\alpha] = 2e_\alpha,\ [h_\alpha, f_\alpha] = -2f_\alpha$$

The way in which Φ sits in the Euclidean space $V_{\mathbb{R}} := V \otimes_{\mathbb{Q}} \mathbb{R}$ can be axiomatized giving rise to the abstract notion of *root system.*

Definition. Given a euclidean space E (with positive scalar product (u, v)) and a finite set Φ of non zero vectors in E we say that Φ is a reduced root system if:

(1) The elements of Φ span E.
(2) $\forall \alpha \in \Phi, s \in \mathbb{R}$ we have $s\alpha \in \Phi$ if and only if $s = \pm 1$.
(3) The numbers

$$< \alpha|\beta >:= \frac{2(\alpha, \beta)}{(\beta, \beta)}$$

are integers (called Cartan integers).

(4) For every $\alpha \in \Phi$ consider the reflection $r_\alpha : x \to x - \frac{2(x,\alpha)}{(\alpha,\alpha)}\alpha$ then $r_\alpha(\Phi) = \Phi$.

The dimension of E is also called the rank of the root system.

The fourth axiom implies that the subgroup of orthogonal transformations generated by the reflections r_α is a finite group (identified to the group of permutations that it induces on Φ).

This group is usually indicated by W and called the *Weyl group*, it is a basic group of symmetries similar in some way to the symmetric group with its canonical permutation representation.

It is in fact possible to describe it also in terms of the associated algebraic or compact group as in Theorem 16.2.

A root system Φ is called reducible if we can divide it as $\Phi = \Phi^1 \cup \Phi^2$ into mutually orthogonal subsets otherwise it is irreducible.

An isomorphism between two root systems Φ_1, Φ_2 is a 1-1 correspondence between the two sets of roots which preserves the Cartan integers.

It can be easily verified that any isomorphism between two irreducible root systems is induced by the composition of an isometry of the ambient Euclidean spaces and a omothety.

The next key to the theory of roots and of Lie algebras is to introduce *simple roots.* For this call a vector $v \in E$ *regular* if v is not orthogonal to any $\alpha \in \Phi$. Thus the regular vectors form the complement in E of the union of the *root hyperplanes* $H_\alpha := \{x \in E| \, (x, \alpha) = 0.\}$.

This set is clearly open and it is a finite union of connected components.

For every regular vector v we can decompose Φ into 2 parts

$$\Phi^+ := \{\alpha \in \Phi|(v, \alpha) > 0\}, \ \Phi^- := \{\alpha \in \Phi|(v, \alpha) < 0\}$$

from the axioms and the definition of regular vector it follows that $\Phi^- = -\Phi^+, \ \Phi = \Phi^+ \cup \Phi^-$.

One of the main ideas is to consider a root $\alpha \in \Phi^+$ as *decomposable* if $\alpha = \beta + \gamma, \ \beta, \gamma \in \Phi^+$¡ then

Theorem. *The set Δ of indecomposable roots in Φ^+ form a basis of E and every element of Φ^+ is a linear combination of Δ with non negative integer coefficients.*

Δ is called the set of simple roots (associated to Φ^+).

Given the set of simple roots $\alpha_1, \ldots, \alpha_n$ (of rank n) the matrix $C := (c_{ij})$ with entries the Cartan integers $< \alpha_i | \alpha_j >$ is called the Cartan matrix of the root system.

One can also characterize root systems by the property of the Cartan matrix. This is an integral matrix A which satisfies the following properties

(1) $a_{ii} = 2, \ a_{ij} \leq 0, \text{ if } i \neq j, \quad \text{if } a_{ij} = 0 \text{ then } a_{ji} = 0.$
(2) It is symmetrizable, i.e. there is a diagonal matrix D with entries positive integers d_i such that DC is symmetric.
(3) DC is positive definite.

If the root system is irreducible we have a corresponding irrreducibility property for C: one cannot reorder the rows and columns (with the same permutation) and make C in block form $C = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}$.

In the Theory of Kac-Moody algebras one considers Cartan matrices which satisfy only the first property: there is a rich theory which we shall not discuss.

For the applications to Lie algebras, having given the simple roots $\alpha_1, \ldots, \alpha_n$ one can choose $e_i \in L_{\alpha_i}, \ f_i \in L_{-\alpha_i}$ so that $e_i, f_i, h_i := [e_i, f_i]$ are $sl(2, \mathbb{C})$ triples and moreover $[h_i, e_j] :=< \alpha_i | \alpha_j > e_j, \ [h_i, f_j] := - < \alpha_i | \alpha_j > f_j$ and one obtains a fundamental Theorem of Chevalley Serre (use the notation $a_{ij} =< \alpha_i | \alpha_j >$).

Theorem. *The Lie algebra L is generated by the $3n$ elements $e_i, f_i, h_i, \ i = 1, \ldots, n$.*

The defining relations are

$$\begin{aligned} [h_i, h_j] &= 0 \\ [h_i, e_j] &= a_{ij} e_j \\ [h_i, f_j] &= a_{ij} f_j \\ [e_i, f_j] &= \delta_{ij} h_i \\ ad(e_i)^{1-a_{ij}}(e_j) &= 0 \\ ad(f_i)^{1-a_{ij}}(f_j) &= 0 \end{aligned}$$

In the classical groups the description of these objects is easy and follows the lines of Chapter 2.

(1) $sl(n, \mathbb{C})$ a Cartan algebra is formed by the diagonal matrices $h := \sum_{i=1}^n \alpha_i e_{ii}, \ \sum_i \alpha_i = 0$. The spaces L_α are the 1-dimensional spaces generated by the root vectors $e_{ij}, \ i \neq j$ and

$$[h, e_{ij}] = (\alpha_i - \alpha_j) e_{ij}$$

thus the linear forms

$$\sum_{i=1}^n \alpha_i e_{ii} \to \alpha_i - \alpha_j$$

are the roots of $sl(n, \mathbb{C})$.

(2) $so(2n,\mathbb{C})$, in block form a matrix $A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ satisfies $A^t I_{2n} = -I_{2n} A$ if and only if $d = -a^t$ and b, c are skew symmetric.

A Cartan subalgebra is formed by the diagonal matrices

$$h := \sum_{i=1}^{n} \alpha_i (e_{ii} - e_{n+i,n+i})$$

Root vectors are the elements

$$e_{ij} - e_{n+j,n+i},\ i \neq j \leq n,\ e_{i,n+j} - e_{j,n+i},\ \ i \neq j \leq n,\ e_{n+i,j} - e_{n+j,i}\ i \neq j \leq n$$

with roots

$$\alpha_i - \alpha_j,\ \alpha_i + \alpha_j,\ -\alpha_i - \alpha_j,\ i \neq j \leq n$$

(3) $so(2n+1,\mathbb{C})$, in block form a matrix $A := \begin{pmatrix} a & b & e \\ c & d & f \\ m & n & p \end{pmatrix}$ satisfies $A^t I_{2n+1} = -I_{2n+1} A$ if and only if $d = -a^t$, b, c are skew symmetric, $p = 0,\ n = -e^t,\ m = -f^t$.

A Cartan subalgebra is formed by the diagonal matrices

$$h := \sum_{i=1}^{n} \alpha_i (e_{ii} - e_{n+i,n+i})$$

Root vectors are the same as for $so(2n,\mathbb{C})$ plus the vectors

$$e_{i,2n+1} - e_{2n+1,i+n}, e_{n+i,2n+1} - e_{2n+1,i},\ i = 1, \dots, n$$

with roots

$$\pm\alpha_i,\ i = 1, \dots, n$$

(4) $sp(2n,\mathbb{C})$, in block form a matrix $A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ satisfies $A^t J_{2n} = -J_{2n} A$ if and only if $d = -a^t$ and b, c are symmetric.

A Cartan subalgebra is formed by the diagonal matrices

$$h := \sum_{i=1}^{n} \alpha_i (e_{ii} - e_{n+i,n+i})$$

Root vectors are the elements

$$e_{ij} - e_{n+j,n+i},\ i \neq j \leq n,\ e_{i,n+j} + e_{j,n+i},\ \ i, j \leq n,\ e_{n+i,j} + e_{n+j,i}\ i, j \leq n$$

with roots

$$\alpha_i - \alpha_j,\ \alpha_i + \alpha_j,\ -\alpha_i - \alpha_j,\ i \neq j,\ \pm 2\alpha_i,\ i = 1, \dots, n$$

One of the main tools of root theory is deduced by the decomposition of Φ into two parts $\Phi : + = \Phi^+ \cup \Phi^-$ the *positive and negative roots* one has $\Phi^- = -\Phi^+$.

In the examples we have that for $sl(n, \mathbb{C})$ the positive roots are the elements $\alpha_i - \alpha_j,\ i < j$ the corresponding root vectors $e_{ij} <\ \ i < j$ span the Lie subalgebra of strictly upper triangular matrices, similarly for negative roots

$$\mathfrak{u}^+ := \oplus_{i<j} \mathbb{C} e_{ij},\ \mathfrak{u}^- := \oplus_{i>j} \mathbb{C} e_{ij}$$

The simple roots and the root vectors associated to simple roots are

$$\alpha_i - \alpha_{i+1},\ e_{i,i+1}$$

For $so(2n, \mathbb{C})$ we set

$$\Phi^+ := \alpha_i - \alpha_j, i < j,\ \alpha_i + \alpha_j,\ i \neq j$$

The simple roots and the root vectors associated to simple roots are

$$\alpha_i - \alpha_{i+1},\ \alpha_{n-1} + \alpha_n,\ e_{i,i+1} - e_{n+i+1,n+i},\ e_{n-1,2n} - e_{n,2n-1}$$

$$\mathfrak{u}^+ := \oplus_{i<j} \mathbb{C}(e_{ij} - e_{n+j,n+i}) \oplus_{i \neq j, \leq n} \mathbb{C}(e_{i,n+j} - e_{j,n+i}),$$
$$\mathfrak{u}^- := \oplus_{i>j} \mathbb{C}(e_{ij} - e_{n+j,n+i}) \oplus_{i \neq j, \leq n} \mathbb{C}(e_{n+i,j} - e_{n+j,i})$$

$so(2n+1, \mathbb{C})$ we set

$$\Phi^+ := \alpha_i - \alpha_j,\ \alpha_i + \alpha_j,\ i < j \leq n,\ \alpha_i$$

The simple roots and the root vectors associated to simple roots are

$$\alpha_i - \alpha_{i+1},\ \alpha_n,\ e_{i,i+1} - e_{n+i+1,n+i},\ e_{n,2n+1} - e_{2n+1,2n}$$

$$\mathfrak{u}^+ := \oplus_{i<j} \mathbb{C}(e_{ij} - e_{n+j,n+i}) \oplus_{i \neq j} \mathbb{C}(e_{i,n+j} - e_{j,n+i}) \oplus_{i=1}^{n} \mathbb{C}(e_{i,2n+1} - e_{2n+1,i}),$$
$$\mathfrak{u}^- := \oplus_{i>j} \mathbb{C}(e_{ij} - e_{n+j,n+i}) \oplus_{i \neq j} \mathbb{C}(e_{n+i,j} - e_{n+j,i}) \oplus_{i=n+1}^{2n} \mathbb{C}(e_{i,2n+1} - e_{2n+1,i})$$

For $sp(2n, \mathbb{C})$ we set

$$\Phi^+ := \alpha_i - \alpha_j,\ \alpha_i + \alpha_j,\ i < j, 2\alpha_i$$

The simple roots and the root vectors associated to simple roots are

$$\alpha_i - \alpha_{i+1}, 2\alpha_n, \quad e_{i,i+1} - e_{n+i+1,n+i},\ e_{n,2n} + e_{2n,n}$$

$$\mathfrak{u}^+ := \oplus_{i<j} \mathbb{C}(e_{ij} - e_{n+j,n+i}) \oplus_{i,j \leq n} \mathbb{C}(e_{i,n+j} + e_{j,n+i}),$$
$$\mathfrak{u}^- := \oplus_{i>j} \mathbb{C}(e_{ij} - e_{n+j,n+i}) \oplus_{i \neq j} \mathbb{C}(e_{n+i,j} + e_{n+j,i})$$

For all these (as well as for all semisimple) Lie algebras we have the direct sum decomposition (as vector space)

$$L = \mathfrak{u}^+ \oplus \mathfrak{t} \oplus \mathfrak{u}^-$$

One sets

$$\mathfrak{b}^+ := \mathfrak{u}^+ \oplus \mathfrak{t},\ \mathfrak{b}^+ := \mathfrak{u}^- \oplus \mathfrak{t}$$

these are called two opposite Borel subalgebras.

33.2 Weights in representations, highest weight theory.

For a semisimple Lie algebra L all finite dimensional representations are completely reducible and have bases formed of weight vectors under the Cartan subalgebra $\mathfrak{t}$.

The weights one gets are elements of a particular lattice Λ of $\mathfrak{t}^*$ called the *weight lattice*, it can be described as follows.

$$\Lambda := \{\lambda \in \mathfrak{t}^* | < \lambda, \alpha >= \frac{2(\lambda, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z},\ \forall \alpha \in \Phi\}$$

One also sets

$$\Lambda^+ := \{\lambda \in \Lambda | < \lambda, \alpha > \geq 0,\ \forall \alpha \in \Phi^+\}$$

Λ^+ is called the *set of dominant weights.*

In particular we can determine uniquely elements $\omega_i \in \Lambda$ with $< \omega_i, \alpha_j >= \delta_{ij},\ \forall \alpha_j \in \Delta$ and prove easily that

$$\Lambda^+ = \{\sum_{i=1}^n m_i \omega_i,\ m_i \in \mathbb{N}\}$$

$(\omega_1, \omega_2, \dots, \omega_n)$ is called the *set of fundamental weights* (relative to Φ^+).

We should make some basic remarks, in the case of classical groups, relative to weights for the Cartan subalgebra and for the maximal torus.

Starting with $GL(n, \mathbb{C})$ with its maximal torus T made of diagonal matrices with non zero entries x_i we have described its Lie algebras as the diagonal matrices with entries α_i we then have.

Given a rational representation of $GL(n, \mathbb{C})$, a vector v is a weight vector for T if and only if it is a weight vector for $\mathfrak{t}$ and the two weights are related as

$$\prod_{i=1}^n x_i^{m_i},\ \sum_{i=1}^n m_i \alpha_i$$

similarly for the other classical groups. In fact from $x_i = e^{\alpha_i}$ the claim follows. We rite then the maximal weigths lor the Lie algebra annd the group.

For $sl(n+1)$

$$\sum_{i \leq k} \alpha_i,\ i \leq n,\ \prod_{i \leq k} x_i$$

For $so(2n)$ the weights

$$\sum_{i\leq k}\alpha_i,\ i\leq n-2,\ \frac{1}{2}(\sum_{i=1}^{n-1}\alpha_i\pm\alpha_n)$$

This shows already that the last two weights do not exponentiate to integral weights of the maximal torus of $SO(2n,\mathbb{C})$ the reason is that there is a double covering of this group the Spin group which possesses these two representations which do not factor through $SO(2n,\mathbb{C})$.

For $so(2n+1)$ the weights

$$\sum_{i\leq k}\alpha_i,\ i\leq n-1,\ \frac{1}{2}(\sum_{i=1}^{n}\alpha_i)$$

Same discussion of the Spin group.

For $sp(2n)$

$$\sum_{i\leq k}\alpha_i,\ i\leq n$$

Theorem. *Given a finite dimensional irreducible representation M of a semisimple Lie algebra L the space of vectors $\{m\in M|\ \mathfrak{u}^+m=0\}$ is 1 dimensional and a weight space under $\mathfrak{t}$, the weight is called the highest weight of M.*

This subspace is the unique 1-dimensional subspace of M stable under the Borel subalgebra $\mathfrak{b}^+$.

Two irreducible representations of L are isomorphic if and only if they have the same highest weight.

The set of highest weights coincides with the set of dominant weights.

It is quite important to remark that

Proposition. *If $u\in M,\ v\in N$ are two highest weight vectors of weight $\lambda,\ \mu$ respectively, then $u\otimes v$ is a highest weight vector of weight $\lambda+\mu$.*

Proof. We have $e(u\otimes v)=eu\otimes v+u\otimes ev$ for every element of the Lie algebra from which the claim follows. $\square$

33.3 Examples for representations of classical groups.

In all the 3 cases we start with a vector space V with a given basis which in the orthogonal and symplectic case is respectively hyperbolic or symplectic. We shall use then the notations $e_1,\dots,e_n,\ e_1,\dots,e_n,f_1,\dots,f_n$ etc. as in Inserire

The first and most basic example, which in fact needs to be developed first when studying the theory which has been summarized in the previous section, is $sl(2,\mathbb{C})$ for which one takes the usual basis

$$e:=\begin{matrix}0&1\\0&0\end{matrix},\ f:=\begin{matrix}0&0\\1&0\end{matrix},\ h:=\begin{matrix}1&0\\0&-1\end{matrix}$$

these elements satisfy the commutation relations

$$[e,f]=h,\ [h,e]=2e,\ [h,f]=-2f.$$

We know from the standard theory developed up to now that the symmetric powers $S^k(V)$ of the 2-dimensional vector space V are irreducible representations of $sl(2,\mathbb{C})$, it is not hard to prove that these exhaust all possible irreducible representations, we sketch the standard proof.

Let N be an irrreducible representation. One starts choosing some eigenvector v for h of some eigenvalue a, one immediately sees that

$$2ev=(he-eh)v=hev-aev,\ hev=(a+2)ev$$

hence ev is an eigenvector of weight $a+2$, hence proceding this way we find an eivenvector v_0 of weight a_0 and such that $ev_0=0$.

Next we consider the elements

$$v_i:=\frac{1}{i!}f^iv_0=\frac{1}{i}fv_{i-1}$$

and, arguing as before we see that they are eigenvectors for h of weight a_0-2i.

From $[e,f]=h$ se see that $ev_i=(a_0-2i)v_i+fev_{i-1}$ and thus, recursively we check that $ev_i=(a_0-i+1)v_{i-1}$.

Finally since we are assuming that the representation is finite dimensional we must have $v_{k+1}=0$ for some minimal k hence $0=ev_{k+1}=(a_0-k)v_k$ which implies $a_0=k$ and $\dim N=k+1$ has as basis the vectors $v_i,\ i=0,\dots,k$ with the explicit action

$$hv_i=(k-2i)v_i,\ fv_i=(i+1)v_{i+1},\ ev_i=(k-i+1)v_{i-1}$$

Call this representation V_k, we identify the basic vector space V with N_1 and next we can identify N_k with $S^k(V)$ noticing that, if V has basis w_0,w_1 the elements $\binom{k}{i}w_0^{k-i}w_1^i$ behave as the elements v_i under the action of the elements e,f,h.

For $sl(n,\mathbb{C})$ we wish to let to the reader to verify that the representation on a tensor power associated to a partition $\lambda:=h_1,h_2,\dots,h_n$ and dual partition $1^{a_1}2^{a_2}\dots n^{a_n}$ has as highest weight vector the vector a_TU of 24.1 and this has weight

$$\omega_\lambda\sum_{j=1}^{n-1}a_j\omega_j \tag{33.3.1}$$

Set $V=\mathbb{C}^n$. The proof is reduced, by the last proposition of the previous paragraph, to remark that the vector $e_1\wedge e_2\wedge\cdots\wedge e_k$ is a highest weight vector of the exterior power $\wedge^kV$ of weight ω_k and that by construction a_TU is the tensor product of these highest weight vectors in $V^{\otimes a_1}\otimes(\wedge^2V)^{\otimes a_2}\otimes\dots(\wedge^nV)^{\otimes a_n}$ but by definition of special linear group $\wedge^nV$ is the trivial representation of $sl(n,\mathbb{C})$.

One shoud remark that, when we take a $2n$ dimensional symplectic space with basis $e_1,\dots,e_n,f_1,\dots,f_n$ the elements $e_1\wedge e_2\wedge\cdots\wedge e_k,\ k\le n$ are still highest weight vectors of weight ω_i but, as we shall see in §35.6, the exterior powers are no more irreducible, for the orthogonal Lie algebras besides the exterior powers we need to discuss the spin representations.

33.4 Highest weights and invariants under U, geometric examples from Cauchy formulas.

We pass now to groups; the Lie algebras $\mathfrak{u}^+$, $\mathfrak{b}^+$, $\mathfrak{t}$ are Lie algebras of three algebraic subgroups of the algebraic group called U^+, B^+, T and one has that $B^+ = U^+T$ as semidirect product (in particular $T = B^+/U^+$).

As an example in $SL(n, \mathbb{C})$ we have that T is the subgroup of diagonal matrices, B^+ the subgroup of upper triangular matrices and U^+ the subgroup of strictly upper triangular matrices, that is the upper triangular matrices with 1 on the diagonal or what is the same with all eigenvalues 1 (unipotent elements).

For an irreducible representation of G the highest weight vector of weight λ is the unique (up to scalars) vector invariant under U^+ and it is an eigenvector under T, we denote its eigenvalue with the same symbol λ. Remark that, if we define character of a group G, any homomorphism of G to the multiplicative group $\mathbb{C}^*$ any algebraic character of B^+ is trivial on U^+ (cf.) and it is just induced by an (algebraic) character of T.

We shall use the word character, for a torus T to mean an algebraic character.

The geometric interpretation of highest weights is obtained as follows.

Consider an action of a reductive group G on an affine algebraic variety V, let $A[V]$ be the coordinate ring of V which is a rational representation under the induced action of G.

Thus $A[V]$ can be decomposed into a direct sum of irreducible representations.

If $f \in A[V]$ is a highest weight vector of some weight λ we have for every $b \in B^+$ that $bf = \lambda(b)f$ and conversely a function f with this property is a highest weight vector.

Notice that if f_1, f_2 are highest weight vectors of weight λ_1, λ_2 then $f_1 f_2$ is a highest weight vector of weight $\lambda_1 + \lambda_2$.

Now unless f is constant the set

$$S_f := \{x \in V \mid f(x) = 0\}$$

is a hypersurface of V and it is clearly stable under B^+, conversely if V satisfies the property that every hypersurface is defined by an equation (for instance if V is an affine space) we have the converse fact that to a B^+ stable hypersurface is associated a highest weight vector.

Of course, as usual in algebraic geometry, this correspondence is not bijective but we have to take into consideration multiplicities,

Lemma. *If $A[V]$ is a unique factorization domain then we have a 1-1 correspondence between irreducible B^+ stable hypersurfaces and irreducible (as polynomials) highest weight vectors.*

Proof. It is enough to show that, if $f = \prod_i g_i$ is a highest weight vector factored into irreducible polynomials then the g_i are also highest weight vectors.

For this take an element $b \in B^+$ we have $f = (bf) = \prod_i (bg_i)$, since B^+ acts as a group of automorphisms the bg_i are irreducible and thus the elements bg_i must equal the g_j up to some possible permutation and some scalars. Since the action of B^+ on $A[V]$ is rational there is a B^+ stable subspace $U \subset A[V]$ containing the elements g_i.

Consider the induced action of B^+ on the projective space of lines of U. By assumption the lines through the elements g_i are permuted by B^+. Since B^+ is connected the only possible algebraic actions of B^+ on a finite set are trivial and it follows that the g_i are eigenvectors under B^+ from which one deduces, from the previous remarks, that they are U^+ invariant and $bg_i = \chi_i(b)g_i$ where χ_i are characters of B and in fact are dominant weights. $\square$

We want to illustrate this point in 3 classical examples showing how this bears on representation theory.

1. The space $M_{n,m}(\mathbb{C})$ of $n \times m$ matrices with action of $GL(n,\mathbb{C}) \times GL(m,\mathbb{C})$ given by $(A,B)X := AXB^{-1}$.

2. The space $M_n^+(\mathbb{C})$ of symmetric $n \times n$ matrices with action of $GL(n,\mathbb{C})$ given by AXA^t.

3. The space $M_n^-(\mathbb{C})$ of skew symmetric $n \times n$ matrices with action of $GL(n,\mathbb{C})$ given by AXA^t.

We leave to the reader the simple

Exercise In each case there are only finitely many orbits under the given group action, two matrices are in the same orbit if and only if they have the same rank.

If V_k denotes the matrices of rank k (in all cases) we have that the closure $\overline{V}_k$ is

(33.4.1) $$\overline{V}_k := \cup_{j \leq k} \overline{V}_j.$$

This implies that the varieties $\overline{V}_k$ are the only varieties invariant under the given group action, they are irreducible and thus the ideals defining them are the only invariant ideals equal to their radical and they are all prime ideals. We shall see how this is interpreted for the second fundamental Theorem of invariant Theory in §34.

More interesting is the fact that there are also finitely many orbits under the action of a Borel subgroup, we will not compute all the orbits but restrict to analyze the invariant hypersurfaces.

1. To distinguish between the two groups we let $T(n), T(m), U^+(n), U^+(m), \ldots$ etc. denote the torus unipotent etc of the two groups.

We take as Borel subgroup the subgroup $B(n)^- \times B(m)^+$ of pairs (A,B) where A is a lower and B is an upper triangular matrix.

Given a matrix $X \in M_{n,m}(\mathbb{C})$ the matrix AXB^{-1} is obtained from X by elementary row and column operations of the following types:

a) multiply a row or a column by a non zero scalar.

b) Add to the i^{th} row the j^{th} row, with $j < i$, multiplied by some number.

c) Add to the i^{th} column the j^{th} column, with $j < i$, multiplied by some number.

This is the usual Gaussian elimination on rows and columns of X without performing any exchange.

The usual remark about these operations is that, for every $k \leq n, m$ they do not change the rank of the $k \times k$ minor X_k of X extracted from the first k rows and the first k columns. Moreover if we start from a matrix X with the property that, for every $k \leq n, m$ we have $\det(X_k) \neq 0$ then the standard algorithm of Gaussian elimination proves that, under the action of $B(n)^- \times B(m)^+$ this matrix is equivalent to the matrix I with entries 1 on the diagonal and 0 elsewhere. We deduce

Theorem. *The open set of matrices $X \in M_{n,m}(\mathbb{C})$ with $\det(X_k) \neq 0$ form a unique orbit under the group $B(n)^- \times B(m)^+$.*

The only $B(n)^- \times B(m)^+$ stable hypersurfaces of $M_{n,m}(\mathbb{C})$ are the ones defined by the equations $\det(X_k) = 0$ which are irreducible.

Proof. We have already remarked that the first part of the Theorem is the content of Gaussian elimination as for the second since the complement of this open orbit is the union of the hypersurfaces of equations $\det(X_k) = 0$ it is clearly enough to prove that these equations are irreducible.

In order to do this we apply the previous Lemma, if $\det(X_k)$ is not irreducible it is a product of highest weight vectors g_i and its weight is the sum of the weights of the g_i which are dominant weights. If we prove that the weight of $d_k := \det(X_k)$ is a fundamental weight we are done once we remark that the only polynomials which belong to the 0 weight are constant.

Thus we compute the weight of $\det(X_k)$ one can do it of course directly but we can also remember the computation of 27.4, in any case given a pair $D_1, D_2 \in T(n) \times T(m)$ of diagonal matrices with entries x_i, y_j respectively we have

$$(D_1, D_2)d_k(X) := d_k(D_1^{-1} X D_2) = \prod_{i=1}^{k} x_i^{-1} y_i d_k(X) \tag{33.4.2}$$

hence the weight of d_k is $\prod_{i=1}^{k} x_i^{-1} \prod_{i=1}^{k} y_i$ which is the highest weight of $(\wedge^k \mathbb{C}^n)^* \otimes \wedge^k \mathbb{C}^m$ a fundamental weight. $\square$

2. and 3. are treated as follows. One thinks of a symmetric or skew symmetric matrix as the matrix of a form.

Again we choose as Borel subgroup $B(n)^-$ and Gaussian elimination is the algorithm of putting the matrix of the form in normal form by a triangular change of basis.

The generic orbit is obtained when the given form has maximal rank on all the subspaces spanned by the first k basis vectors $k \leq n$, which in the symmetric case means that the form is non degenerate on these subspaces while in the skew case it means that the form is non degenerate on the even dimensional subspaces.

On symmetric matrices this is essentially the Gram Schmidt algorithm we get that

Theorem. *The open set of symmetric matrices $X \in M_n^+(\mathbb{C})$ with $\det(X_k) \neq 0$ form a unique orbit under the group $B(n)^-$.*

The only $B(n)^-$ stable hypersurfaces of $M_n^+(\mathbb{C})$ are the ones defined by the equations $s_k := \det(X_k) = 0$ which are irreducible.

Proof. Here we can procede as in the linear case except at one point, when we arrive at computing the character of s_k we discover that it is $\prod_{i=1}^k x_i^{-2}$ that is it is twice a fundamental weight.

Hence a priori we could have $s_k = ab$ with a, b with weight $\prod_{i=1}^k x_i^{-1}$, to see that this is not possible set the variables $x_{ij} = 0,\ i \neq j$ getting $s_k = \prod_{i=1}^k x_{ii}$ hence a, b should specialyze to two factors of $\prod_{i=1}^k x_{ii}$, but clearly these factors never have as weight $\prod_{i=1}^k x_i^{-1}$ hence s_k is irreducible. $\square$

3. We choose as Borel subgroup $B(n)^-$ and perform Gaussian elimination on skew symmetric matrices. In this case for every k for which $2k \leq n$ consider the minor X_{2k}, the condition that this skew symmetric matrix be non singular is that the Pfaffian $p_k := Pf(X_{2k})$ is non zero.

Theorem. *The open set of skew symmetric matrices $X \in M_n^-(\mathbb{C})$ with $Pf(X_{2k}) \neq 0$ form a unique orbit under the group $B(n)^-$.*

The only $B(n)^-$ stable hypersurfaces of $M_n^-(\mathbb{C})$ are the ones defined by the equations $p_k := Pf(X_{2k}) = 0$ which are irreducible.

Proof. If $Pf(X_{2k}) \neq 0$ for all k we easily see that we can construct in a triangular form a symplectic basis for the matrix X hence the first part, for the rest again it suffices to prove that the polynomials p_k are irreducible. In fact we can compute their weight which, by the formula 12.9.11, is $\prod_{i=1}^{2k} x_i^{-1}$ a fundamental weight. $\square$

We should deduce now the Cauchy formulas which are behind this Theory.

We do it in the symmetric and skew symmetric case which are the new formulas.

In the first case we have that the full list of highest weight vectors is the set of monomials $\prod_{k=1}^n s_k^{m_k}$ with weight $\prod_{k=1}^n \prod_{i=1}^k x_i^{-2m_k}$.

If we denote by V the fundamental representation of $GL(n, \mathbb{C})$ we have that

$$\prod_{k=1}^n \prod_{i=1}^k x_i^{-2m_k} = \prod_{i=1}^n x_i^{-\sum_{i \leq k} 2m_k}$$

is the highest weight of $S_\lambda(V)^*$ where λ is the partition $\lambda_k = 2\sum_{i \leq k} m_k$.

In the skew case we obtain the monomials $\prod_{2k\leq n} p_k^{m_k}$ with weight $\prod_{i=1}^n x_i^{-\sum_{i\leq 2k} m_k}$.

We deduce that

Theorem. *As representation of* $GL(V)$ *the ring* $S[S^2[V]]$ *decomposes*

$$S[S^2[V]] := \oplus_{\lambda=2h_1,2h_2,\ldots,2h_n} S_\lambda(V) \tag{33.4.3}$$

If $\dim(V) = 2n,\ 2n+1$ *as representation of* $GL(V)$ *the ring* $S[\wedge^2[V]]$ *decomposes*

$$S[\wedge^2[V]] := \oplus_{\lambda=h_1,h_1,h_2,h_2,\ldots,h_n,h_n} S_\lambda(V) \tag{33.4.4}$$

Proof. We have that the space of symmetric forms on V is, as representation, $S^2[V]^*$ and the action is in matrix notations, given by $(A, X) \to (A^{-1})^t X A^{-1}$ so $S[S^2[V]]$ is the ring of polynomials on this space and we can apply the previous Theorem. Similar considerations for the skew case. $\square$

In more pictorial language thinking of λ as the shape of a diagram we can say that, the diagrams appearing in $S[S^2[V]]$ have even columns while the ones appearing in $S[\wedge^2[V]]$ have even rows.

It is then convenient to have a short notation for these concepts we shall write

$$\lambda \vdash^{er} n,\ \lambda \vdash^{ec} n$$

to express the fact that it has even rows, resp. even columns.

We should remark that the previous Theorem corresponds to identities of characters.

According to formula 27.4.3, given a linear operator A on a vector space U its action on the symmetric algebra has as graded character $\frac{1}{det(1-tA)}$ if $e_1, \ldots, e_n$ is a basis of V and X is the matrix $Xe_i = x_i e_i$ we have that $e_i e_j,\ i \leq j$ is a basis of $S^2[V]$ and $e_i \wedge e_j,\ i < j$ a basis of $\wedge^2[V]$ from 27.4.3 and the previous Theorem we therefore deduce

$$\frac{1}{\prod_{i\leq j}(1 - x_i x_j)} = \sum_m \sum_{\lambda \vdash^{ec} m} S_\lambda(x_1, \ldots, x_n) \tag{33.4.5}$$

$$\frac{1}{\prod_{i<j}(1 - x_i x_j)} = \sum_m \sum_{\lambda \vdash^{er} m} S_\lambda(x_1, \ldots, x_n)$$

The formulas we have developed are special types of Plethysm formulas, which are any type of formulas which can describe the composition $S_\lambda(S_\mu(V))$ of Schur functors, these formulas although always theoretically computable are quite formidable and not at all understood.

§34 The second fundamental theorem

34.1 We need to discuss now the relations among invariants. We shall take a geometric approach reserving a combinatorial approach in the section on tableaux.

The study of relations among invariants procedes as follows. Express the first fundamental theorem in matrix form, deduce the prime ideal theory from the highest weight theory.

For the general linear group we had described invariant theory through the multiplication map $f : M_{p,m} \times M_{m,q} \to M_{p,q},\ f(X, Y) := YX$

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

We have also remarked that, by elementary linear algebra, the multiplication map f has, as image, the subvariety 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 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 $M_{p,q}$.

For the group $O(n, \mathbb{C})$ we have considered the space of $n \times m$ matrices Y with the action of $O(n, \mathbb{C})$ given by multiplication XY then the mapping $Y \to Y^t Y$ from the space of $n \times m$ matrices to the symmetric $m \times m$ matrices of rank $\leq n$ is a quotient under the orthogonal group. Again the determinants of the $m + 1 \times m + 1$ minors of these matrices define this subvariety set theoretically.

Similarly for the symplectic group we had considered the space of $2n \times m$ matrices Y with the action of $Sp(2n, \mathbb{C})$ given by multiplication XY then the mapping $Y \to Y^t J_n Y$ from the space of $2n \times m$ matrices to the antisymmetric $m \times m$ matrices of rank $\leq 2n$ is a quotient under the symplectic group.

In this case the correct relations are not the determinants of the minors but rather the Pfaffians of the principal minors of order $2(n + 1)$.

We have thus identified three types of *determinantal varieties* for each of which we want to determine the ideal of relations.

We will make use of the Plethysm formulas developed in the previous section.

According to 33.4 we know that the determinantal varieties are the only varieties invariant under the appropriate group action, according to the matrix formulation of the first fundamental Theorem they are also the varieties which have ring of invariants as coordinate rings, we want to describe the ideals of definition and their coordinate rings as representations.

We should recall the theory of 27.4.7 and complement it with the theory for symmetric and skew symmetric case.

According to 27.4.7 given two vector spaces V, W we have the decomposition

$$\mathcal{P}(\hom(V, W)) = \oplus_\lambda S_\lambda(W^*) \otimes S_\lambda(V)$$

Moreover if we think of $\mathcal{P}(\hom(V,W))$ as the polynomial ring $\mathbb{C}[x_{ij}]$ the subspace D_k of it spanned by the determinants of the minors of order k of the matrix $X := (x_{ij})$ is identified with the subrepresentation $D_k = \wedge^k V \otimes \wedge^k W^*$. We define

$$I_k := \mathcal{P}(\hom(V,W))D_k \tag{34.1.1}$$

to be the determinantal ideal generated by the determinants of all the $k \times k$ minors of X.

Consider now $\mathcal{P}(S^2(V)^*)$ as the polynomial ring $\mathbb{C}[x_{ij}],\ x_{ij} = x_{ji}$ let again $X := (x_{ij})$ be a symmetric matrix of variables.

We want to see how to identify, in the same lines as 27.4.7, the subspace J_k of $\mathcal{P}(S^2(V)^*)$ spanned by the determinants of the minors of order k of the matrix X.

We know that, given a symmetric form A on V it induces a symmetric form on $V^{\otimes m}$ and thus by restriction a symmetric form, which we will denote by $S_\lambda(A)$ on $S_\lambda(V)$, thus an element of $S^2(S_\lambda(V)^*)$.

According to the theory

we have that $S_{2\lambda}(V)^*$ appears with multiplicity 1 in $S^2(S_\lambda(V)^*)$ and we claim that the dual map

$$S_{2\lambda}(V) \to S^2(S_\lambda(V)) \to \mathcal{P}(S^2(V)^*),\ \phi \to < \phi | S_\lambda(A) >$$

is non 0 and hence it identifies $S_{2\lambda}(V)$ with its corresponding factor in the decomposition 33.4.3.

In particular when we apply this to $\wedge^k A$ we see that if $A = (a_{ij})$ is the matrix of the given form in a basis $e_1, \dots, e_n$ the matrix of $\wedge^k A$ in the basis $e_{i_1} \wedge e_{i_2} \wedge_{i_k}$ is given by the formula

$$\wedge^k A(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}, e_{j_1} \wedge e_{j_2} \wedge \cdots \wedge e_{j_k}) = \det(a_{i_r, j_s}),\ r, s = 1, \dots, k \tag{34.1.2}$$

the determinant 34.1.2 is the determinant of the minor extracted from the rows $i_1, \dots, i_k$ and the columns $j_1, \dots, j_k$ of A.

We have thus naturally defined a map $S^2(\wedge^k V) \to \mathcal{P}(S^2(V)^*)$ with image the space J_k (spanned by the determinants of the minors of order k).

We want to prove in fact that $J_k = S_{2^k}(V)$ is irreducible with highest weight vector s_k.

To see this we need to analyze the decomposition of $S^2(\wedge^k V)$ into irreducible representations and show that the only irreducible representation in this decomposition which belongs to partitions with even columns is $S_{2^k}(V)$. This follows from the formula of section 32,

inserire

Consider now $\mathcal{P}(\wedge^2(V)^*)$ as the polynomial ring $\mathbb{C}[x_{ij}],\ x_{ij} = -x_{ji}$ let again $X := (x_{ij})$ be a skew symmetric matrix of variables.

We carry out a similar analysis for the subspace P_k of $\mathcal{P}(\wedge^2(V)^*)$ spanned by the Pfaffians of the principal minors of order $2k$ of the matrix X.

In this case the analysys is simpler the extrior power map $\wedge^2 V^* \to \wedge^{2k} V^*,\ A \to A^k$ gives for $A = \sum_{i<j} a_{ij} e_i \wedge e_j$ that

$$A^k = \sum_{i_1<i_2<\ldots i_{2k}} [i_1, i_2, \ldots, i_{2k}] e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_{2k}} \tag{34.1.3}$$

where $[i_1, i_2, \ldots, i_{2k}]$ denotes the Pfaffian of the principal minor of A extracted from the row and column indeces $i_1 < i_2 < \ldots i_{2k}$. One deduces immediately by duality the required map

$$\wedge^{2k} V \to S[\wedge^2[V]]$$

with image the space P_k.

According to the results of the previous section and the discussion we have just performed we claim

Second fundamental Theorem.

(1) *The only $GL(V) \times GL(W)$ invariant prime ideals in $\mathcal{P}(\hom(V, W))$ are the ideals I_k. As representations we have that (using 27.4.7)*

(34.1.4)
$$I_k = \oplus_{\lambda, ht(\lambda) \ge k} S_\lambda(W^*) \otimes S_\lambda(V),\ \mathcal{P}(\hom(V, W))/I_k = \oplus_{\lambda, ht(\lambda) < k} S_\lambda(W^*) \otimes S_\lambda(V)$$

(2) *The only $GL(V)$ invariant prime ideals in $\mathcal{P}(S^2(V)^*)$ are the determinantal ideals*

$$I_k^+ := \mathcal{P}(S^2(V)^*) J_k$$

generated by the determinants of the $k \times k$ minors of the symmetric matrix X. As representations we have that

$$I_k^+ = \oplus_{\lambda \vdash ^{ec}, ht(\lambda) \ge k} S_\lambda(V),\ \mathcal{P}(S^2(V)^*)/I_k^+ = \oplus_{\lambda \vdash ^{ec}, ht(\lambda) < k} S_\lambda(V) \tag{34.1.5}$$

(3) *The only $GL(V)$ invariant prime ideals in $\mathcal{P}(\wedge^2(V)^*)$ are the Pfaffian ideals*

$$I_k := \mathcal{P}(\wedge^2(V)^*) P_k$$

generated by the Pfaffians of the principal $2k \times 2k$ minors. As representations we have that

$$I_k^- = \oplus_{\lambda \vdash ^{er}, ht(\lambda) \ge 2k} S_\lambda(V),\ \mathcal{P}(\wedge^2(V)^*)/I_k^- = \oplus_{\lambda \vdash ^{er}, ht(\lambda) < 2k} S_\lambda(V) \tag{34.1.6}$$

Proof. In each of the 3 cases we know that the set of highest weights is the set of monomials M in a certain number of elements x_i highest weights of certain representations N_i.

The fact is that every invariant subspace J is identified by the set $I \subset M$ of highest weight vectors that it contains and, if J is an ideal, the set I is closed under multiplication by M, moreover if it is a prime ideal it follows that, if a monomial $\prod x_i^{m_i} \in I$ then at least one x_i appearing with non 0 exponent must be in I, it follows then that the ideal J is necessarily generated by the subspaces N_i which are contained in it.

Now to conclude it is enough to remark, in the case of determinantal ideals that for every k we have

(34.1.7) $$D_{k+1} \subset I_k,\ J_{k+1} \subset I_k^+,\ P_{k+1} \subset I_k^-.$$

in fact the first two statements follow immediately by a row or column expansion of a determinant of a $k+1 \times k+1$ minor in terms of determinants of $k \times k$ minors, as for the Pfaffians recall the formula 34.1.3 from which we deduce a similar Laplace expansion for Pfaffians

$$A^{k+1} = \sum_{i_1 < i_2 < \dots i_{2(k+1)}} [i_1, i_2, \dots, i_{2(k+1)}] e_{i_1}, \wedge e_{i_2} \wedge \cdots \wedge e_{i_{2(k+1)}} =$$
$$A \wedge \sum_{i_1 < i_2 < \dots i_{2k}} [i_1, i_2, \dots, i_{2k}] e_{i_1}, \wedge e_{i_2} \wedge \cdots \wedge e_{i_{2k}}$$

From which we deduce the typical Laplace expansion

$$[1, 2, \dots, 2(k+1)] = \sum_{i<j} (-1)^{i+j-1} a_{ij} [1, 2, \dots, i-1, i+1, \dots, j-1, j+1, \dots, 2(k+1)]$$

therefore in each case the ideal is generated by D_k, J_k^+, P_k^- for the minimal index k for which this space is contained in the ideal. □

§35 The second fundamental theorem for intertwiners

35.1 We want to apply the results of the previous section to intertwiners, we need first some remarks of general nature, let us go back to the formula 27.4.7 $\mathcal{P}[\hom(V, W)] = S[W^* \otimes V] = \oplus_\lambda S_\lambda(W^*) \otimes S_\lambda(V)$ and let us apply it to $\dim V = n,\ W = \mathbb{C}^n$.

Using the standard basis e_i of $\mathbb{C}^n$ we identify $V^{\otimes n}$ with a suitable subspace of $S[\mathbb{C}^n \otimes V]$ by encoding the element $\prod_{i=1}^n e_i \otimes v_i$ with $v_1 \otimes v_2 \otimes \cdots \otimes v_n$.

The previous mapping identifies $V^{\otimes n} \subset S[\mathbb{C}^n \otimes V]$ as weight space under the torus of diagonal matrices X with entries x_i, and $Xe_i = x_i e_i$ for which

$$X \prod_{i=1}^n e_i \otimes v_i = \prod_i x_i \prod_{i=1}^n e_i \otimes v_i$$

we claim that

Lemma. *We can identify* $V^{\otimes n} := \{u \in S[\mathbb{C}^n \otimes V] | Xu = \prod_i x_i u\}$ *and the symmetric group* $S_n \subset GL(n, \mathbb{C})$ *of permutation matrices acts on* $V^{\otimes n}$ *with the usual permutation action.*

Proof. We are implicitely using the action of X on the first factor of $\mathbb{C}^n \otimes V$ and leave the proof to the reader. $\square$

Let us denote by $\chi := \prod_{i=1}^n x_i$ this is a character of the diagonal torus T invariant under the symmetric group (and generates the group of these characters), 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

$$S_\lambda(\mathbb{C}^n)^\chi := \{u \in S_\lambda(\mathbb{C}^n) | Xu = \prod_i x_i u, \ \forall X \in T\} \tag{35.1.1}$$

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 $S_\lambda(\mathbb{C}^n)^0$, 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

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

hence

$$V^{\otimes n} := S^n(\mathbb{C}^n \otimes V)^\chi = \oplus_{\lambda \vdash n} S_\lambda((\mathbb{C}^n)^*)^\chi \otimes S_\lambda(V) = \oplus_{\lambda \vdash n} M_\lambda \otimes S_\lambda(V) \tag{35.1.2}$$

we get the required identification. $\square$

Now we apply this to various examples, first we want to identify the group algebra $\mathbb{C}[S_n]$ with the spcace P^n of polynomials in the variables x_{ij} wich are multilinear in rigth and left indeces, that is to say we consider the span P^n of those monomials $x_{i_1,j_1} x_{i_2,j_2} \dots x_{i_n,j_n}$ such that both $i_1, i_2, \dots, i_n$ and $j_1, j_2, \dots, j_n$ are permutations of $1, 2, \dots, n$, of course a monomial of this type can be uniquely displayed as

$$x_{1,\sigma^{-1}(1)} x_{2,\sigma^{-1}(2)} \dots x_{n,\sigma^{-1}(n)} = x_{\sigma(1),1} x_{\sigma(2),2} \dots x_{\sigma(n),n}$$

which defines the required map

$$\Phi : \sigma \to x_{\sigma(1),1} x_{\sigma(2),2} \dots x_{\sigma(n),n} \tag{35.1.3}$$

Remark that this space of polynomials is a weight space with respect to the product $T \times T$ of two maximal tori of diagonal matrices under the induced $GL(n, \mathbb{C}) \times GL(n, \mathbb{C})$ action on $\mathbb{C}[x_{ij}]$, let us call by χ_1, χ_2 the to weights.

Remark also that this mapping is equivariant under the left and rigth action of S_n on $\mathbb{C}[S_n]$ which correspond respectively to

$$x_{i,j} \to x_{\sigma(i),j}, \ x_{i,j} \to x_{i,\sigma(j)} \tag{35.1.4}$$

We now fix an $m-$dimensional vector space V and recall the basic symmetry homomorphism

(35.1.5) $$i_n : \mathbb{C}[S_n] \to End(V^{\otimes n})$$

recall that we also have a homomorphism given by the F.F.T. with the notations of 22.4 $f : \mathbb{C}[x_{ij}] \to \mathbb{C}[< \alpha_i | v_j >]$ mapping P^n onto the space of multilinear invariants of n covector and n vector variables which we will denote by $\mathcal{I}_n$ spanned by the elements $\prod_{i=1}^n < \alpha_{\sigma(i)} | v_i >$. Finally we have the canonical isomorphism $j : \sigma \to \prod_{i=1}^n < \alpha_{\sigma(i)} | v_i >$ (cf. 22.3.2). We have a commutative diagram

(35.1.6)
$$\begin{array}{ccc} \mathbb{C}[S_n] & \xrightarrow{\Phi} & P_n \\ {\scriptstyle i_n}\downarrow & & {\scriptstyle f}\downarrow \\ End_{GL(V)} V^{\otimes n} & \xrightarrow{j} & \mathcal{I}_n \end{array}$$

from which we deduce that the Kernel of i_n can be identified, via the map Φ, with the intersection of P_n with the determinantal ideal I_{m+1} in $\mathbb{C}[x_{ij}]$.

A determinant $(i_1, i_2, \ldots, i_{m+1} | j_1, j_2, \ldots, j_{m+1})$ of $X = (x_{ij})$, multiplied by any monomial in the x_{ij}, is clearly a weight vector for $T \times T$.

In order to get a weight vector in P_n we must consider the products

(35.1.7) $$(i_1, i_2, \ldots, i_{m+1} | j_1, j_2, \ldots, j_{m+1}) x_{i_{m+2}, j_{m+2}} \ldots x_{i_n, j_n}$$

with $i_1, i_2, \ldots, i_n$ and $j_1, j_2, \ldots, j_n$ both permutations of $1, 2, \ldots, n$.

Theorem. *Under the isomorphism Φ the space $P_n \cap I_{m+1}$ corresponds to the two sided ideal of $\mathbb{C}[S_n]$ generated by the element*

(35.1.8) $$A_{m+1} := \sum_{\sigma \in S_{m+1}} \epsilon_\sigma \sigma$$

antisymmetrizer on $m+1$ elements (chosen arbitrarily from the given n elements).

Proof. From 35.1.4 it follows that the element corresponding to a typical element of type 35.1.7 is of the form $\sigma A \tau^{-1}$ where A corresponds to $(1, 2, \ldots, m+1 | 1, 2, \ldots, m+1) x_{m+2,m+2} \ldots x_{n,n} = \sum_{\sigma \in S_{m+1}} \epsilon_\sigma x_{\sigma(1),1} x_{\sigma(2),2} \ldots x_{\sigma(m+1),m+1}$ which clearly corresponds to A_{m+1}. $\square$

Now we want to recover in a new form the result of 24.1.1.

First recall that, as any group algebra, $\mathbb{C}[S_n]$ decomposes into the sum of its minimal ideals corresponding to irreducible representations

$$\mathbb{C}[S_n] = \oplus_{\lambda \vdash n} M_\lambda^* \otimes M_\lambda$$

We first decompose

$$\mathbb{C}[x_{ij}] = \oplus_\lambda S_\lambda(\mathbb{C}^n)^* \otimes S_\lambda(\mathbb{C}^n), \; I_{m+1} = \oplus_{ht\lambda} S_\lambda(\mathbb{C}^n)^* \otimes S_\lambda(\mathbb{C}^n)$$

then we pass to the weight space

(35.1.9)

$$P^n = \mathbb{C}[x_{ij}]^{\chi_1,\chi_2} = \oplus_\lambda S_\lambda(\mathbb{C}^n)^{*\,\chi_1} \otimes S_\lambda(\mathbb{C}^n)^{\chi_2},$$

$$P^n \cap I_{m+1} = \oplus_{ht\lambda\geq m+1} S_\lambda(\mathbb{C}^n)^{*\,\chi_1} \otimes S_\lambda(\mathbb{C}^n)^{\chi_2} = \oplus_{\lambda\vdash n,\, ht\lambda\geq m+1} M_\lambda^* \otimes M_\lambda$$

As a consequence the image of $\mathbb{C}[S_n]$ in $End(V^{\otimes n})$ is $\oplus_{\lambda\vdash n,\, ht\lambda\leq m} M_\lambda^* \otimes M_\lambda$ as in 24.1.1.

35.2 Now we pass to the orthogonal and symplectic group.

First of all we formulate some analogue of P^n.

In both the symmeric or skew symmetric case we take a matrix $Y = (y_{ij})$ of even size $2n$ and consider the span of the *multilinear monomials*

$$y_{i_1,j_1} y_{i_2,j_2} \cdots y_{i_n,j_n}$$

such that $i_1, i_2, \ldots, i_n, j_1, j_2, \ldots, j_n$ is a permutation of $1, 2, \ldots, 2n$.

This is again the weight space of the weight χ of the diagonal group T of $GL(2n, \mathbb{C})$.

Although at first sigth the space looks the same in the two cases it is not so and we will call these two cases P_+^{2n}, P_-^{2n}.

In fact the symmetric group S_{2n} acts on the polynomial ring $\mathbb{C}[y_{ij}]$ (as subgroup of $GL(2n, \mathbb{C})$) by $\sigma(y_{ij}) = y_{\sigma(i)\sigma(j)}$.

It is clear that S_{2n} acts transitively on the given set of multilinear monomials.

First of all we want to understand it as representation on $P_\pm^n$ and for this we consider the special monomial

$$M_0 := y_{1,2} y_{3,4} \cdots y_{2n-1,2n}.$$

Let H_n be the subgroup of S_{2n} which fixes the partition of $\{1, 2, \ldots, 2n\}$ formed by the n sets with 2 elements $\{2i-1, 2i\}$, $i = 1, \ldots, n$. Clearly $H_n = S_n \ltimes \mathbb{Z}/(2)^n$ is the obvious semidirect product where S_n acts in the diagonal way on odd and even numbers $\sigma(2i-1) = 2\sigma(i) - 1$, $\sigma(2i) = 2\sigma(i)$ and $\mathbb{Z}/(2)^n$ is generated by the transpositions $(2i-1, 2i)$.

In either case H_n is the stabilyzer of the line through M_0. In the symmetric case H_n fixes M_0 while in the skew symmetric case it induces on this line the sign of the permutation (remark that $S_n \subset S_{2n}$ is made of even permutations).

We deduce

Proposition. *P_+^n, as representation of S_{2n} coincides with the permutation representation $Ind_{H_n}^{S_{2n}}\mathbb{C}$ associated to S_{2n}/H_n.*

P_-^n, as representation of S_{2n} coincides with the representation $Ind_{H_n}^{S_{2n}}\mathbb{C}(\epsilon)$ induced to S_{2n} from the sign representation $\mathbb{C}(\epsilon)$ of H_n.

Next one can describe these representations in terms of irreducible representations using the formulas 33.4.3,4 getting

Theorem. *As representation of S_{2n} the space P_+^n decomposes*

$$P_+^n = S[S^2[V]]^\chi := \oplus_{\lambda \vdash^{ec} 2n} S_\lambda(V)^\chi = \oplus_{\lambda \vdash^{ec} 2n} M_\lambda \tag{35.2.1}$$

P_-^n decomposes

$$P_-^n = S[\wedge^2[V]]^\chi := \oplus_{\lambda \vdash^{er} 2n} S_\lambda(V)^\chi = \oplus_{\lambda \vdash^{er} 2n} M_\lambda \tag{35.2.2}$$

Remark. By Frobenius reciprocity this theorem computes the multiplicity of the trivial and the sign representation of H_n in any irreducible representation of S_{2n}, both appear with multiplicity at most 1 and in fact the trivial appears when λ has even columns and the sign when it has even rows.

Next we apply the same ideas as in 35.1 to intertwiners, we do it only in the simplest case.

35.3 1. Orthogonal case. Let V be an orthogonal space of dimension m and consider the space $\mathcal{I}_{2n}^+$ of multilinear invariants in $2n$ variables $u_i \in V,\ i = 1, \ldots, 2n$.

$\mathcal{I}_{2n}^+$ is spanned by the monomials $(u_{i_1}, u_{i_2})(u_{i_3}, u_{i_4}) \ldots (u_{i_{2n-1}}, u_{i_{2n}})$ where $i_1, i_2, \ldots, i_n, j_1, j_2, \ldots, j_n$ is a permutation of $1, 2, \ldots, 2n$.

Let $y_{ij} = y_{ji}$ be symmetric variables.

Under the map

$$\mathbb{C}[y_{ij}] \to \mathbb{C}[(u_i, u_j)],\ y_{ij} \to (u_i, u_j)$$

the space P_+^n maps surjectively onto $\mathcal{I}_{2n}^+$ with kernel $P_+^n \cap I_{m+1}^+$.

The same proof as in 35.1 shows that

Theorem. *As representation of S_{2n} we have*

$$P_+^n \cap I_{m+1}^+ = \oplus_{\lambda \vdash^{ec} 2n,\ ht(\lambda) \geq m+1} M_\lambda,\ \mathcal{I}_{2n}^+ = \oplus_{\lambda \vdash^{ec} 2n,\ ht(\lambda) \leq m} M_\lambda$$

The interpretation of the relations in the algebras of intertwiners $End_{O(V)} V^{\otimes n}$ is more complicated and we shall not explicit it in full.

1. Symplectic case. Let V be an symplectic space of dimension $2m$ and consider the space $\mathcal{I}_{2n}^-$ of multilinear invariants in $2n$ variables $u_i \in V,\ i = 1, \ldots, 2n$.

$\mathcal{I}_{2n}^-$ is spanned by the monomials $[u_{i_1}, u_{i_2}][u_{i_3}, u_{i_4}] \ldots [u_{i_{2n-1}}, u_{i_{2n}}]$ where $i_1, i_2, \ldots, i_n, j_1, j_2, \ldots, j_n$ is a permutation of $1, 2, \ldots, 2n$.

Let $y_{ij} = -y_{ji}$ be antisymmetric variables.

Under the map

$$\mathbb{C}[y_{ij}] \to \mathbb{C}[(u_i, u_j)],\ y_{ij} \to [u_i, u_j]$$

the space P_-^n maps surjectively onto $\mathcal{I}_{2n}^-$ with kernel $P_-^n \cap I_{m+1}^+$.

The same proof as in 35.1 shows that

Theorem. *As representation of S_{2n} we have*

$$P_-^n \cap I_{m+1}^- = \oplus_{\lambda \vdash^{er} 2n,\ ht(\lambda) \geq 2m+2} M_\lambda,\ \mathcal{I}_{2n}^- = \oplus_{\lambda \vdash^{ec} 2n,\ ht(\lambda) \leq 2m} M_\lambda$$

the interpretation of the relations in the algebras of intertwiners $End_{Sp(V)} V^{\otimes n}$ is again more complicated and we shall not explicit it in full.

We want nevertheless explicit part of the structure of the intertwiners algebra, the one relative to traceless tensors $T^0(V^{\otimes n})$.

In both the orthogonal and symplectic case the idea is similar, let us first develop the symplectic case which is simpler.

Consider a monomial $M := y_{i_1,j_1} y_{i_2,j_2} \dots y_{i_n,j_n}$ we can assume by symmetry $i_k < j_k$ for all k, we see by the formula 32.1.8 that the operator $\phi_M : V^{\otimes n} \to V^{\otimes n}$ involves at least one contraction unless all the indeces i_k are $\leq n$ and instead $j_k > n$.

Let us dentote by $\overline{\phi}_M$ the restriction of the operator ϕ_M to $T^0(V^{\otimes n})$, we have seen that $\overline{\phi}_M = 0$ if the monomial contains a variable $y_{ij},\ y_{n+i,n+j},\ i,j \leq n$.

Thus $M \to \overline{\phi}_M$ map factors through the space $\overline{P}_-^n$ of monomials obtained setting to 0 one of the two previous types of variables.

The only monomials that remain are of type $M_\sigma := \prod_{i=1}^n y_{i,n+\sigma(i)}, \sigma \in S_n$ and M_σ corresponds to the invariant

$$\prod [u_i, u_{n+\sigma(i)}] = [u_{\sigma^{-1}(1)} \otimes u_{\sigma^{-1}(2)} \otimes \cdots \otimes u_{\sigma^{-1}(n)}, u_{n+1} \otimes u_{n+2} \otimes \cdots \otimes u_{2n}]$$

which corresponds to the map induced by the permutation σ on $V^{\otimes n}$.

We have just identified $\overline{P}_-^n$ to the group algebra $\mathbb{C}[S_n]$ and the map to

$$\rho : \overline{P}_-^n = \mathbb{C}[S_n] \to End_{Sp(V)}(T^0(V^{\otimes n})),\ \rho(M) := \overline{\phi}_M$$

to the canonical map to the algebra of operators induced by the symmetric group and, since $T^0(V^{\otimes m})$ is a sum of isotypic components the map ρ is surjective.

Let us identify the image of $P_-^n \cap I_{m+1}^-$ in $\overline{P}_-^n = \mathbb{C}[S_n]$ which is in the kernel of ρ.

Take the Pfaffian of the principal minor of Y of indeces $i_1, i_2, \dots, i_{2m+2}$ and evaluate after setting $y_{ij} = y_{n+i,n+j} = 0,\ i,j \leq n$ let us say that h of these indeces are $\leq n$ and $2m+2-h$ are $> n$.

The specialized matrix has block form $\begin{pmatrix} 0 & Z \\ -Z^t & 0 \end{pmatrix}$ and the minor extracted from the indeces $i_1, i_2, \dots, i_{2m+2}$ has a square block matrix of 0 of size the maximum between the numbers $h, 2m+2-h$.

Since the maximum dimension of an isotropic space for a non degenerate symplectic form on a $2m+2$ dimensional space is $m+1$ we deduce that the only case in which this Pfaffian can be non 0 is when $h = 2m+2-h = m+1$ and in this case the minor has also a block form $\begin{pmatrix} 0 & W \\ -W^t & 0 \end{pmatrix}$ and its Pfaffian equals up to sign $\det(Z)$.

Thus arguing as in the linear case we see that the image of $P_-^n \cap I_{m+1}^-$ in $\overline{P}_-^n = \mathbb{C}[S_n]$ is the ideal generated by the antisymmetrizer on $m+1$ elements.

Theorem. *The algebra* $End_{Sp(V)}(T^0(V^{\otimes n}))$ *equals the algebra* $\mathbb{C}[S_n]$ *modulo the ideal generated by the antisymmetrizer on* $m+1$ *elements.*

Proof. We have already seen that the given algebra is a homomorphic image of the group algebra od S_n modulo the given ideal. In order to prove that there are no further relations remark that, if $U \subset V$ is the subspace spanned by $e_1, \ldots, e_m$ it is a (maximal) isotropic subspace and thus $U^{\otimes n} \subset T^0(V^{\otimes n})$. On the other hand by the linear theory the kernel of the action of $\mathbb{C}[S_n]$ on $U^{\otimes n}$ coincides with the ideal generated by the antisymmetrizer on $m+1$ elements and the claim follows. $\square$

35.4 We are now ready to exhibit the list of irreducible rational representations of $SP(V)$.

First of all, using the double centralizer Theorem we have a decomposition

$$T^0(V^{\otimes n}) = \oplus_{\lambda \vdash n,\ ht(\lambda) \leq m} M_\lambda \otimes T_\lambda(V)$$

where we have indicated with $T_\lambda(V)$ the irreducible representation of $SP(V)$ paired to M_λ.

We should remark then that we can construct, as in 24.1 the tensor $a_T(e_1 \otimes e_2 \otimes \ldots \otimes e_{n_1}) \otimes (e_1 \otimes e_2 \otimes \ldots \otimes e_{n_2}) \otimes \ldots \otimes (e_1 \otimes e_2 \otimes \ldots \otimes e_{n_t}) \in T_\lambda(V)$ where the partition λ has rows $n_1, n_2, \ldots, n_t$.

We ask the reader to verify that it is a highest weight vector for $T_\lambda(V)$ with highest weight $\sum_{j=1}^t \omega_{n_j}$.

Since $SP(V)$ is contained in the special linear group, from Proposition 15.2 all irreducible representations appear in tensor powers of V, since $T^0(V^{\otimes m})$ contains all the irreducible representations appearing in $V^{\otimes n}$ and not in $V^{\otimes k},\ k < n$ we deduce

Theorem. *The irreducible representations* $T_\lambda(V), ht(\lambda) \leq m$ *are a complete list of inequivalent irreducible representations of* $SP(V)$.

From the computation of the highest weights and the basic theorem 33.2 it follows that this is also the list of irreducible representations of the Lie algebra $sp(2m, \mathbb{C})$.

35.5 We want to explicit also in the orthogonal case part of the structure of the intertwiners algebra, the one relative to traceless tensors $T^0(V^{\otimes n})$.

We let V be an $m-$dimensional orthogonal space.

Consider a monomial $M := y_{i_1,j_1} y_{i_2,j_2} \ldots y_{i_n,j_n}$ we can assume by symmetry $i_k < j_k$ for all k, we see by the formula 32.1.8 that the operator $\phi_M : V^{\otimes n} \to V^{\otimes n}$ involves at least one contraction unless all the indeces i_k are $\leq n$ and instead $j_k > n$.

Let us denote by $\overline{\phi}_M$ the restriction of the operator ϕ_M to $T^0(V^{\otimes n})$, we have seen that $\overline{\phi}_M = 0$ if the monomial contains a variable $y_{ij},\ y_{n+i,n+j},\ i,j \leq n$.

Thus $M \to \overline{\phi}_M$ map factors through the space $\overline{P}_+^n$ of monomials obtained setting to 0 one of the two previous types of variables.

The only monomials that remain are of type $M_\sigma := \prod_{i=1}^n y_{i,n+\sigma(i)}, \sigma \in S_n$ and M_σ corresponds to the invariant

$$\prod (u_i, u_{n+\sigma(i)}) = (u_{\sigma^{-1}(1)} \otimes u_{\sigma^{-1}(2)} \otimes \cdots \otimes u_{\sigma^{-1}(n)}, u_{n+1} \otimes u_{n+2} \otimes \cdots \otimes u_{2n})$$

which corresponds to the map induced by the permutation σ on $V^{\otimes n}$.

We have just identified $\overline{P}_+^n$ to the group algebra $\mathbb{C}[S_n]$ and the map to

$$\rho : \overline{P}_+^n = \mathbb{C}[S_n] \to End_{Sp(V)}(T^0(V^{\otimes n})),\ \rho(M) := \overline{\phi}_M$$

to the canonical map to the algebra of operators induced by the symmetric group and, since $T^0(V^{\otimes n})$ is a sum of isotypic components the map ρ is surjective.

Let us identify the image of $P_+^n \cap I_{m+1}^+$ in $\overline{P}_+^n = \mathbb{C}[S_n]$ which is in the kernel of ρ.

Take the determinant of an $m+1 \times m+1$ minor of Y extracted from the row indeces $i_1, i_2, \dots, i_{m+1}$ column indeces $j_1, j_2, \dots, j_{m+1}$ and evaluate after setting $y_{ij} = y_{n+i,n+j} = 0,\ i, j \leq n$ let us say that h of the row indeces are $\leq n$ and $m+1-h$ are $> n$ and also k of the column indeces are $\leq n$ and $m+1-k$ are $> n$.

The specialized matrix has block form $\begin{pmatrix} 0 & Z \\ W & 0 \end{pmatrix}$ where Z is an $h \times m+1-k$ and W an $m+1-h \times k$ matrix.

If this matrix has non 0 determinant the image of the first k basis vectors must be linearly independent hence $m+1-h \geq k$ and similarly $h \leq m+1-k$ hence $h+k = m+1$ or Z is square $h \times h$ and W is square $k \times k$ matrix.

Up to sign the determinant of this matrix is $\det(W)\det(Z)$.

This determinant is again a weight vector and, multiplied by a monomial in the y_{ij} can give rise to an element of P_+^n if and only if the indeces $i_1, i_2, \dots, i_{m+1}$ and $j_1, j_2, \dots, j_{m+1}$ are all distinct.

Up to a permutation of the indeces we may assume then that these two sets of indeces are $1, 2, \dots, h, n+h+1, n+h+2, \dots, n+h+k$ and $h+1, h+2, \dots, h+k, n+1, n+2, \dots, n+h$ so that using the symmetry $y_{n+h+i,h+j} = y_{h+j,n+h+i}$ we obtain

$$\det(W)\det(Z) =$$

$$\sum_{\sigma \in S_h} y_{\sigma(1),n+1} y_{\sigma(2),n+2} \cdots y_{\sigma(h),n+h} \sum_{\sigma \in S_k} y_{h+\sigma(1),n+h+1} y_{h+\sigma(2),n+h+2} \cdots y_{h+\sigma(k),n+h+k}$$

and we can take as multiple of this the one multiplied by $\prod_{t=1}^{n-h-k} y_{h+k+t,n+h+k+t}$ A variation of the argument of the linear case shows that this element corresponds in $\mathbb{C}[S_n]$ to the antisymmetrizer relative to the partition consisting of the parts $(1, 2, \dots, h), (h+1, \dots, h+k), \dots, t, \dots$ which is the element

$$\sum_{\sigma \in S_h \times S_k} \epsilon_\sigma \sigma$$

Since a determinant of a minor of order $> m+1$ can be developed into a combination of monomials times determinants of order $m+1$ we get that in the ideal of $\mathbb{C}[S_n]$ corresponding to $P_+^n \cap I^- + m + 1$ we have all the ideal generated by the products of two antisymmetrizers on two disjoint sets adding to $\geq m+1$ elements.

Now by the description of the Young symmetrizers it follows that each Young symmetrizer relative to a partition with the first two columns adding to a number $\geq m+1$ is in the ideal generated by such products and thus we see that the image of $P_+^n \cap I^- + m + 1$ in $\overline{P}_+^n = \mathbb{C}[S_n]$ contains the ideal generated by all the Young symmetrizers relative to diagrams with the first two columns adding to a number $\geq m+1$.

Theorem. *The algebra $End_{Sp(V)}(T^0(V^{\otimes n}))$ equals the algebra $\mathbb{C}[S_n]$ modulo the ideal generated by all the Young symmetrizers relative to diagrams with the first two columns adding to a number $\geq m+1$.*

Proof. We have already seen that the given algebra is a homomorphic image of the group algebra od S_n modulo the given ideal. In order to prove that there are no further relations we have to show that, if λ is a partition with the first two columns adding uo to at most m then we can find a non 0 traceless tensor u with $a_T u \neq 0$ where a_T is a Young symmetrizer of type λ.

For this we cannot argue as simply as in the simplectic case but need a variation of the theme.

First of all consider the diagram of λ filled in increasing order from up to down and rigth to left with the numbers $1, 2, \ldots, n$ e.g.

$$\begin{array}{cccc} 1 & 5 & 8 & 11 \\ 2 & 6 & 9 & \\ 3 & 7 & 10 & \\ 4 & & & \end{array}$$

next suppose we fill it as a tableau with some of the basis vectors e_i, f_j e.g.

$$\begin{array}{cccc} e_2 & f_3 & e_1 & e_1 \\ f_2 & e_4 & f_4 & \\ e_3 & e_3 & f_1 & \\ f_4 & & & \end{array}$$

and to such a display we associate a tensor in which we plce in the i^{th} position the vector placed in the tableau in the case labeled with i, e.g. in our previous example:

$$e_2 \otimes f_2 \otimes e_3 \otimes f_4 \otimes f_3 \otimes e_4 \otimes e_3 \otimes e_1 \otimes f_4 \otimes f_1 \otimes e_1$$

Assume first $m = 2p$ is even if the first column has $\leq p$ elements we can work in the subspace $e_1, \ldots, e_p$ as for the symplectic group.

Otherwise let $p+s, p-t,\ s \leq t$ be the lengths of the first two columns.

We first fill the diagram with e_i in the i^{th} row $i \leq p$ and we are left with s rows with just 1 element, we fill this with $f_p, f_{p-1}, \ldots, f_{p-s}$. This tensor is symmetric in the row positions.

When we perform a contraction on this tensor there are s possible contractions which are non 0 (in fact 1) and correspond to certain pairs of indeces of the first column occupied with pairs e_{p-i}, f_{p-i}, notice that if we exchange these two positions the contraction does not change.

It follows that, when we antisymmetrize this element we get a required non zero traceless tensor in M_λ.

The odd case is similar. $\square$

We are now ready to exhibit the list of irreducible rational representations of $SO(V)$.

First of all, using the double centralizer Theorem we have a decomposition

$$T^0(V^{\otimes n}) = \oplus_{\lambda\vdash n,\ h_1+h_2\leq m} M_\lambda \otimes T_\lambda(V)$$

where we have indicated with $T_\lambda(V)$ the irreducible representation of $O(V)$ paired to M_λ and h_1, h_2 the first two columns of λ.

Since the determinant representation of $O(V)$ is contained in $V^{\otimes m}$ is of order 2 and so equal to its inverse, from Proposition 15.2 all irreducible representations appear in tensor powers of V, since $T^0(V^{\otimes n})$ contains all the irreducible representations appearing in $V^{\otimes n}$ and not in $V^{\otimes k},\ k < n$ we deduce

Theorem. *The irreducible representations $T_\lambda(V), h_1 + h_2 \leq m$ are a complete list of inequivalent irreducible representations of $O(V)$.*

As for the special orthogonal group one should only verify to which diagram belongs $T_\lambda(V) \wedge^m V$ but clearly this belongs to the diagram in which to λ has been added a first column of length m.

We can now pass to $SO(V)$, in this case there is one more invariant $[v_1, \dots, v_m]$ which gives rise to new intertwiners.

First of all let us analyze, for $k \leq m$ the operator

$$* : \wedge^k V \to \wedge^{m-k} V$$

defined by the implicit formula (using the induced scalar product on tensor and exterior powers):

$$(*(v_1 \wedge \cdots \wedge v_k), v_{k+1} \wedge \cdots \wedge v_m) = [v_1, \dots, v_m]$$

In an orthonormal oriented basis u_i we have $*(u_{i_1} \wedge \cdots \wedge u_{i_k}) = \epsilon u_{j_1} \wedge \cdots \wedge u_{j_{m-k}}$ where $i_1, \dots, i_k, j_1, \dots, j_{m-k}$ is a permutation of $1, \dots, m$ and ϵ is the sign of this permutation.

In general let us at least understand what type of intertwiners we obtain on traceless tensors using this new invariant.

Typically we are reduced to study the map γ from $T^0(V^{k+p})$ to $T^0(V^{m-k+p})$ induced by the invariant

$$(\gamma(u_1 \otimes \cdots \otimes u_{p+k}), v_1 \otimes \cdots \otimes v_{p+m-k}) = \prod_{i=1}^{p} (u_i, v_i)[u_{p+1}, \dots, u_{p+k}, v_{p+1}, \dots, v_{p+m-k}$$

the reader can verify that (using the highest weight), this maps $T_\lambda(V)$ to $T_{\lambda'}(V)$ where λ' is obtained from λ subsituting the first column k_1 with $m - k_1$.

From the computation of the highest weights and the basic theorem 33.2 it follows that this is not the list of all the irreducible representations of the Lie algebra $so(m, \mathbb{C})$ the reason is that one has to consider spinorial representations.

35.6 We give here a small complement to the previous theory by analizing the action of the symplectic or orthogonal group of a space V on the exterior algbra.

We start with the symplectic case which in some way is more interesting.

First of all, by Theorem 35.3 we have that the traceless tensors $T^0(V^{\otimes m})$ contain a representation associated to the full antisymmetrizer if and only if $m \leq n$ let us denote by $\wedge^{m,0}(V)$ this representation which, by what we have just seen appears with multiplicity 1 in $\wedge^m(V)$.

Remark next that, by definition, $\wedge^2 V$ contains a canonical bivector

$$\psi := \sum_{i=1}^{n} e_i \wedge f_i$$

invariant under $Sp(V)$.

We want to compute now the $Sp(V)$ equivariant maps between $\wedge^k V,\ \wedge^h V$.

Since the skew symmetric tensors are diect summands in tensor space any $Sp(V)$ equivariant maps between $\wedge^k V,\ \wedge^h V$ can be decomposed as $\wedge^k V \xrightarrow{i} V^{\otimes k} \xrightarrow{p} V^{\otimes h} \xrightarrow{A} \wedge^h V$ where, i is the cannical inclusion, f is some equivariant map and A is the antisymmetrizer.

Now we have seen in 32.1 how to describe equivariant maps $V^{\otimes k} \xrightarrow{p} V^{\otimes h}$ up to the symmetric group action.

If we apply the symmetric groups to i or to A it changes at most the sign, in particular we see that the insertion maps $\wedge^k V \to \wedge^{k+2} V$ can, up to sign be identified with $u \to \psi \wedge u$ which we shall also call ψ, while also for the contraction we have a unique map (up to constant) which can be normalized to

$$(35.6.1) \qquad c : v_1 \wedge v_2 \wedge \cdots \wedge v_k \to \sum_{i<j} (-1)^{i+j-1} < v_i, v_j > v_1 \wedge v_2 \wedge \check{v}_i \ldots \check{v}_j \cdots \wedge v_k$$

the general formula of 32.1 gives in this case that

Lemma. *All the maps between $\wedge^k V,\ \wedge^h V$ are combinations of $\psi^i c^j$.*

In order to understand the commutation relations between these two maps let us set $h := [c, \psi]$, go back to the spin formalism of 12.11 and recall the formulas of the action of the Clifford algebra on the exterior power

$$i(v)(u) := v \wedge u,\ j(\varphi)(v_1 \wedge v_2 \ldots \wedge v_k) := \sum_{t=1}^{k} (-1)^{t-1} < \varphi | v_t > v_1 \wedge v_2 \ldots \check{v}_t \ldots \wedge v_k$$

together with the identity

$$i(v)^2 = j(\varphi)^2 = 0,\ i(v)j(\varphi) + j(\varphi)i(v) =< \varphi | v >$$

Now clearly as operator we have $\psi := \sum_i i(e_i)i(f_i)$ while we claim that also, using the dual basis, $c = \sum_i j(f^i)j(e^i)$ in fact let us drop the symbols i, j compute directly in the Clifford algebra of $V \oplus V^*$ with standard hyperbolic form and compute

$$\sum_i f^i e^i v_1 \wedge v_2 \wedge \cdots \wedge v_k =$$

$$\sum_i \sum_{s<t} (-1)^{t+s} (< e^i | v_t >< f^i | v_s > - < f^i | v_t >< e^i | v_s > v_1 \wedge v_2 \wedge \check{v}_s \ldots \check{v}_t \cdots \wedge v_k$$

$$= \sum_{s<t} (-1)^{t+s-1} < v_s | v_t > v_1 \wedge v_2 \wedge \check{v}_s \ldots \check{v}_t \cdots \wedge v_k$$

Now we can use the commutation relations

$$e_i e^j + e^j e_i = 0,\ i \neq j,\ e_i f^j + f^j e_i = 0,\ f_i e^j + e^j f_i = 0,$$
$$f_i f^j + f^j f_i = 0,\ i \neq j,\ e_i e^i + e^i e_i = 1,\ f_i f^i + f^i f_i = 1$$

to deduce

$$h = [c, \psi] = \sum_{i,j} [f^j e^j, e_i f_i] = \sum_i [f^i e^i, e_i f_i] = \sum_i (f^i e^i e_i f_i - e_i f_i f^i e^i) =$$

$$\sum_i (-f^i e_i e^i f_i + f^i f_i - e_i e^i + e_i f^i f_i e^i) = \sum_i (-f^i e_i e^i f_i + f^i f_i + e^i e_i + f^i e_i e^i f_i) =$$

$$\sum_i (f^i f_i + e^i e_i)$$

Now we claim that on $\wedge^k V$ we have $\sum_i (f^i f_i + e^i e_i)$ acts as $2n - k$ in fact when we consider a vector $u := v_1 \wedge v_2 \ldots \wedge v_k$ with the v_i out of the symplectic basis, the operators $f^i f_i, e^i e_i$ annihilate u if e_i, f_i is one of the vectors $v_1, \ldots, v_k$ otherwise they map u into u itself.

Theorem. *The elements c, ψ, h satisfy the commutation relations of the standard generators e, f, h of $sl(2, \mathbb{C})$.*

Proof. We need only show that $[h, c] = 2c,\ [c, \psi] = -2\psi$ this follows imediately from the fact that c maps $\wedge^k V$ to $\wedge^{k-2} V$, ψ maps $\wedge^k V$ to $\wedge^{k+2} V$ while h has eigenvalue $2n - k$ on $\wedge^k V$. $\square$

We can apply now the representation theory of $sl(2, \mathbb{C})$ to deduce

Theorem.

(1) *We have a direct sum decomposition*

$$\wedge V = \oplus_{m \leq n} \wedge^{m,0} (V) \oplus \psi \wedge (\wedge V)$$

(2)

$$\wedge^k V = \oplus_{2i \geq k-n} \wedge^{k-2i,0} (V) \wedge \psi^i$$

Proof. For every representation M of $sl(2, \mathbb{C})$, we have, by the highest weight theory, a decomposition $M = M^e \oplus fM, \ M^e := \{m \in M | \, em = 0\}$.

Hence the first part follows from the fact that all the contractions reduce to c on skew symmetric tensors or $\wedge V_c = \oplus_{m \leq n} \wedge^{m,0} (V)$.

For any representation M, for every integer a denote by $M_a := \{m \in M | \, hm = am\}$ the space of weight a under h. Decompose M_e into weight spaces $M_{e,a} := M_e \cap M_a$.

One easily verifies that, for every weight space, we have

$$M_a = \oplus_i f^i M_{e,a+2i}$$

now clearly for $M = \wedge V$ we have $\wedge V_{c,k} = \wedge^{m,0}(V), \ \wedge V_{2n-k} = \wedge^k V$ hence the claim. □

We can interpret the first part of the previous theorem as saying that, the quotient algebra $\wedge V / \psi \wedge (\wedge V) = \oplus_{m \leq n} \wedge^{m,0} (V)$ is the direct sum of the fundamental representations of $Sp(V)$ in the same way as the usual exterior algebra is the direct sum of the fundamental representations of $SL(V)$.

35.7 The orthogonal case is different, in this case every contraction vanishes on the skew symmetric tensors which therefore are irreducible under the orthogonal group. In this case though we can still analyze the equivariant maps by invariants and easily see that the only invariant maps correspond to the only alternating invariants and are the $* : \wedge^k V \to \wedge^{m-k} V$, the exterior powers are irreducible and pairwise isomorphic under $SO(V)$. Now in general, since $SO(V)$ has index 2 in $O(V)$ we can apply the analysis of 23.1 and deduce that each irreducible representation M of $O(V)$ remain irreducible under $SO(V)$ or splits into 2 irreducibles according wether it is not or is isomorphic to $M \otimes \epsilon$ where here the sign representation is the one induced by the determinant.

§36 Invariants of matrices

Introduction In this section we will deduce the Invariant Theory of matrices from our previous work.

36.1

We are interested now in the following problem, describe the ring of invariants of the action of the general linear group $GL(n, \mathbb{C})$ acting by simultaneous conjugation on m copies of the space $M_n(\mathbb{C})$ of square matrices.

In intrinsic language, we have an $n-$dimensional vector space V and $GL(V)$ ancts on $End(V)^{\oplus m}$.

We will denote by $(X_1, X_2, \dots, X_m)$ an $m-$tuple of $n \times n$ matrices.

Before we formulate and prove the main theorem let us recall the results of 22.1.

The Theorem proved there can be reformulated as follows.

Suppose we are interested in the multilinear invariants of m matrices i.e. the invariant elements of the dual of $End(V)^{\otimes m}$.

First remark that, the dual of $End(V)^{\otimes m}$ can be identified, in a $GL(V)$ equivariant way to $End(V)^{\otimes m}$ by the pairing formula:

$$\langle A_1 \otimes A_2 \cdots \otimes A_m | B_1 \otimes B_2 \cdots \otimes B_m \rangle := tr(A_1 \otimes A_2 \cdots \otimes A_m \circ B_1 \otimes B_2 \cdots \otimes B_m) = \prod tr(A_i B_i).$$

Therefore under this isomorphism the multilineear invariants of matrices are identified with the $GL(V)$ invariants of $End(V)^{\otimes m}$ which in turn are spanned by the elements of the symmetric group, we deduce from Theorem 22.1

Lemma. *The multilinear invariants of m matrices are linearly spanned by the functions:*

$$\phi_\sigma(X_1, X_2, \dots, X_m) := tr(\sigma^{-1} \circ X_1 \otimes X_2 \otimes \cdots \otimes X_m).\ \sigma \in S_m$$

Recall that, from 22.1.3, we have that, if $\sigma = (i_1 i_2 \dots i_h)(j_1 j_2 \dots j_k) \dots (s_1 s_2 \dots s_m)$ is the cycle decomposition of σ then

$$\phi_\sigma(X_1, X_2, \dots, X_m) = tr(X_{i_1} X_{i_2} \dots X_{i_h}) tr(X_{j_1} X_{j_2} \dots X_{j_k}) \dots tr(X_{s_1} X_{s_2} \dots X_{s_m})$$

This explains our convention in defining ϕ_σ.

Theorem, FFT for matrices. *The ring of invariants of matrices under simultaneous conjugation is generated by the elements*

$$tr(X_{i_1} X_{i_2} \dots X_{i_{k-1}} X_{i_k})$$

Proof. The formula means that we take all possible non commutative monomials in the X_i and form their traces.

The proof of this Theorem is an immediate consequence of the previous Lemma and Arhonold method.

The proposed ring of invariants is in fact clearly closed under polarizations and coincides with the invariants, by the previous Lemma, for the multilinear elements, the claim follows.

There is a similar Theorem for the orthogonal and symplectic groups.

Assume that V is equipped with a non degenerate form $< u, v >$ (symmetric or skew symmetric), then we can identify $End(V) = V \otimes V^* = V \otimes V$ by

$$u \otimes v(w) :=< v, w > u$$

Let $\epsilon = \pm 1$ according to the symmetry, i.e. $< a, b >= \epsilon < b, a >$, we then get the formulas

$$(a \otimes b) \circ (c \otimes d) = a \otimes < b, c > d, \qquad tr(a \otimes b) =< b, a >$$
$$< (a \otimes b)c, d >=< b, c >< a, d >= \epsilon < c, (b \otimes a)d > .$$

In particular, using the notion of adjoint X^* of an operator, $< Xa, b >:=< a, X^* b >$ we see that $(a \otimes b)^* = \epsilon(b \otimes a)$.

We can now analyze first the multilinear invariants of m matrices under the group G (orthogonal or symplectic) fixing the given form, recalling the FFT of invariant Theory for such groups.

Compute such an invariant function on $End(V)^{\otimes m} = V^{\otimes 2m}$, write a decomposable element in this space as $X_1 \otimes X_2 \cdots \otimes X_m = u_1 \otimes v_1 \otimes u_2 \otimes v_2 \otimes \ldots u_m \otimes v_m, \quad X_i = u_i \otimes v_i$ then the invariants are spanned by products of m scalar products $< x_i, y_i >$ such that the $2m$ elements x_i, y_i exhaust the list of the u_i, v_j.

Of course in these scalar products a vector u can pe paired with another u or a v (homosexual or etherosexual pairings according to Weyl).

The previous formulas show that, up to a sign, such an invariant can be expressed in the form:

$$\phi_{\overline{\sigma}}(X_1, \ldots, X_m) := tr(Y_{i_1} Y_{i_2} \ldots Y_{i_h}) tr(Y_{j_1} Y_{j_2} \ldots Y_{j_k}) \ldots tr(Y_{s_1} Y_{s_2} \ldots Y_{s_m})$$

where, for each i the element Y_i is either X_i or X_i^*.

Combinatorially this can be pictured by a *marked permutation* $\overline{\sigma}$.

e.g.

$$\overline{\sigma} = \{3, 2, \overline{1}, \overline{5}, 4\}, \quad \phi_{\overline{\sigma}}(X_1, \ldots, X_5) := tr(X_3 X_1^* X_2) tr(X_4 X_5^*)$$

we deduce then the:

Theorem, FFT for matrices. *The ring of invariants of matrices under simultaneous conjugation by the group G (orthogonal or symplectic) is generated by the elements*

$$tr(Y_{i_1} Y_{i_2} \ldots Y_{i_{k-1}} Y_{i_k}), \; Y_{i_h} = X_{i_h} \text{ or } X_{i_h}^*.$$

Proof. Same as before.

We suggest the reader to formalize the previous analysis in the spirit of universal algebra as follows.

Definition. An algebra with trace is an associative algebra R equipped with a unary (linear) map $tr : R \to R$ satisfying the following axioms:

(1) $tr(a)b = btr(a), \; \forall a, b \in R.$
(2) $tr(tr(a)b) = tr(a)tr(b), \; \forall a, b \in R.$
(3) $tr(ab) = tr(ba), \; \forall a, b \in R.$

An algebra with involution is an associative algebra R equipped with a unary (linear) map $* : R \to R,\ x \to x^*$ satisfying the following axioms:

$$(x^8)^* = x,\ (xy)^* = y^* x^*,\ \forall x, y \in R$$

For an algebra with involution and a trace we shall assume the compatibility condition $tr(x^*) = tr(x), \forall x$.

As always happens in universal algebra for these structures one can construct free algebras on a set S (or on the vector space spanned by S).

Explicitely in the category of associative algebras the free algebra on S is obtained introducing variables $x_s,\ s \in S$ and considering the algebra having as basis all the *words* or monomials in the x_s.

For algebras with involution we have to add also the *adjoints* x_s^* an independent set of variables.

When we introduce a trace we have to add to these free algebras a set of commuting indeterminates $t(M)$ as M runs on the set of monomials.

Here in order to preserve the axioms we also need to impose the identifications $t(AB) = t(BA)$ (cyclic equivalence)
$t(A^*) = t(A)$ (adjoint symetry) for all monomials A, B.

In all cases the free algebra with trace is the tensor product of the free algebra (without trace) and the polynomial ring in the elements $t(A)$, this polynomial ring will be called the free trace ring (with or without involution).

The free algebras F_S by definition have the property that, given elements $f_s \in F_S,\ s \in S$ there is a unique homomorphism $F_S \to F_S$ (compatible with the structures (trace, involution) and mapping x_s to f_s for all s.

In particular we can rescale independently the x_s and thus speak about multihomogeneous, in particular multilinear elements in F_S.

Assume $S = \{1, 2, \ldots, m\}$ and let us describe the multilinear elements in the various cases.

1. For the free associative algebra in $m-$variables the multilinear elements (in all the variables) correspond to permutations in S_m as:

$$x_{i_1} x_{i_2} \ldots x_{i_m}$$

2. For the free algebra with involution the multilinear elements (in all the variables) correspond to marked permutations in S_m as:

$$y_{i_1} y_{i_2} \cdots y_{i_m}$$

where $y_i = x_i$ if i is unmarked, while $y_i = x_i^*$ if i is marked.

3. For the free associative algebra with trace in $m-$variables the multilinear elements (in all the variables) correspond to permutations in S_{m+1} according to the following rule.

Take such a permutation, write it into cycles isolating the cycle containing $m+1$ as follows:

$\sigma = (i_1 i_2 \dots i_h)(j_1 j_2 \dots j_k) \dots (r_1 \dots r_p)(s_1 s_2 \dots s_q m+1)$ set:

$$\psi_\sigma(x_1, x_2, \dots, x_m) = tr(x_{i_1} x_{i_2} \dots x_{i_h}) tr(x_{j_1} x_{j_2} \dots x_{j_k}) \dots tr(x_{r_1} x_{r_2} \dots x_{r_m}) x_{s_1} x_{s_2} \dots x_{s_m}.$$

4. For the free associative algebra with trace and involution in $m-$variables the multilinear elements (in all the variables) correspond to marked permutations in S_{m+1} but there are some equivalences due to the symmetry of trace under involution.

In fact it may be interesting to isolate, in the case of trace algebras, the part T_{m+1} of the multilinear elements in $m+1$ variables lying in the trace ring and compare it with the full set A_m of multilinear elements in m variables using the map $c_m : A_m \to T_{m+1},\ M \to t(M x_{m+1})$.

We leave it to the reader to verify that this map is a linear isomorphism.

Example of T_3, A_2 in the case of involutions:

$x_1 x_2, t(x_1 x_2 x_3);\ x_2 x_1, t(x_2 x_1 x_3);$

$$x_1 x_2, t(x_1 x_2 x_3);\ x_2 x_1, t(x_2 x_1 x_3);$$

$$x_1^* x_2, t(x_1^* x_2 x_3);\ x_2^* x_1, t(x_2^* x_1 x_3);$$

$$x_1 x_2^*, t(x_1 x_2^* x_3);\ x_2 x_1^*, t(x_2 x_1^* x_3);$$

$$t(x_1)x_2, t(x_1)t(x_2x_3);\ t(x_2)x_1, t(x_2)t(x_1x_3);\ t(x_1)x_2^*, t(x_1)t(x_2^*x_3);\ t(x_2)x_1^*, t(x_2)t(x_1^*x_3);$$

$$t(x_1)t(x_2), t(x_1)t(x_2)t(x_3);\ t(x_1x_2), t(x_1x_2)t(x_3);\ t(x_1x_2^*), t(x_1x_2^*)t(x_3);$$

Let us also denote for convenience by R the free algebra with trace in infinitely many variables and by Tr the polynomial ring of traces in R, the formal trace in R is denoted by $t : R \to Tr$.

We formalize these remarks as follows, we define maps

$$\Psi : \mathbb{C}[S_{m+1}] \to R, \Psi(\sigma) := \psi_\sigma),\ \Phi : \mathbb{C}[S_m] \to R, \Phi(\sigma) := \phi_\sigma)$$

We have a simple relation between these maps, if $\sigma \in S_{m+1}$ then:

$$\Phi(\sigma) = t(\Psi(\sigma) x_{m+1})$$

Remark finally that, for a permutation $\sigma \in S_m$ and an element $a \in \mathbb{C}[S_m]$ we have

$$\Phi(\sigma a \sigma^{-1}) = \Phi(a)(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(m)}).$$

We shall now take advantage of this formalism and study more general algebras, we want to study the set of $G-$equivariant maps

$$f : End(V)^{\oplus m} \to End(V)$$

here $m = \dim V$ and G is either $GL(V)$ or the orthogonal or symplectic group of a form.

We shall denote this space by $R_m(n)$ in the $GL(V)$ case and by $R_m^o(n)$ $R_m^s(n)$ respectively in the orthogonal and symplectic cases.

First remark that the scalar multiples of the identity $\mathbb{C}1_V \subset End(V)$ form the trivial representation and an equivariant map with values in $\mathbb{C}1_V$ can be canonically identified with the ring of invariants of matrices, we shall denote this ring by $T_m(n), T^o_m(n), T^s_m(n)$ in the 3 corresponding cases.

Next remark that $End(V)$ has an algebra structure and a trace both compatible with the group action, we deduce that under pointwise multiplication of the values the space $R_m(n)$ is a (possibly non commutative) algebra and moreover applying the trace function we deduce that:

$R_m(n)$ is an algebra with trace and the trace takes values in $T_m(n)$.

For $R^o_m(n)$ $R^s_m(n)$ we have same statement plus an involution.

Finally observe that the *coordinate maps* $(X_1, X_2, \ldots, X_m) \to X_i$ are clearly equivariant, we shall denote them (by the usual abuse of notations) by X_i we have:

Theorem. *In the case of $GL(V)$, $R_m(n)$ is generated as algebra over $T_m(n)$ by the variables X_i.*

$R^o_m(n)$ $R^s_m(n)$ are generated as algebra over $T^o_m(n), T^s_m(n)$ by the variables X_i, X_i^*.

Proof. Let us give the proof of the first statement, the others are similar.

Given an equivariant map $f(X_1, \ldots, X_m)$ in $m-$variables construct the invariant function of $m+1$ variables $g(X_1, \ldots, X_m, X_{m+1}) := tr(f(X_1, \ldots, X_m)X_{m+1}$.

By the structrure theorem of invariants g can be expressed as a linear combination of elements of the form $tr(M_1)tr(M_2)\ldots tr(M_k)$ where the M_i's are monomials in the variables $X_i,\ i = 1, \ldots, m+1$.

By construction g is linear in X_{m+1} thus we can assume that each term

$$tr(M_1)tr(M_2)\ldots tr(M_k)$$

is linear in X_{m+1} and in particular (using the cyclic equivalence of trace) we may assume that X_{m+1} appears only in M_k and $M_k = N_k X_{m+1}$ (N_k does not contain X_{m+1} and

$$tr(M_1)tr(M_2)\ldots tr(M_{k-1})tr(M_k) = tr((tr(M_1)tr(M_2)\ldots tr(M_{k-1})N_k)X_{m+1}).$$

It follows that we can construct a polynomial $h(X_1, \ldots, X_m)$ (non commutative) in the variables $X_i,\ i \leq m$ with coefficients invariants and such that:

$$tr(h(X_1, \ldots, X_m)X_{m+1}) = tr(f(X_1, \ldots, X_m)X_{m+1})$$

Now we can use the fact that the trace is a non degenerate bilinear form and that X_{m+1} is an independent variable to deduce that $h(X_1, \ldots, X_m) = f(X_1, \ldots, X_m)$ as desired.

Our next task is to use the second fundamental Theorem to understand the relations between the invariants that we have constructed, for this we will again use the language of universal algebra, we have to view the algebras constructed as quotients of the corresponding free algebras and we have to deduce some information on the kernel of this map.

Recall that in an algebra R with some extra operations an ideal I must be stable under these operations so that R/I can inherit the structure.

In an algebra with trace or involution thus I must be stable under the trace or the involution.

Let us call by F_m, F_m^i the free algebra with trace in $m-$variables and the free algebra with trace and involution in $m-$variables. We have the canonical maps (compatible with trace and when it applies with the involution)

$$\pi : F_m \to R_m(n),\ \pi^o : F_m^i \to R_m^o(n),\ \pi^s : F_m^i \to R_m^s(n)$$

We have already seen that in the free algebras we have the operation of substituting the variables x_s by any elements f_s, then one has the following:

Definition. An ideal of a free algebra, stable under all the substitutions of the variables, is called a T-ideal (or an ideal of polynomial identities).

The reason for this notation is the following, given an algebra R a morphism of the free algebra in R consists in evaluating the variables x_s in some elements r_s, the intersection of all the kernels of all possible morphisms are those expressions of th efree algebra which vanish identically when evaluated in R, it is clear that they form a T-ideal, the ideal of polynomial identities of R. Conversely if $I \subset F_S$ is a T-ideal it is easily seen that it is the ideal of polynomial identities of F_S/I.

Of course an intersection of T-ideals is again a T-ideal and thus we can speak of the T-ideal generated by a set of elements (identities).

Going back to the algebra $R_m(n)$ (or the other ones for G) we also see that we can compose any equivariant map $f(X_1, \ldots, X_m)$ with any m nmaps $g_i(X_1, \ldots, X_m)$ getting $f(g_1(X_1, \ldots, X_m), \ldots, g_m(X_1, \ldots, X_m)$ we thus see that also in $R_m(n)$ we have the morphisms given by substitutions of variables clearly substitution in the free algebra is compatible with substitution in the algebras $R_m(n), R_m^o(n).R_m^s(n)$ and thus we see that the kernels of the maps $\pi,\ \pi^o, \pi^s$ are all T-ideals.

They are called respectively:

The ideal of trace identities of matrices.

The ideal of trace identities of matrices with orthogonal involution.

The ideal of trace identities of matrices with symplectic involution.

We can apply the language of substitutions in the free algebra and thus define polarization and restitution operators and then we see immediately (working with infinitely many variables) that

Lemma. *two T-ideals which contain the same multilinear elements coincide.*

Thus we can deduce that the kernels of $\pi,\ \pi^o, \pi^s$ are generated as T-ideals by their multilinear elements.

We are going to prove in fact that they are generated as T-ideals by some special identities.

So let us first analyze the case of $R(n) := \cup_m R_m(n),\ T(n) := \cup_m T_m(n)$.

The multilinear elements of degree m (in the first variables) are contained in $R_m(n)$ and are of the form

$$\sum_{\sigma \in S_{m+1}} a_\sigma \psi_\sigma(X_1, X_2, \dots, X_m).$$

From the theory developed we know that the map

$$\Psi : \mathbb{C}[S_{m+1}] \to R_m(n), \quad \sum_{\sigma \in S_{m+1}} a_\sigma \sigma \to \sum_{\sigma \in S_{m+1}} a_\sigma \psi_\sigma(X_1, X_2, \dots, X_m)$$

has as kernel the ideal generated by the antisymmetrizer in $n+1$ elements.

Thus, first of all, the multilinear identities appera only for $m \geq n$ and for $m = n$ there is a unique identity (up to scalars).

Thus the first step consists in identifying the identity $A_n(x_1, \dots, x_n)$ corresponding to the antisymmetrizer

$$A_n(x_1, \dots, x_n) := \sum_{\sigma \in S_{n+1}} \epsilon_\sigma \psi_\sigma(x_1, x_2, \dots, x_n)$$

with ϵ_σ the sign of the permutation.

For this recall that there is a canonical identity, homogeneous of degree n in 1 variable, which is the Cayley-Hamilton identity, i.e. $\chi_X(X) = 0$ where $\chi_X(t) := \det(t - X) = t^n - tr(X)t^{n-1} + \dots$ is the characteristic polynomial.

Thus the formal identity is $P_n(x) := x^n - t(x)x^{n-1} + \dots$.

If we fully polarize this identity we get a multilinear trace identity $CH(x_1, \dots, x_n)$ for $n \times n$ matrices, whose terms not containg traces arise from the polarization of x^n and are thus of the form $\sum_{\tau \in S_n} x_{\tau(1)} x_{\tau(2)} \dots x_{\tau(n)}$.

By the uniqueness of the identities in degree n we must have that the polarized Cayley Hamilton is a multiple of the identity corresponding to the antisymmetrizer and to compute the scalar we may look in the two identitites at the terms not containg a trace.

Clearly $x_1 x_2 \dots x_n = \psi_{(12\dots nn+1)}$ and $\epsilon_{(12\dots nn+1)} = (-1)^n$ and thus we have finally:

Proposition. $A_n(x_1, \dots, x_n) = (-1)^n CH(x_1, \dots, x_n)$

Example $n = 2$ **(polarize** $P_2(x)$**)**

$$A_2 := x_1x_2 + x_2x_1 - t(x_1)x_2 - t(x_1)x_2 - t(x_1x_2) + t(x_1)t(x_2)$$

$$P_2(x) = x^2 - t(x)x + det(x) = x^2 - t(x)x + \frac{1}{2}(t(x)^2 - t(x^2))$$

Exercise Using this formal description (and restitution) write combinatorial expressions for the coefficients of the characteristic polynomial of X in terms of $tr(X^i)$ (i.e. expressions for elementary symmetric functions in terms of Newton power sums).

Let us look at some implications for relations among traces.

According to the general principle of correspondence between elements in R and T we deduce the trace relation

$$T_{n+1}(x_1, \ldots, x_n, x_{n+1}) := \sum_{\sigma \in S_{n+1}} \epsilon_\sigma \phi_\sigma(x_1, x_2, \ldots, x_n, x_{n+1}) = t(A_n(x_1, \ldots, x_n)x_{n+1})$$

Recall that $\phi_\sigma(x_1, x_2, \ldots, x_m) = tr(x_{i_1}x_{i_2}\ldots x_{i_h})tr(x_{j_1}x_{j_2}\ldots x_{j_k})\ldots tr(x_{s_1}x_{s_2}\ldots x_{s_m})$.

Example $n = 2$

$$T_3 := t(x_1x_2x_3) + t(x_2x_1x_3) - t(x_1)t(x_2x_3) - t(x_1)t(x_2x_3) - t(x_1x_2)t(x_3) + t(x_1)t(x_2)t(x_3).$$

Observe also that, when we restitute T_{n+1} we get $n!t(P_n(x)x) = T_{n+1}(x, x, \ldots, x)$, the vanishing of this expression for $n \times n$ matrices is precisely the identity which expresses the $n+1$ power sum of n indeterminates, as polynomial in the lower power sums, e.g $n = 2$:

$$t(x^3) = \frac{3}{2}t(x)t(x^2) - \frac{1}{2}t(x)^3.$$

Before we pass to the general case we should remark that any substitution map in R sends the trace ring Tr in itself, thus it also makes sense to speak of a T-ideal in Tr, in particular we have the T-ideal of trace relations, kernel of the evaluation map of Tr into $T(n)$, we whish to prove:

Theorem. *The T-ideal of trace relations, kernel of the evaluation map of Tr into $T(n)$ is generated (as T-ideal) by the trace relation T_{n+1}.*

The T-ideal of (trace) identities of $n \times n$ matrices is generated (as T-ideal) by the Cayley Hamilton identity.

Proof. From all the remarks made it is sufficient to prove that a multilinear trace relation (resp. identity) (in the first m variables) is in the T-ideal generated by T_{n+1} resp.A_n.

Let us first look at trace relations. By the description of trace identities it is enough to look at the relations of the type $\Phi(\tau(\sum_{\sigma \in S_{n+1}} \epsilon_\sigma \sigma)\gamma)$, $\tau, \gamma \in S_{m+1}$.

Write this as $\Phi(\gamma^{-1}(\gamma\tau(\sum_{\sigma \in S_{n+1}} \epsilon_\sigma \sigma))\gamma)$ and we have seen that conjugation corresponds to permutation of variables, an operation allowed in the T-ideal, thus we can assume that $\gamma = 1$.

We start with the remark $(m \geq n)$:

Remark: Splitting the cycles Every permutation τ in S_{m+1} can be written as a product $\tau = \alpha \circ \beta$ where $\beta \in S_{n+1}$ and, in each cycle of α there is at most 1 element in $1, 2, \ldots, n+1$.

This is an excercise on permutations it is based on the observation that

$$(xxxxx\ a\ yyyyy\ b\ zzzz\ c\ \ldots) = (xxxxx\ a)(yyyyy\ b)(zzzz\ c)(\ldots)(abc\ldots)$$

From this it follows that (up to a sign) we can assume that τ has the property that in each cycle of τ there is at most 1 element in $1, 2, \ldots, n+1$.

Having made this assumptionm we have to observe that, if $\sigma \in S_{n+1}$ the cycle decomposition of the permutation then the cycle decomposition of $\tau\sigma$ is obtained by formally substituting in the cycles of σ, to every element $a \in [1, \ldots, n+1]$ the cycle of τ containg a (written formally with a at the end) and then add all the cycles of τ not containing elements $\leq n+1$.

If we interpret this operation in terms of the corresponding trace element $\phi_\sigma(x_1, \ldots, x_{m+1})$ we see that the resulting element is the product of a trace element corresponding to the cycles of τ not containing elements $\leq n+1$ and an element obtained by subtituting in $\phi_\sigma(x_1, \ldots, x_{m+1})$ the variables x_i by monomials (which one reads off from the cycle decomposition of τ).

As a result we finally prove that a trace relation is in the T-ideal generated by T_{n+1}.

Let us pass now to trace identities.

First we remark, that by definition of ideal in a trace ring the relation T_{n+1} is a consequence of Cayley Hamilton.

A trace polynomial $f(x_1, x_2, \ldots, x_m)$ is a trace identity for matrices, if and only if $t(f(x_1, x_2, \ldots, x_m)x_{m+1})$ is a trace relation, thus from the previous proof we have that $t(f(x_1, x_2, \ldots, x_m)x_{m+1})$ is a linear combination of elements of type

$$AT_{n+1}(M_1, \ldots, M_{n+1}) = At(A_n(M_1, \ldots, M_n)M_{n+1}).$$

Now we have to consider two cases, the variable x_{m+1} appears in A or in one of the M_i. In the first case we have $AT_{n+1}(M_1, \ldots, M_{n+1}) = t(A_n(M_1, \ldots, M_n)M_{n+1})t(Bx_{m+1})$, in the second, due to the antisymmetry of AT we can assume that x_{m+1} appears in $M_{n+1} = Bx_{m+1}C$ hence

$$\begin{aligned} At(A_n(M_1, \ldots, M_n)M_{n+1}) &= At(A_n(M_1, \ldots, M_n)Bx_{m+1}C) = \\ &= t(CAA_n(M_1, \ldots, M_n)Bx_{m+1}) \end{aligned}$$

$CAA_n(M_1, \ldots, M_n)B$ is also clearly a consequernce of Cayley Hamilton.