3

Theory of Algebraic Forms

Summary. This is part of classical invariant theory, which is not a very appropriate name, referring to the approach to invariant theory presented in Hermann Weyl's *The Classical Groups* in which the connection between invariant theory, tensor algebra, and representation theory is stressed. One of the main motivations of Weyl was the role played by symmetry in the developments of quantum mechanics and relativity, which had just taken shape in the 30–40 years previous to the appearance of his book.

Invariant theory was born and flourished in the second half of the 19^{th} century due to the work of Clebsch and Gordan in Germany, Cayley and Sylvester in England, and Capelli in Italy, to mention some of the best known names. It was developed at the same time as other disciplines which have a strong connection with it: projective geometry, differential geometry and tensor calculus, Lie theory, and the theory of associative algebras. In particular, the very formalism of matrices, determinants, etc., is connected to its birth.

After the prototype theorem of I.T., the theory of symmetric functions, 19^{th}-century I.T. dealt mostly with binary forms (except for Capelli's work attempting to lay the foundations of what he called *Teoria delle forme algebriche*). One of the main achievements of that period is Gordan's proof of the finiteness of the ring of invariants. The turning point of these developments has been Hilbert's theory, with which we enter into the methods that have led to the development of commutative algebra. We shall try to give an idea of these developments and how they are viewed today.

In this chapter we will see a few of the classical ideas which will be expanded on in the next chapters. In particular, polarization operators have a full explanation in the interpretation of the Cauchy formulas by representation theory; cf. Chapter 9.

1 Differential Operators

1.1 Weyl Algebra

On the polynomial ring $\mathbb{C}[x_1, \dots, x_n]$ acts the algebra of *polynomial differential operators, $W(n) := \mathbb{C}\big[x_1, \dots, x_n, \frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\big]$,*[11] which is a noncommutative algebra, the multiplication being the composition of operators.

[11] This is sometimes known as the Weyl algebra, due to the work of Weyl on commutation relations in quantum mechanics.

In noncommutative algebra, given any two elements a, b, one defines their commutator $[a, b] = ab - ba$, and says that $ab = ba + [a, b]$ is a *commutation relation.*

The basic commutation relations for $W(n)$ are

$$(1.1.1) \qquad [x_i, x_j] = \left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right] = 0, \quad \left[\frac{\partial}{\partial x_i}, x_j\right] = \delta_i^j.$$

In order to study $W(n)$ it is useful to introduce a notion called the *symbol.* One starts by writing an operator as a linear combination of terms $x_1^{h_1} \cdots x_n^{h_n} \frac{\partial^{k_1}}{\partial x_1} \cdots \frac{\partial^{k_n}}{\partial x_n}$. The symbol $\sigma(P)$ of a polynomial differential operator P is obtained in an elementary way by taking the terms of higher degree in the derivatives and substituting commutative variables ξ_j for the operators $\frac{\partial}{\partial x_j}$;

$$\sigma\left(x_1^{h_1} \cdots x_n^{h_n} \frac{\partial}{\partial x_1}^{k_1} \cdots \frac{\partial}{\partial x_n}^{k_n}\right) = x_1^{h_1} \cdots x_n^{h_n} \xi_1^{k_1} \cdots \xi_n^{k_n}.$$

In a more formal way, one can filter the algebra of operators by the degree in the derivatives and take the associated graded algebra. Let us recall the method.

Definition 1. A filtration of an algebra R consists of an increasing sequence of subspaces $0 = R_0 \subset R_1 \subset R_2 \subset \cdots \subset R_i \subset \cdots \subset$ such that

$$\bigcup_{i=0}^{\infty} R_i = R, \quad R_i R_j \subset R_{i+j}.$$

Given a filtered algebra R, one can construct the *graded algebra* $\mathrm{Gr}(R) := \oplus_{i=0}^{\infty} R_{i+1}/R_i$. The multiplication in $\mathrm{Gr}(R)$ is induced by the product $R_i R_j \subset R_{i+j}$, so that if $a \in R_i,\ b \in R_j$, the class of the product ab in R_{i+j}/R_{i+j-1} depends only on the classes of a modulo R_{i-1} and of b modulo R_{j-1}. In this way $\mathrm{Gr}(R)$ becomes a *graded algebra* with R_i/R_{i-1} being the elements of degree i. For a filtered algebra R and an element $a \in R$, one can define the symbol of a as an element $\sigma(a) \in \mathrm{Gr}(R)$ as follows. We take the minimum i such that $a \in R_i$ and set $\sigma(a)$ to be the class of a in R_i/R_{i-1}.

In $W(n)$ one may consider the filtration given by the degree in the derivatives for which

$$(1.1.2) \qquad W(n)_i := \left\langle x_1^{h_1} \cdots x_n^{h_n} \frac{\partial}{\partial x_1}^{k_1} \cdots \frac{\partial}{\partial x_n}^{k_n} \mid \sum k_j \leq i \right\rangle.$$

Let us denote by $S(n)$ the resulting graded algebra and by $\xi_i := \sigma\left(\frac{\partial}{\partial x_i}\right) \in S(n)_1$ the symbol. We keep for the variables x_i the same notation for their symbols since they are in $R_0 = S_0$. From the commutation relations 1.1.1, it follows that the classes ξ_i of $\frac{\partial}{\partial x_i}$ and of x_i in $S(n)$ commute and we have:

Proposition 1. *$S(n)$ is the commutative algebra of polynomials:*

$$S(n) := \mathbb{C}[x_1, \ldots, x_n, \xi_1, \ldots, \xi_n]$$

Proof. The easy proof is left to the reader. One uses the fact that the monomials in the $\frac{\partial}{\partial x_i}$ are a basis of $W(n)$ as a module over the polynomial ring $\mathbb{C}[x_1, \dots, x_n]$. $\square$

There is a connection between differential operators and group actions. Consider in fact the linear group $GL(n, \mathbb{C})$ which acts on the space $\mathbb{C}^n$ and on the algebra of polynomials by the formula $f^g(x) = f(g^{-1}x)$. If $g \in GL(n, \mathbb{C})$ and $P \in W(n)$, consider the operator

$$P^g(f(x)) := {}^g\big(P^{g^{-1}}(f(x))\big), \ \forall f \in \mathbb{C}[x_1, \dots, x_n].$$

In other words, if we think of $g : \mathbb{C}[x_1, \dots, x_n] \to \mathbb{C}[x_1, \dots, x_n]$ as a linear map, we have $P^g = g \circ P \circ g^{-1}$.

It is clear that the action on linear operators is an automorphism of the algebra structure, and we claim that it transforms $W(n)$ into itself. Since $W(n)$ is generated by the elements x_i, $\frac{\partial}{\partial x_i}$, it is enough to understand how $GL(n, \mathbb{C})$ acts on these elements. On an element x_i, it will give a linear combination $g(x_i) = \sum_j a_{ji} x_j$. As for $g\big(\frac{\partial}{\partial x_i}\big)$, let us first remark that this operator acts as a *derivation*.

We will investigate derivations in the following chapters but now give the definition:

Definition 2. A derivation of an algebra R is a linear operator $D : R \to R$ such that

$$D(ab) = D(a)b + aD(b). \tag{1.1.3}$$

A direct computation shows that if $\phi : R \to R$ is an algebra automorphism and $D : R \to R$ a derivation, then $\phi \circ D \circ \phi^{-1}$ is also a derivation.

In particular, we may apply these remarks to $g \circ \frac{\partial}{\partial x_i} \circ g^{-1}$. Let $g^{-1}(x_i) = \sum_j b_{ji} x_j$. We have that

$$g \circ \frac{\partial}{\partial x_i} \circ g^{-1}(x_j) = g \circ \frac{\partial}{\partial x_i}\Big(\sum_h b_{hj} x_h\Big) = b_{ij}.$$

Proposition 2. *We have* $g \circ \frac{\partial}{\partial x_i} \circ g^{-1} = \sum_j b_{ij} \frac{\partial}{\partial x_j}$.

Proof. A derivation of $\mathbb{C}[x_1, \dots, x_n]$ is completely determined by its values on the variables x_i. By the above formulas, $g \circ \frac{\partial}{\partial x_i} \circ g^{-1}(x_j) = b_{ij} = \sum_k b_{ik} \frac{\partial}{\partial x_k}(x_j)$; hence the statement. $\square$

The above formulas prove that the space of derivatives behaves as the dual of the space of the variables and that the action of the group is by inverse transpose. This of course has an intrinsic meaning: if V is a vector space and $\mathcal{P}(V)$ the ring of polynomials on V, we have that $V^* \subset \mathcal{P}(V)$ are the linear polynomials. The space V can be identified intrinsically with the space spanned by the derivatives. If $v \in V$, we can define the *derivative* D_v *in the direction of* v in the usual way:

$$D_v f(x) := \frac{d}{dt} f(x + tv)_{t=0}. \tag{1.1.4}$$

Intrinsically the algebra of differential operators is generated by V, V^*, with the commutation relations:

$$(1.1.5) \qquad \phi \in V^*,\ v \in V, \quad [\phi, v] = -\langle \phi \mid v \rangle.$$

The action of $GL(V)$ preserves the filtration, and thus the group acts on the graded algebra $S(n)$. We may identify this action as that on the polynomials on $V \oplus V^*$, (cf. Chapter 9).

Remark. The algebra of differential operators is a graded algebra under the following notion of degree. We say that an operator has degree k if it maps the space of homogeneous polynomials of degree h to the polynomials of degree $h + k$ for every h. For instance, the variables x_i have degree 1 while the derivatives $\frac{\partial}{\partial x_j}$ have degree -1. In particular we may consider the differential operators of degree 0.

For all i, let $\mathbb{C}[x_1, \ldots, x_n]_i$ be the space of homogeneous polynomials of degree i.

Exercise. Prove that over a field F of characteristic 0, any linear map from the space $F[x_1, \ldots, x_n]_i$ to the space $F[x_1, \ldots, x_n]_j$ can be expressed via a differential operator of degree $j - i$.

2 The Aronhold Method, Polarization

2.1 Polarizations

Before proceeding, let us recall, in a language suitable for our purposes, the usual Taylor–Maclaurin expansion.

Consider a function $F(x)$ of a vector variable $x \in V$. Under various types of assumptions we have a development for the function $F(x+y)$ of two vector variables.

For our purposes, we may restrict our considerations to polynomials and develop $F(x + y) := \sum_{i=0}^{\infty} F_i(x, y)$, where by definition $F_i(x, y)$ is homogeneous of degree i in y (of course for polynomials the sum is really finite). Therefore, for any value of a parameter λ, we have $F(x + \lambda y) := \sum_{i=0}^{\infty} \lambda^i F_i(x, y)$.

If F is also homogeneous of degree k we have

$$\sum_{i=0}^{\infty} \lambda^k F_i(x, y) = \lambda^k F(x + y) = F(\lambda(x + y)) = F(\lambda x + \lambda y) = \sum_{i=0}^{\infty} \lambda^i F_i(\lambda x, y),$$

and we deduce that $F_i(x, y)$ is also homogeneous of degree $k - i$ in x.

Given two functions F, G, we clearly have that

$$F(x+y)G(x+y) = \sum_{i=0}^{\infty} \sum_{a+b=i} F_a(x,y)G_b(x,y)$$

is the decomposition into homogeneous components relative to y.

The operator $D = D_{y,x}$ defined by the formula $D_{y,x}F(x) := F_1(x,y)$ is clearly linear, and also by the previous formula we have $D(FG) = D(F)G + F\,D(G)$. These are the defining conditions of a *derivation*.

If we use coordinates $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$, we have that $D_{y,x} = \sum_{i=1}^{n} y_i \frac{\partial}{\partial x_i}$.

Definition. The operator $D_{y,x} = \sum_{i=1}^{n} y_i \frac{\partial}{\partial x_i}$ is called a *polarization* operator.

The effect of applying $D_{y,x}$ to a bihomogeneous function of two variables x, y is to decrease by one the degree of the function in x and raise by one the degree in y.

Assume we are now in characteristic 0, we have then the standard theorem of calculus:

Theorem. $F(x+y) = \sum_{i=0}^{\infty} \frac{1}{i!} D_{y,x}^i F(x) \quad \left(= \sum_{i=0}^{\infty} F_i(x,y)\right)$.

Proof. We reduce to the one variable theorem and deduce that

$$F(x+\lambda y) := \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \frac{d^i}{d\lambda^i} F(x+\lambda y)_{\lambda=0}.$$

Then

$$F_i(x,y) = \frac{1}{i!} \frac{d^i}{d\lambda^i} F(x+\lambda y)_{\lambda=0},$$

and this is, by the chain rule, equal to $\frac{1}{i!} D_{y,x}^i F(x)$. □

2.2 Restitution

Suppose now that we consider the action of an invertible linear transformation on functions. We have $(gF)(x+y) = F(g^{-1}x + g^{-1}y)$. Hence we deduce that the polarization operator commutes with the action of the linear group. The main consequence is:

Proposition. *If $F(x)$ is an invariant of a group G, so are the polarized forms $F_i(x,y)$.*

Of course we are implicitly using the (direct sum) linear action of G on pairs of variables.

Let us further develop this idea. Consider now any number of vector variables and, for a polynomial function F, homogeneous of degree m, the expansion

$$F(x_1 + x_2 + \cdots + x_m) = \sum_{h_1, h_2, \ldots, h_m} F_{h_1, h_2, \ldots, h_m}(x_1, x_2, \ldots, x_m),$$

where $\sum h_i = m$, and the indices h_i represent the degrees of homogeneity in the variables x_i. A repeated application of the Taylor–Maclaurin expansion gives

$$(2.2.1) \qquad F_{h_1,h_2,\dots,h_m}(x_1, x_2, \dots, x_m) = \frac{1}{h_1!h_2!\cdots h_n!} D^{h_1}_{x_1x} D^{h_2}_{x_2x} \cdots D^{h_m}_{x_mx} F(x).$$

In particular, in the expansion of $F(x_1 + x_2 \cdots + x_m)$ there will be a term which is linear in all the variables x_i.

Definition. The term $F_{1,1,\dots,1}(x_1, \dots, x_m)$ multilinear in all the variables x_i is called the *full polarization of the form* F.

Let us write $PF := F_{1,1,\dots,1}(x_1, \dots, x_m)$ to stress the fact that this is a linear operator. It is clear that if $\sigma \in S_m$ is a permutation, then

$$F(x_1 + x_2 \cdots + x_n) = F(x_{\sigma 1} + x_{\sigma 2} \cdots + x_{\sigma n}).$$

Hence we deduce that the polarized form satisfies the *symmetry* property:

$$PF(x_1, \dots, x_m) = PF(x_{\sigma 1}, \dots, x_{\sigma m}).$$

We have thus found that:

Lemma. *The full polarization is a linear map from the space of homogeneous forms of degree m to the space of symmetric multilinear functions in m (vector) variables.*

Now let us substitute for each variable x_i the variable $\lambda_i x$ (the λ_i's being distinct numbers). We obtain:

$$\begin{aligned}(\lambda_1 + \lambda_2 \cdots + \lambda_m)^m F(x) &= F((\lambda_1 + \lambda_2 \cdots + \lambda_m)x) \\ &= F(\lambda_1 x + \lambda_2 x \cdots + \lambda_m x) \\ &= \sum_{h_1,h_2,\dots,h_m} F_{h_1,h_2,\dots,h_m}(\lambda_1 x, \lambda_2 x, \dots, \lambda_m x) \\ &= \sum_{h_1,h_2,\dots,h_m} \lambda_1^{h_1} \lambda_2^{h2} \dots \lambda_m^{h_m} F_{h_1,h_2,\dots,h_m}(x, x, \dots, x).\end{aligned}$$

Comparing the coefficients of the same monomials on both sides, we get

$$\binom{m}{h_1 h_2 \cdots h_m} F(x) = F_{h_1,h_2,\dots,h_m}(x, x, \dots, x).$$

In particular,

$$m!F(x) = PF(x, x, \dots, x).$$

Since we are working in characteristic zero we can also rewrite this identity as

$$F(x) = \frac{1}{m!} PF(x, x, \dots, x).$$

For a polynomial $G(x_1, x_2, \dots, x_m)$ of m vector variables, the linear operator

$$R : G(x_1, x_2, \dots, x_m) \mapsto \frac{1}{m!} G(x, x, \dots, x)$$

is called the *restitution* in the classical literature. We have:

Theorem. *The maps P, R are inverse isomorphisms, equivariant for the group of all linear transformations, between the space of homogeneous forms of degree m and the space of symmetric multilinear functions in m variables.*

Proof. We have already proved that $RPF = F$. Now let $G(x_1, x_2, \dots, x_m)$ be a symmetric multilinear function. In order to compute PRG we must determine the multilinear part of $\frac{1}{m!} G\left(\sum x_i, \sum x_i, \dots, \sum x_i\right)$.

By the multilinearity of G we have that

$$G\left(\sum x_i, \sum x_i, \dots, \sum x_i\right) = \sum G(x_{i_1}, x_{i_2}, \dots, x_{i_m}),$$

where the right sum is over all possible sequences of indices $i_1 i_2 \cdots i_m$ taken out of the numbers $1, \dots, m$. But the multilinear part is exactly the sum over all of the sequences without repetitions, i.e., the permutations. Thus

$$PRG = \frac{1}{m!} \sum_{\sigma \in S_m} G(x_{\sigma 1}, x_{\sigma 2}, \dots, x_{\sigma m}).$$

Since G is symmetric, this sum is in fact G. □

Remark. We will mainly use the previous theorem to reduce the computation of invariants to the multilinear ones. At this point it is not yet clear why this should be simpler, but in fact we will see that in several interesting cases this turns out to be true, and we will be able to compute all of the invariants by this method. This sequence of ideas is sometimes referred to as *Aronhold's method.*

2.3 Multilinear Functions

In order to formalize the previous method, consider an infinite sequence of n-dimensional vector variables $x_1, x_2, \dots, x_k, \dots$. We usually have two conventions: each x_i is either a column vector $x_{1i}, x_{2i}, \dots, x_{ni}$ or a row vector $x_{i1}, x_{i2}, \dots, x_{in}$. In other words we consider the x_{ij} as the coordinates of the space of $n \times \infty$ or of $\infty \times n$ matrices or of the space of sequences of column (resp. row) vectors.

Let $A := \mathbb{C}[x_{ij}]$ be the polynomial ring in the variables x_{ij}. For the elements of A we have the notion of being homogeneous with respect to one of the vector variables x_i, and A is in this way a *multigraded ring.*[12]

We denote by $A_{h_1, h_2, \dots, h_i, \dots}$ the multihomogeneous part relative to the degrees $h_1, h_2, \dots, h_i, \dots,$ in the vector variables $x_1, x_2, \dots, x_i, \dots$.

[12] Notice that we are not considering the separate degrees in all of the variables x_{ij}.

We have of course the notions of multihomogeneous subspace or subalgebra. For each pair i, j of indices, we consider the corresponding polarization operator

$$(2.3.1) \qquad D_{ij} = \sum_{h=1}^{n} x_{hi} \frac{\partial}{\partial x_{hj}}, \quad \text{or} \quad D'_{ij} = \sum_{h=1}^{n} x_{ih} \frac{\partial}{\partial x_{jh}}, \text{ in row notation.}$$

We view these operators as acting as derivations on the ring $A = \mathbb{C}[x_{ij}]$.

Given a function F homogeneous in the vector variables $x_1, x_2, \dots, x_m$ of degrees $h_1, h_2, \dots, h_m$, we can perform the process of polarization on each of the variables x_i as follows: Choose from the infinite list of vector variables m disjoint sets X_i of variables, each with h_i elements. Then fully polarize the variable x_i with respect to the chosen set X_i.

The result is multilinear and symmetric in each of the sets X_i. The function F is recovered from the polarized form by a sequence of restitutions.

We should remark that a restitution is a particular form of polarization since, if a function F is linear in the variable x_i, the effect of the operator D_{ji} on F is that of replacing in F the variable x_j with x_i.

Definition. A subspace V of the ring A is said to be *stable under polarization* if it is stable under all polarization operators.

Remark. Given a polynomial F, F is homogeneous of degree m with respect to the vector variable x_i if and only if $D_{ii}F = mF$.

From this remark one can easily prove the following:

Lemma. *A subspace V of A is stable under the polarizations D_{ii} if and only if it is multihomogeneous.*

2.4 Aronhold Method

In this section we will use the term *multilinear function* in the following sense:

Definition. We say that a polynomial $F \in A$ is multilinear if it is homogeneous of degree 0 or 1 in each of the variables x_i.

In particular we can list the indices of the variables $i_1, \dots, i_k$ in which F is linear (the variables which appear in the polynomial) and say that F is multilinear in the x_{i_j}.

Given a subspace V of A we will denote by V_m the set of multilinear elements of V.

Theorem. *Given two subspaces V, W of A stable under polarization and such that $V_m \subset W_m$, we have $V \subset W$.*

Proof. Since V is multihomogeneous it is enough to prove that given a multihomogeneous function F in V, we have $F \in W$. We know that F can be obtained by restitution from its fully polarized form $F = RPF$. The hypotheses imply that $PF \in V$ and hence $PF \in W$. Since the restitution is a composition of polarization operators and W is assumed to be stable under polarization, we deduce that $F \in W$. □

Corollary. *If two subspaces V, W of A are stable under polarization and $V_m = W_m$, then $V = W$.*

We shall often use this corollary to compute invariants. The strategy is as follows. We want to compute the space W of invariants in A under some group G of linear transformations in n-dimensional space. We produce a list of invariants (which are more or less obvious) forming a subspace V closed under polarization. We hope to have found all invariants and try to prove $V = W$. If we can do it for the multilinear invariants we are done.

3 The Clebsch–Gordan Formula

3.1 Some Basic Identities

We start with two sets of *binary variables* (as the old algebraists would say):

$$x := (x_1, x_2); \quad y := (y_1, y_2).$$

In modern language these are linear coordinate functions on the direct sum of two copies of a standard two-dimensional vector space. We form the following matrices of functions and differential operators:

$$\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial y_1} \\ \frac{\partial}{\partial x_2} & \frac{\partial}{\partial y_2} \end{pmatrix} = \begin{pmatrix} \Delta_{xx} & \Delta_{xy} \\ \Delta_{yx} & \Delta_{yy} \end{pmatrix}.$$

We define

$$(x, y) := \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = x_1 y_2 - y_1 x_2$$

$$\Omega := \det \begin{pmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial y_1} \\ \frac{\partial}{\partial x_2} & \frac{\partial}{\partial y_2} \end{pmatrix} = \frac{\partial}{\partial x_1}\frac{\partial}{\partial y_2} - \frac{\partial}{\partial x_2}\frac{\partial}{\partial y_1}.$$

Since we are in a noncommutative setting the determinant of the product of two matrices need not be the product of the determinants. Moreover, it is even necessary to specify what one means by the determinant of a matrix with entries in a noncommutative algebra.

A direct computation takes care of the noncommutative nature of differential polynomials and yields

$$(3.1.1) \qquad (x, y)\Omega = \det \begin{pmatrix} \Delta_{xx} + 1 & \Delta_{xy} \\ \Delta_{yx} & \Delta_{yy} \end{pmatrix} = (\Delta_{xx} + 1)\Delta_{yy} - \Delta_{yx}\Delta_{xy}.$$

This is a basic identity whose extension to many variables is the *Capelli identity* on which one can see the basis for the representation theory of the linear group.

Now let $f(x,y)$ be a polynomial in x_1, x_2, y_1, y_2, homogeneous of degree m in x and n in y. Notice that

$$\Delta_{xx} f(x,y) = mf(x,y), \quad \Delta_{yy} f(x,y) = nf(x,y), \tag{3.1.2}$$

$$\begin{aligned} &\Delta_{yx} f(x,y) \text{ has bidegree } (m-1, n+1), \\ &\Delta_{xy} f(x,y) \text{ has bidegree } (m+1, n-1). \end{aligned} \tag{3.1.3}$$

Observe that $(\Delta_{xx}+1)\Delta_{yy} f(x,y) = (m+1)nf(x,y)$; thus the identity

$$(x,y)\Omega f(x,y) = (m+1)nf(x,y) - \Delta_{yx}\Delta_{xy} f(x,y). \tag{3.1.4}$$

The identity 3.1.4 is the beginning of the Clebsch–Gordan expansion.

We have some commutation rules:

$$\begin{aligned} &[\Omega, \Delta_{yx}] = [\Omega, \Delta_{xy}] = 0, \ [\Omega, \Delta_{xx}] = \Omega = [\Omega, \Delta_{yy}], \\ &[(x,y), \Delta_{yx}] = [(x,y), \Delta_{xy}] = 0, \\ &[(x,y), \Delta_{xx}] = -(x,y) = [(x,y), \Delta_{yy}]. \end{aligned} \tag{3.1.5}$$

Finally

$$\Omega(x,y) - (x,y)\Omega = 2 + \Delta_{xx} + \Delta_{yy}. \tag{3.1.6}$$

In summary: In the algebra generated by the elements

$$\Delta_{xx}, \Delta_{xy}, \Delta_{yx}, \Delta_{yy},$$

the elements $\Delta_{xx} + \Delta_{yy}$ and the elements $\Omega^k (x,y)^k$ are in the center.

The Clebsch–Gordan expansion applies to a function $f(x,y)$ of bidegree (m,n), say $n \leq m$:

Theorem (Clebsch–Gordan expansion).

$$f(x,y) = \sum_{h=0}^{n} c_{h,m,n} (x,y)^h \Delta_{yx}^{n-h} \Delta_{xy}^{n-h} \Omega^h f(x,y) \tag{3.1.7}$$

with the coefficients $c_{h,m,n}$ to be determined by induction.

Proof. The proof is a simple induction on n. Since $\Omega f(x,y)$ has bidegree $(m-1, n-1)$ and $\Delta_{xy} f(x,y)$ has bidegree $(m+1, n-1)$ we have by induction on n

$$(x,y)\Omega f(x,y) = (x,y) \sum_{h=0}^{n-1} c_{h,m-1,n-1} (x,y)^h \Delta_{yx}^{n-1-h} \Delta_{xy}^{n-1-h} \Omega^h \Omega f(x,y)$$

$$\Delta_{xy} f(x,y) = \sum_{h=0}^{n-1} c_{h,m+1,n-1} (x,y)^h \Delta_{yx}^{n-1-h} \Delta_{xy}^{n-1-h} \Omega^h \Delta_{xy} f(x,y).$$

Now using 3.1.4 and 3.1.5 we obtain

$$\sum_{h=0}^{n-1} c_{h,m-1,n-1}(x,y)^{h+1}\Delta_{yx}^{n-h-1}\Delta_{xy}^{n-h-1}\Omega^{h+1}f(x,y)$$

$$= (m+1)nf(x,y) - \sum_{h=0}^{n-1} c_{h,m+1,n-1}(x,y)^h\Delta_{yx}^{n-h}\Delta_{xy}^{n-h}\Omega^h f(x,y),$$

so

$$\begin{aligned}(m+1)nf(x,y) &= \sum_{h=1}^{n} c_{h-1,m-1,n-1}(x,y)^h\Delta_{yx}^{n-h}\Delta_{xy}^{n-h}\Omega^h f(x,y) \\ &\quad + \sum_{h=0}^{n-1} c_{h,m+1,n-1}(x,y)^h\Delta_{yx}^{n-h}\Delta_{xy}^{n-h}\Omega^h f(x,y) \\ &= (m+1)nf(x,y).\end{aligned}$$

Thus, $c_{h,m,n} = \frac{1}{(m+1)n}[c_{h-1,m-1,n-1} + c_{h,m+1,n-1}]$. □

We should make some comments on the meaning of this expansion, which is at the basis of many developments. As we shall see, the decomposition given by the Clebsch–Gordan formula is unique, and it can be interpreted in the language of representation theory.

In Chapter 10 we shall prove that the spaces P_m of binary forms of degree m constitute the full list of the irreducible representations of $Sl(2)$. In this language the space of forms of bidegree (m, n) can be thought of as the tensor product of the two spaces P_m, P_n of homogeneous forms of degree m, n, respectively.

The operators Δ_{xy}, Δ_{yx}, and Ω, (x, y) are all $Sl(2)$-equivariant.

Since $f(x, y)$ has bidegree (m, n) the form $\Delta_{xy}^{n-h}\Omega^h f(x, y)$ is of bidegree $(m + n - 2h, 0)$, i.e., it is an element of P_{m+n-2h}.

We thus have the $Sl(2)$-equivariant operator

$$\Delta_{xy}^{n-h}\Omega^h : P_m \otimes P_n \to P_{m+n-2h}.$$

Similarly, we have the $Sl(2)$-equivariant operator

$$(x,y)^h\Delta_{yx}^{n-h} : P_{m+n-2h} \to P_m \otimes P_n.$$

Therefore the Clebsch–Gordan formula can be understood as the expansion of an element f in $P_m \otimes P_n$ in terms of its components in the irreducible representations P_{m+n-2h}, $0 \le h \le \min(m, n)$ of this tensor product

$$P_m \otimes P_n = \bigoplus_{h=0}^{\min(m,n)} P_{m+n-2h}.$$

At this point the discussion can branch. Either one develops the basic algebraic identities for dimension greater than two (Capelli's theory), or one can enter into the developments of the theory of binary forms, i.e., the application of what has been done in the invariant theory of the group $Sl(2)$. In the next section we discuss Capelli's theory, leaving the theory of binary forms for Chapter 15.

4 The Capelli Identity

4.1 Capelli Identity

Capelli developed the algebraic formula generalizing the Clebsch–Gordan expansion, although, as we shall see, some of the formulas become less explicit.

In order to keep the notations straight we consider in general m, n-ary variables:

$$\underline{x}_i := x_{i1}\, x_{i2}\, \cdots\, x_{in};\ i = 1, \ldots, m,$$

which we can view as rows of an $m \times n$ matrix of variables:

$$X := \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ x_{i1} & x_{i2} & \ldots & x_{in} \\ \ldots & \ldots & \ldots & \ldots \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{pmatrix}. \tag{4.1.1}$$

In the same way we form an $n \times m$ matrix of derivative operators:

$$Y := \begin{pmatrix} \frac{\partial}{\partial x_{11}} & \frac{\partial}{\partial x_{21}} & \ldots & \frac{\partial}{\partial x_{m1}} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial}{\partial x_{1i}} & \frac{\partial}{\partial x_{2i}} & \ldots & \frac{\partial}{\partial x_{mi}} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial}{\partial x_{1n}} & \frac{\partial}{\partial x_{2n}} & \ldots & \frac{\partial}{\partial x_{mn}} \end{pmatrix}. \tag{4.1.2}$$

The product XY is a matrix of differential operators. In the i, j position it has the polarization operator $\Delta_{i,j} := \sum_{h=1}^{n} x_{ih} \frac{\partial}{\partial x_{jh}}$.

Now we want to repeat the discussion on the computation of the determinant of the matrix $XY = (\Delta_{i,j})$. For an $m \times m$ matrix $\mathcal{U}$ in noncommutative variables $u_{i,j}$, the determinant will be computed multiplying from left to right as

$$\det(\mathcal{U}) := \sum_{\sigma \in \mathcal{S}_n} \epsilon_\sigma u_{1,\sigma(1)} u_{2,\sigma(2)} \cdots u_{m,\sigma(m)}. \tag{4.1.3}$$

In our computations we shall take the elements $u_{i,j} = \Delta_{i,j}$ if $i \neq j$, while $u_{ii} = \Delta_{ii} + m - i$. Taking this we find the determinant of the matrix

$$\begin{pmatrix} \Delta_{1,1} + m - 1 & \Delta_{1,2} & \ldots & \Delta_{1,m-1} & \Delta_{1,m} \\ \Delta_{2,1} & \Delta_{2,2} + m - 2 & \ldots & \Delta_{2,m-1} & \Delta_{2,m} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \Delta_{m-1,1} & \Delta_{m-1,2} & \ldots & \Delta_{m-1,m-1} + 1 & \Delta_{m-1,m} \\ \Delta_{m,1} & \Delta_{m,2} & \ldots & \Delta_{m,m-1} & \Delta_{m,m} \end{pmatrix}. \tag{4.1.4}$$

We let $C = C_m$ be the value of this determinant computed as previously explained and call it *the Capelli element*. We have assumed on purpose that m, n are not necessarily equal and we will get a computation of C similar to what would happen in the commutative case for the determinant of a product of rectangular matrices.

Rule. When we compute such determinants of noncommuting elements, we should remark that some of the usual formal laws of determinants are still valid:

(1) The determinant is *linear* in the rows and columns.
(2) The determinant changes sign when transposing two rows.

Note that property (2) is NOT TRUE when we transpose columns!

In fact it is useful to develop a commutator identity. Recall that in an associative algebra R the commutator $[r, s] := rs - sr$ satisfies the following (easy) identity:

$$[r, s_1 s_2 \cdots s_k] = \sum_{i=1}^{k} s_1 \cdots s_{i-1}[r, s_i]s_{i+1} \cdots s_k. \tag{4.1.5}$$

A formula follows:

Proposition.

$$[r, \det(\mathcal{U})] = \sum_{i=1}^{m} \det(\mathcal{U}_i), \tag{4.1.6}$$

where $\mathcal{U}_i$ is the matrix obtained from $\mathcal{U}$ by replacing its i^{th} column c_i with $[r, c_i]$.

There is a rather peculiar lemma that we will need. Given an $m \times m$ matrix $\mathcal{U}$ in noncommutative variables $u_{i,j}$, let us consider an integer k, and for all $t = 1, \dots, m$, let us set $\mathcal{U}_{t,k}$ to be the matrix obtained from $\mathcal{U}$ by setting equal to 0 the first k entries of the t^{th} row.

Lemma. $\sum_{t=1}^{m} \det(\mathcal{U}_{t,k}) = (m - k)\det(\mathcal{U})$.

Proof. By induction on k. For $k = 0$, it is trivial. Remark that:

$$\det(\mathcal{U}_{t,k}) = \det(\mathcal{U}_{t,k-1}) - \det\begin{pmatrix} u_{11} & \dots & u_{1,k-1} & 0 & u_{1,k+1} & \dots & u_{1m} \\ \dots & \dots & \dots & \dots & & & \\ u_{t-1,1} & \dots & u_{t-1,k-1} & 0 & u_{t-1,k+1} & \dots & u_{t-1,m} \\ 0 & \dots & 0 & u_{t,k} & 0 & \dots & 0 \\ u_{t+1,1} & \dots & u_{t+1,k-1} & 0 & u_{t+1,k+1} & \dots & u_{t+1,m} \\ \dots & \dots & \dots & \dots & & & \\ u_{m1} & u_{m2} & \dots & 0 & \dots & \dots & u_{mn} \end{pmatrix}.$$

Thus

$$\sum_{t=1}^{m} \det(\mathcal{U}_{t,k}) = \sum_{t=1}^{m} \det(\mathcal{U}_{t,k-1}) - \det(\mathcal{U}) = (m - k + 1)\det(\mathcal{U}) - \det(\mathcal{U})$$
$$= (m - k)\det(\mathcal{U}).$$

□

Theorem. *The value of the Capelli element is given by*

$$C = \begin{cases} 0 & \text{if } m > n, \\ \det(X)\det(Y) & \text{if } m = n, \\ \sum_{1 \le i_1 < i_2 < \cdots < i_m \le n} X_{i_1 i_2 \cdots i_m} Y_{i_1 i_2 \cdots i_m} & \text{if } m < n \quad \text{(Binet's form)}, \end{cases}$$

where by $X_{i_1 i_2 \cdots i_m}$, respectively $Y_{i_1 i_2 \cdots i_m}$, we mean the determinants of the minors formed by the columns of indices $i_1 < i_2 < \cdots < i_m$ of X, respectively by the rows of indices $i_1 < i_2 < \cdots < i_m$ of Y.

Proof. Let us give the proof of Capelli's theorem. It is based on a simple computational idea. We start by introducing a new $m \times n$ matrix Ξ of indeterminates $\xi_{i,j}$, disjoint from the previous ones, and perform our computation in the algebra of differential operators in the variables x, ξ. Consider now the product ΞY. It is a product of matrices with elements in a commutative algebra, so one can apply to it the usual rules of determinants to get

(4.1.7)

$$\det(\Xi Y) = \begin{cases} 0, & \text{if } m > n, \\ \det(\Xi)\det(Y) & \text{if } m = n, \\ \sum_{1 \le i_1 < i_2 < \cdots < i_m \le n} \Xi_{i_1 i_2 \cdots i_m} Y_{i_1 i_2 \cdots i_m} & \text{if } m < n \quad \text{(Binet's form)}. \end{cases}$$

The matrix ΞY is an $m \times m$ matrix of differential operators. Let us use the notations:

$$\eta_{i,j} = \sum_{h=1}^{n} \xi_{ih} \frac{\partial}{\partial x_{jh}}, \quad \gamma_{i,j} = \sum_{h=1}^{n} x_{ih} \frac{\partial}{\partial \xi_{jh}}, \quad D_{i,j} = \sum_{h=1}^{n} \xi_{ih} \frac{\partial}{\partial \xi_{jh}}.$$

We now apply to both sides of the identity 4.1.17 the product of commuting differential operators $\gamma_i := \gamma_{ii} = \sum_{h=1}^{n} x_{ih} \frac{\partial}{\partial \xi_{ih}}$. The operator γ_i, when applied to a function linear in $\underline{\xi}_i$, has the effect of replacing $\underline{\xi}_i$ with $\underline{x}_i$. In particular we can apply the operators under consideration to functions which depend only on the $\underline{x}$ and not on the $\underline{\xi}$. For such a function we get

$$\begin{cases} \prod_{i=1}^{m} \gamma_i \sum_{\sigma \in S_n} \epsilon_\sigma \eta_{1,\sigma(1)} \eta_{2,\sigma(2)} \cdots \eta_{n,\sigma(n)} f(\underline{x}) = 0 & \text{if } m > n, \\ \det(X)\det(Y) f(\underline{x}) & \text{if } m = n, \\ \sum_{1 \le i_1 < i_2 < \cdots < i_m \le n} X_{i_1 i_2 \cdots i_m} Y_{i_1 i_2 \cdots i_m} f(\underline{x}) & \text{if } m < n. \end{cases}$$

So it is only necessary to show that

$$\prod_{i=1}^{m} \gamma_i \sum_{\sigma \in S_n} \epsilon_\sigma \eta_{1,\sigma(1)} \eta_{2,\sigma(2)} \cdots \eta_{n,\sigma(n)} f(\underline{x}) = C f(\underline{x}).$$

We will prove this formula by induction, by introducing the determinants C_{k-1} of the matrices:

$$C_{k-1} := \begin{pmatrix} \Delta_{1,1} + m - 1 & \Delta_{1,2} & \dots & \Delta_{1,m-1} & \Delta_{1,m} \\ \dots & \dots & \dots & \dots & \dots \\ \Delta_{k-1,1} & \dots & \Delta_{k-1,k-1} + m - k + 1 & \dots & \Delta_{k-1,m} \\ \eta_{k,1} & \eta_{k,2} & \dots & \eta_{k,m-1} & \eta_{k,m} \\ \dots & \dots & \dots & \dots & \dots \\ \eta_{m,1} & \eta_{m,2} & \dots & \eta_{m,m-1} & \eta_{m,m} \end{pmatrix}.$$

We want to show that for every k:

$$\prod_{i=1}^{k} \gamma_i \sum_{\sigma \in \mathcal{S}_n} \epsilon_\sigma \eta_{1,\sigma(1)} \eta_{2,\sigma(2)} \cdots \eta_{n,\sigma(n)} f(\underline{x}) = C_k f(\underline{x}).$$

By induction, this means that we need to prove

$$\gamma_k C_{k-1} f(\underline{x}) = C_k f(\underline{x}). \tag{4.1.8}$$

Since $\gamma_k f(\underline{x}) = 0$ this is equivalent to proving that $[\gamma_k, C_{k-1}] f(\underline{x}) = C_k f(\underline{x})$.

By 4.1.6 we have $[\gamma_k, C_{k-1}] = \sum_{j=1}^{m} C_{k-1}^{j}$ where C_{k-1}^{j} is the determinant of the matrix obtained from C_{k-1} by substituting its j^{th} column c_j with the commutator $[\gamma_k, c_j]$.

In order to prove 4.1.8 we need the following easy commutation relations:

$$[\gamma_i, \eta_{hk}] = \delta_h^i \Delta_{ik} - \delta_k^i D_{hi}, \quad [\gamma_i, \Delta_{hk}] = -\delta_k^i \gamma_{hi}.$$

Thus we have

$$[\gamma_k, c_j] = \begin{vmatrix} 0 \\ 0 \\ \vdots \\ \Delta_{kj} \\ \vdots \\ 0 \end{vmatrix}, \ j \neq k, \qquad [\gamma_k, c_k] = \begin{vmatrix} -\gamma_{1k} \\ \vdots \\ -\gamma_{k-1,k} \\ \Delta_{kk} - D_{kk} \\ -D_{k+1,k} \\ \vdots \\ -D_{mk} \end{vmatrix}.$$

By the linearity in the rows we see that $[\gamma_k, C_{k-1}] f(\underline{x}) = (C_k + U) f(\underline{x})$, where U is the determinant of the matrix obtained from C_{k-1} by substituting its k^{th} column with

$$\begin{vmatrix} -\gamma_{1k} \\ \vdots \\ -\gamma_{k-1,k} \\ -(m-k) - D_{kk} \\ -D_{k+1,k} \\ \vdots \\ -D_{mk} \end{vmatrix}.$$

We need to show that $Uf(\underline{x}) = 0$. In the development of $Uf(\underline{x})$, the factor D_{kk} is applied to a function which does not depend on $\underline{\xi}_k$, and so it produces 0. We need to analyze the determinants ($s = 1, \dots, k-1, \quad j = 1, \dots, m-k$):

$$U_s := \begin{pmatrix} \Delta_{1,1} + m - 1 & \Delta_{1,2} & \dots & 0 & \dots & \Delta_{1,m} \\ \dots & \dots & \dots & -\gamma_{s,k} & \dots & \dots \\ \Delta_{k-1,1} & \dots & \dots & 0 & \dots & \Delta_{k-1,m} \\ \eta_{k,1} & \eta_{k,2} & \dots & 0 & \dots & \eta_{k,m} \\ \dots & \dots & \dots & 0 & \dots & \\ \eta_{m,1} & \eta_{m,2} & \dots & 0 & \dots & \eta_{m,m} \end{pmatrix},$$

$$U'_j := \begin{pmatrix} \Delta_{1,1} + m - 1 & \Delta_{1,2} & \dots & 0 & \dots & \Delta_{1,m} \\ \dots & \dots & \dots & 0 & \dots & \dots \\ \Delta_{k-1,1} & \dots & \dots & 0 & \dots & \Delta_{k-1,m} \\ \eta_{k,1} & \eta_{k,2} & \dots & 0 & \dots & \eta_{k,m} \\ \dots & \dots & \dots & -D_{k+j,k} & \dots & \\ \eta_{m,1} & \eta_{m,2} & \dots & 0 & \dots & \eta_{m,m} \end{pmatrix}.$$

For the first U_s, exchange the k and s rows, changing sign, and notice that in the development of $U_s f(\underline{x})$ the only terms that intervene are the ones in which the factor $\gamma_{s,k}$ is followed by one of the $\eta_{k,t}$ since the other terms do not depend on $\underline{\xi}_k$. For these terms $\gamma_{s,k}$ commutes with all factors except for $\eta_{k,t}$ and the action of $\gamma_{s,k}\eta_{k,t}$, is the same as of $\Delta_{s,t}$

For the last U'_j, exchange the k and $k+j$ rows, changing sign, and notice that in the development of $U'_j f(\underline{x})$, the only terms that intervene are the ones in which the factor $D_{k+j,k}$ is followed by one of the $\eta_{k,t}$, since the other terms do not depend on $\underline{\xi}_k$. For these terms $D_{k+j,k}$ commutes with all factors except for $\eta_{k,t}$, and the action of $D_{k+j,k}\eta_{k,t}$ is the same as that of $\eta_{k+j,t}$. We need to interpret the previous terms. Let $\mathcal{A}$ be the matrix obtained from C_{k-1} dropping the k^{th} row and k^{th} column, and A its determinant. We have that

$$Uf(\underline{x}) = -(m-k)Af(\underline{x}) + \left(\sum_{s+1}^{k-1} U_s + \sum_{j=1}^{m-k} U'_j\right) f(\underline{x}).$$

We have then that the determinants U_s, U'_j are the $m-1$ terms $\det(\mathcal{A}_{t,k})$ which sum to $(m-k)A$, and thus by the previous lemma we finally obtain 0. □

In 5.2 we will see how to obtain a generalization of the Clebsch–Gordan formula from the Capelli identities.

5 Primary Covariants

5.1 The Algebra of Polarizations

We want to deduce some basic formulas which we will interpret in representation theory later. The reader who is familiar with the representation theory of semisimple Lie algebras will see here the germs of the theory of highest weights, central characters, the Poincaré–Birkhoff–Witt, the Peter–Weyl and Borel–Weil theorems.

The theory is due to Capelli and consists in the study the algebra of differential operators $\mathcal{U}_m(n)$, generated by the elements $\Delta_{ij} = \sum_{h=1}^n x_{ih}\frac{\partial}{\partial x_{jh}},\ i, j = 1, \dots, m.$

A similar approach can also be found in the book by Deruyts ([De]).

One way of understanding $\mathcal{U}_m(n)$ is to consider it as a subalgebra of the algebra $\mathbb{C}\big[x_{ih}, \frac{\partial}{\partial x_{jk}}\big],\ i, j = 1, \dots, m,\ h, k = 1, \dots, n$, of all polynomial differential operators. The first observation is the computation of their commutation relations:

$$[\Delta_{ij}, \Delta_{hk}] = \delta_h^j \Delta_{ik} - \delta_k^i \Delta_{hj}. \tag{5.1.1}$$

In the language that we will develop in the next chapter this means that the m^2 operators Δ_{ij} span a *Lie algebra* (Chapter 4, 1.1), in fact the Lie algebra gl_m of the general linear group.

As for the associative algebra $\mathcal{U}_m(n)$, it can be thought of as a representation of the universal enveloping algebra of gl_m (Chapter 5, §7).

If we list the m^2 operators Δ_{ij} as $u_1, u_2, \dots, u_{m^2}$ in some given order, we see immediately by induction that the monomials $u_1^{h_1}u_2^{h_2}\cdots u_{m^2}^{h_{m^2}}$ are linear generators of $\mathcal{U}_m$.

Theorem. *If $n \geq m$, the monomials $u_1^{h_1}u_2^{h_2}\cdots u_{m^2}^{h_{m^2}}$ are a linear basis of $\mathcal{U}_m(n)$.*[13]

Proof. We can prove this theorem using the method of computing the symbol of a differential operator. The symbol of a polynomial differential operator in the variables $x_i, \frac{\partial}{\partial x_j}$ is obtained in an elementary way, by taking the terms of higher degree in the derivatives and substituting for the operators $\frac{\partial}{\partial x_j}$ the commutative variables ξ_j. In a more formal way, one can filter the algebra of operators by the degree in the derivatives and take the associated graded algebra. The symbol of Δ_{ij} is $\sum_{h=1}^n x_{ih}\xi_{jh}$. If $n \geq m$ these polynomials are algebraically independent and this implies the linear independence of the given monomials. $\square$

Remark. If $n < m$, the previous theorem does not hold. For instance, we have the Capelli element which is 0. One can analyze all of the relations in a form which is the noncommutative analogue of determinantal varieties.

The previous theorem implies that as an abstract algebra, $\mathcal{U}_m(n)$ is independent of n, for $n \geq m$. We will denote it by $\mathcal{U}_m$. $\mathcal{U}_m$ is the *universal enveloping algebra* of gl_m (cf. Chapter 5, §7).

5.2 Primary Covariants

Let us now introduce the notion of a *primary covariant*.[14] We will see much later that this is the germ of the theory of highest weight vectors.

Let X be as before a rectangular $m \times n$ matrix of variables.

[13] This implies the Poincaré–Birkhoff–Witt theorem for linear Lie algebras.

[14] The word covariant has a long history and we refer the reader to the discussion in Chapter 15.

Definition. A $k \times k$ minor of the matrix X is primary if it is extracted from the first k rows of X. A primary covariant is a product of determinants of primary minors of the matrix X.

Let us introduce a convenient notation for the primary covariants. Given a number $p \leq m$ and p indices $1 \leq i_1 < i_2 < \cdots < i_p \leq n$, denote by the symbol $|i_1, i_2, \ldots, i_p|$ the determinant of the $p \times p$ minor of the matrix X extracted from the first p rows and the columns of indices $i_1, i_2, \ldots, i_p$. By definition a primary covariant is a product of polynomials of type $|i_1, i_2, \ldots, i_p|$.

If we perform on the matrix X an elementary operation consisting of adding to the i^{th} row a linear combination of the preceding rows, we see that the primary covariants do not change. In other words primary covariants are invariant under the action of the group U^- of strictly lower triangular matrices acting by left multiplication. We shall see in Chapter 13 that the converse is also true: a polynomial in the entries of X invariant under U^- is a primary covariant.

Then, the decomposition of the ring of polynomials in the entries of X as a representation of $GL(m) \times GL(n)$, together with the highest weight theory, explains the true nature of primary covariants.

Let $f(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_m)$ be a function of the vector variables $\underline{x}_i$, such that f is multihomogeneous of degrees $\underline{h} := (h_1, h_2, \ldots, h_m)$. Let us denote by $\mathcal{U}_m$ the algebra generated by the operators $\Delta_{i,j},\ i, j \leq m$.

Lemma. *If $i \neq j \leq p$, then Δ_{ij} commutes with $|i_1, i_2, \ldots, i_p|$ while*

$$\Delta_{ii}|i_1, i_2, \ldots, i_p| = |i_1, i_2, \ldots, i_p|(1 + \Delta_{ii}).$$

Proof. Immediate. □

Theorem (Capelli–Gordan expansion). *Given a multidegree $\underline{h}$, there exist elements $C_i(\underline{h}), D_i(\underline{h}) \in \mathcal{U}_m$ depending only on the multidegree $\underline{h}$, such that for any function f of multidegree $\underline{h}$ we have*

$$f = \sum_i C_i(\underline{h}) D_i(\underline{h}) f. \tag{5.2.1}$$

Moreover, for all i, the polynomial $D_i(\underline{h}) f$ is a primary covariant.

Proof. We apply first induction on the total degree. For a fixed total degree N, we apply induction on the set of multidegrees ordered opposite to the lexicographic order. Thus the basis of this second induction is multidegree $(N, 0, 0, \ldots, 0)$, i.e., for a function only in the variable $\underline{x}_1$ which is clearly a primary covariant. In general, assume that the degrees $h_i = 0$ for $i > k$. For $h_k = p$, we apply the Capelli identities expressing the value of the Capelli determinant C_p (associated to the $\Delta_{i,j},\ i, j \leq p$). If $p > n$, we get $0 = C_p f$, and developing the determinant C_p from right to left we get a sum

$$0 = (\Delta_{11} + p - 1) \cdots (\Delta_{pp}) f + \sum_{i<p} A_{ij} \Delta_{ip} f$$

$$= \prod_{i=1}^{p} (h_i + p - i) f + \sum_{i<p} A_{ij} \Delta_{ip} f,$$

where the A_{ip} are explicit elements of $\mathcal{U}_m$. The functions $\Delta_{ip} f$ have multidegree higher than that of f, and we can apply induction.

If $p \leq n$ we also apply the Capelli identity, but now we have

$$\sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} X_{i_1 i_2 \cdots i_p} Y_{i_1 i_2 \cdots i_p} f = (\Delta_{11} + p - 1) \cdots (\Delta_{pp}) f + \sum_{i<p} A_{ij} \Delta_{ip} f$$

from which

$$f = \prod_{i=1}^{p} (h_i + p - i)^{-1} \Big(\sum_{1 \leq i_1 < i_2 < \cdots < i_p \leq n} X_{i_1 i_2 \cdots i_p} Y_{i_1 i_2 \cdots i_p} f - \sum_{i<p} A_{ip} \Delta_{ip} f \Big).$$

Now the induction can be applied to both the functions $\Delta_{ip} f$ and $Y_{i_1 i_2 \cdots i_p} f$. Next one has to use the commutation rules of the elements $X_{i_1 i_2 \cdots i_p}$ with the Δ_{ij}, which imply that the elements $X_{i_1 i_2 \cdots i_p}$ contribute determinants which are primary covariants. □

A classical application of this theorem gives another approach to computing invariants. Start from a linear action of a group G on a space and take several copies of this representation. We have already seen that the polarization operators commute with the group G and so preserve the property of being invariant. In Section 2.4 we have discussed a possible strategy of computing invariants by reducing to multilinear ones with polarization.

The Capelli–Gordan expansion offers a completely opposite strategy. In the expansion 5.2.1, if f is a G-invariant so are the primary covariants $D_i(\underline{h}) f$. Thus we deduce that each invariant function of several variables can be obtained under polarization from invariant primary covariants, in particular, from invariants which depend at most on n vector variables. We will expand on these ideas in Chapter 11, §1 and §5, where they will be set forth in the language of representation theory.

5.3 Cayley's Ω Process

There is one more classical computation which we shall reconsider in the representation theory of the linear group.

Let us take $m = n$ so that the two matrices X, Y defined in 4.1 are square matrices.

By definition, $D := \det(X) = |1, 2, \ldots, n|$, while $\det(Y)$ is a differential operator classically denoted by Ω and called *Cayley's Ω process.* We have $C = |1, 2, \ldots, n| \Omega = D\Omega$.

From Lemma 5.2 we have the commutation relations:

(5.3.1) $$[\Delta_{ij}, D] = 0, \ i \neq j, \quad [\Delta_{ii}, D] = D.$$

We have a similar result for Ω:

Lemma. *If $i \neq j$, then Δ_{ij} commutes with Ω, while*

$$\Delta_{ii}\Omega = \Omega(\Delta_{ii} - 1). \tag{5.3.2}$$

Proof. Let us apply 4.1.6. The operator Δ_{ij} commutes with all of the columns of Ω except for the i^{th} column ω_i with entries $\frac{\partial}{\partial x_{it}}$. Now $\left[\Delta_{ij}, \frac{\partial}{\partial x_{it}}\right] = -\frac{\partial}{\partial x_{jt}}$, from which $[\Delta_{ij}, \omega_i] = -\omega_j$. The result follows immediately. □

Let us introduce a more general determinant, analogous to a characteristic polynomial. We denote it by $C_m(\rho) = C(\rho)$ and define it as

$$\begin{pmatrix} \Delta_{1,1} + m - 1 + \rho & \Delta_{1,2} & \dots & \Delta_{1,m} \\ \Delta_{2,1} & \Delta_{2,2} + m - 2 + \rho & \dots & \Delta_{2,m} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \Delta_{m-1,1} & \Delta_{m-1,2} & \dots & \Delta_{m-1,m} \\ \Delta_{m,1} & \Delta_{m,2} & \dots & \Delta_{m,m} + \rho \end{pmatrix}. \tag{5.3.3}$$

We have now a generalization of the Capelli identity:

Proposition.

$$\Omega C(k) = C(k+1)\Omega, \quad |1, 2, \dots, m| C(k) = C(k-1)|1, 2, \dots, m|, \tag{5.3.4}$$

(5.3.5)
$$D^k\Omega^k = C(-(k-1))C(-(k-2)) \cdots C(-1)C, \ \Omega^k D^k = C(k)C(k-1) \cdots C(1).$$

Proof. We may apply directly 5.3.1 and 5.3.2 and then proceed by induction. □

We now develop $C_m(\rho)$ as a polynomial in ρ, obtaining an expression

$$C_m(\rho) = \rho^m + \sum_{i=1}^{m} K_i \rho^{m-i}. \tag{5.3.6}$$

Theorem.

(i) *The elements K_i generate the center of the algebra $\mathcal{U}_m$, generated by the elements Δ_{ij}, $i, j = 1, \dots, m$.*
(ii) *The operator $C_m(\rho)$ applied to the polynomial $\prod_{i=1}^m |1, 2, \dots, i-1, i|^{k_i}$ multiplies it by $\prod_j (\ell_j + m - j + \rho)$, where the numbers ℓ_j are the multidegrees of the given polynomial and $\ell_j := \sum_{i \geq j} k_i$.*

Proof. To prove (i) we begin by proving that the K_i belong to the center of $\mathcal{U}_m$. From 5.3.1, 5.3.2, and 5.3.4 it follows that, for every k, the element $C(k)$ is in the center of $\mathcal{U}_m$. But then this easily implies that all the coefficients of the polynomial are also in the center. To prove that they generate the center we need some more theory, so we postpone the proof to Chapter 5, §7.2, only sketching its steps here. Let $\sigma_{ij} := \sum_{h=1}^m x_{ih}\xi_{jh}$ be the symbol of Δ_{ij}. We think of the σ_{ij} as coordinates of

$m \times m$ matrices. One needs to prove first that the symbol of a central operator is an invariant function under conjugation.

Next one shows that the coefficients of the characteristic polynomial of a matrix generate the invariants under conjugation. Finally, the symbols of the Capelli elements are the coefficients of the characteristic polynomial up to a sign, and then one finishes by induction.

Now we prove (ii). We have, by the definition of polarizations, that $\Delta_{hk}|1, 2, \ldots, i-1, i| = 0$ if $h < k$, or if $k > i$, while $\Delta_{hh}|1, 2, \ldots, i-1, i|^j = j|1, 2, \ldots, i-1, i|^j$, $h \leq i$. Therefore it follows that

$$\begin{aligned} \Delta_{hh} \prod |1, 2, \ldots, i-1, i|^{k_i} &= \Big(\sum_{i \geq h} k_i\Big) \prod |1, 2, \ldots, i-1, i|^{k_i} \\ &= \ell_h \prod |1, 2, \ldots, i-1, i|^{k_i}. \end{aligned}$$

The development of the determinant of 4.1.7 applied to $\prod_{i=1}^m |1, 2, \ldots, i-1, i|^{k_i}$ gives 0 for all of the terms in the upper triangle and thus it is equal to the expression

$$\begin{aligned} &\prod_{i=1}^{m} (\Delta_{ii} + m - i + \rho) \prod_{i=1}^{m} |1, 2, \ldots, i-1, i|^{k_i} \\ &= \prod_{i=1}^{m} (\ell_i + m - i + \rho) \prod_{i=1}^{m} |1, 2, \ldots, i-1, i|^{k_i}. \qquad \square \end{aligned}$$

The significance of this theorem is that the polynomials we have considered are highest weight vectors for irreducible representations of the group $GL(m)$ over which, by Schur's lemma, the operators K_i must act as scalars. Thus we obtain the explicit value of the central character (cf. Chapter 11, §5).