Hanspeter Kraft
Claudio Procesi

CLASSICAL INVARIANT THEORY

A Primer

Preliminary Version, July 1996

Table of Contents

§ 1 Invariant Functions and Covariants

In this paragraph we recall some elementary facts from algebraic geometry and from representation theory. Moreover, we give the basic notions of invariant theory like the ring of invariants and the module of covariants, and explain a number of easy examples. Finally, we describe two important Finiteness Theorems for the ring of invariant polynomial functions on a representation space W of a group G. The first one goes back to HILBERT and states that the invariant ring is finitely generated in case where the linear representation of G on the coordinate ring of W is completely reducible. The second is due to E. NOETHER and shows that for a finite group G the invariant ring is generated by the invariants of degree less or equal to the order of the group G.

1.1 Polynomial functions. In the following K will always denote an infinite field. Let W be a finite dimensional K-vector space. A function $f\colon W \to K$ is called *polynomial* or *regular* if it is given by a polynomial in the coordinates with respect to a basis of W. It is easy to see that this is independent of the choice of a coordinate system of W. We denote by $K[W]$ the K-algebra of polynomial functions on W which is usually called the *coordinate ring* of W or the *ring of regular functions* on W. If $w_1, \dots, w_n$ is a basis of W and $x_1, \dots, x_n$ the dual basis of the dual vector space W^* of W, i.e., the *coordinate functions*, we have $K[W] = K[x_1, \dots, x_n]$. This is a polynomial ring in the x_i because the field K is infinite (see the following exercise).

> **Exercise 1.** Show that the coordinate functions $x_1, \dots, x_n \in K[W]$ are algebraically independent. Equivalently, if $f(a_1, a_2, \dots, a_n) = 0$ for a polynomial f and all $a = (a_1, a_2, \dots, a_n) \in K^n$ then f is the zero polynomial. (Hint: Use induction on the number of variables and the fact that a non-zero polynomial in one variable has only finitely many zeroes.)

A regular function $f \in K[W]$ is called *homogeneous of degree* d if $f(tw) = t^d f(w)$ for all $t \in K$, $w \in W$. Thus $K[W] = \oplus_d K[W]_d$ is a graded K-algebra where $K[W]_d$ denotes the subspace of homogeneous polynomials of degree d. (Recall that an algebra $A = \oplus_i A_i$ is *graded* if the multiplication satisfies $A_i A_j \subset A_{i+j}$). Choosing coordinates as above we see that the monomials $x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n}$ such that $d_1 + d_2 + \cdots + d_n = d$ form a basis of $K[W]_d$.

We have $K[W]_1 = W^*$ and this extends to a canonical identification of $K[W]_d$ with the dth *symmetric power* $\mathrm{S}^d(W^*)$ of W^*:

$$K[W] = \bigoplus_d K[W]_d = \bigoplus_d \mathrm{S}^d(W^*) = \mathrm{S}(W^*)$$

where $\mathrm{S}(W^*)$ is the *symmetric algebra* of W^*.

1.2 Invariants. As usual, we denote by $\mathrm{GL}(W)$ the *general linear group*, i.e., the group of K-linear automorphisms of the K-vector space W. Choosing a basis $(w_1, w_2, \dots, w_n)$ of W we can identify $\mathrm{GL}(W)$ with the group $\mathrm{GL}_n(K)$ of invertible $n \times n$ matrices with entries in K in the usual way: The ith column of the matrix A corresponding to the automorphism $g \in \mathrm{GL}(W)$ is the coordinate vector of $g(w_i)$ with respect to the chosen basis.

Now assume that there is given a subgroup $G \subset \mathrm{GL}(W)$ or, more generally, a group G together with a *linear representation on* W, i.e., a group homomorphism

$$\rho\colon G \to \mathrm{GL}(W).$$

The corresponding *linear action* of G on W will be denoted by $(g, w) \mapsto gw := \rho(g)w$ ($g \in G$, $w \in W$), and we will call W a *G-module*. In the following, representations of groups will play a central role. We assume that the reader is familiar with the basic notion and elementary facts from representation theory. (See Appendix A for a short introduction.)

Exercises

2. Let $\rho\colon G \to \mathrm{GL}(W)$ be a finite dimensional representation and $\rho^*\colon G \to \mathrm{GL}(W^*)$ the *dual* (or *contragredient*) representation, i.e., $\rho^*(g)(\lambda)\,w := \lambda(\rho(g)^{-1}\,w)$, $\lambda \in W^*, w \in W$. Choosing a basis in W and the dual basis in W^* one has the following relation for the corresponding matrices $A = \rho(g)$ and $A^* = \rho^*(g)$: $A^* = (A^{-1})^t$.

3. Show that the natural representation of $\mathrm{SL}_2(K)$ on K^2 is *selfdual* (i.e., equivalent to its dual representation) by finding an invertible matrix S such that $(A^{-1})^t = S\,A\,S^{-1}$ for all $A \in \mathrm{SL}_2$.

Definition. A function $f \in K[W]$ is called *G-invariant* or shortly *invariant* if $f(gw) = f(w)$ for all $g \in G$ and $w \in W$. The invariants form a subalgebra of $K[W]$ called *invariant ring* and denoted by $K[W]^G$.

Recall that the *orbit* of $w \in W$ is defined to be the subset $Gw := \{gw \mid g \in G\} \subset W$ and the *stabilizer* of w it the subgroup $G_w := \{g \in G \mid gw = w\}$. It is clear that a function is G-invariant if and only if it is constant on all orbits of G in W. A subset $X \subset W$ is called *G-stable* if it is a union of orbits, i.e., if one has $gx \in X$ for all $x \in X$, $g \in G$.

Exercise 4. The natural representation of SL_2 (and GL_2) on K^2 has two orbits. The stabilizer of $e_1 := (1, 0)$ is $U := \left\{ \left(\begin{smallmatrix} 1 & s \\ 0 & 1 \end{smallmatrix}\right) \middle| s \in K \right\}$, the subgroup of *upper triangular unipotent matrices*. For any other point $(x, y) \neq (0, 0)$ the stabilizer is conjugate to U.

There is another way to describe the invariant ring. For this we consider the following linear action of G on the coordinate ring $K[W]$, generalizing the dual representation on the linear functions (see Exercise 2):

$$(g, f) \mapsto gf, \quad gf(w) := f(g^{-1}w) \text{ for } g \in G,\ f \in K[W],\ w \in W.$$

This is usually called the *regular representation* of G on the coordinate ring. (The inverse g^{-1} in this definition is necessary in order to get a *left*-action on the space of functions.) Clearly, a function f is invariant if and only if it is a *fixed point* under this action, i.e., $gf = f$ for all $g \in G$. This explains the notation $K[W]^G$ for the ring of invariants. Moreover, it follows from the definition that the subspaces $K[W]_d$ are stable under the action. Hence, the invariant ring decomposes into a direct sum $K[W]^G = \oplus_d K[W]_d^G$ and thus is a graded K-algebra, too.

Exercise 5. Consider the linear action of GL_2 on $K[x, y]$ induced by the natural representation of GL_2 on K^2.

(a) What is the image of x and y under $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(K)$?

(b) Show that $K[x, y]^{\mathrm{GL}_2} = K[x, y]^{\mathrm{SL}_2} = K$.

(c) Show that $K[x, y]^U = K[y]$ where $U \subset \mathrm{SL}_2$ is the subgroup of upper triangular unipotent matrices (see Exercise 4).

Example 1. We start with the two-dimensional representation of the *multiplicative group* $K^* := \mathrm{GL}_1(K)$ on $W = K^2$ given by $t \mapsto \left(\begin{smallmatrix} t & 0 \\ 0 & t^{-1} \end{smallmatrix}\right)$. Then the invariant ring is generated by xy: $K[W]^{K^*} = K[xy]$. In fact, the subspaces $Kx^a y^b \subset K[W]$ are all stable under K^*, and $t(x^a y^b) = t^{b-a} x^a y^b$.

Example 2. Next we consider the *special linear group* $\mathrm{SL}_n(K)$, i.e., the subgroup of $\mathrm{GL}_n(K)$ of matrices with determinant 1, and its representation on the space $\mathrm{M}_n(K)$ of $n \times n$-matrices by left multiplication: $(g, A) \mapsto gA$, $g \in \mathrm{SL}_n$, $A \in \mathrm{M}_n(K)$. Clearly, the *determinant function* $A \mapsto \det A$ is invariant. In fact, we claim that *the invariant ring is generated by the determinant:*

$$K[\mathrm{M}_n]^{\mathrm{SL}_n} = K[\det].$$

Outline of Proof: Let f be an invariant. Define the polynomial $p \in K[t]$ by $p(t) := f(\begin{pmatrix} t & & \\ & 1 & \\ & & \ddots \end{pmatrix})$. Then $f(A) = p(\det A)$ for all invertible matrices A, because A can be written in the form

$$A = g \begin{pmatrix} \det A & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \quad \text{where} \quad g \in \mathrm{SL}_n(K)$$

and f is invariant. We will see in the next section that $\mathrm{GL}_n(K)$ is Zariski-dense in $\mathrm{M}_n(K)$ (Lemma 1.3) which implies that $f(A) = p(\det A)$ for all matrices A. Thus $f \in K[\det]$. □

Exercises

6. Determine the invariant rings $K[\mathrm{M}_2(K)]^U$ and $K[\mathrm{M}_2(K)]^T$ under left multiplication by the subgroup U of upper triangular unipotent matrices (see Exercise 4) and the subgroup $T := \left\{ \left(\begin{smallmatrix} t & 0 \\ 0 & t^{-1} \end{smallmatrix} \right) \middle| t \in K^* \right\}$ of diagonal matrices of SL_2.

7. Let $T_n \subset \mathrm{GL}_n(K)$ be the subgroup of invertible diagonal matrices. If we choose the standard basis in $V := K^n$ and the dual basis in V^* we can identify the coordinate ring $K[V \oplus V^*]$ with $K[x_1, \dots, x_n, \zeta_1, \dots, \zeta_n]$. Show that $K[V \oplus V^*]^{T_n} = K[x_1\zeta_1, x_2\zeta_2, \dots, x_n\zeta_n]$. What happens if one replaces T_n by the subgroup T_n' of diagonal matrices with determinant 1?
(Hint: All monomials in $K[x_1, \dots, x_n, \zeta_1, \dots, \zeta_n]$ are eigenvectors for T_n.)

Example 3. Let $\mathcal{S}_n$ denote the *symmetric group* on n letters and consider the natural representation of $\mathcal{S}_n$ on $V = K^n$ given by $\sigma(e_i) = e_{\sigma(i)}$, or, equivalently,

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

As above, $\mathcal{S}_n$ acts on the polynomial ring $K[x_1, x_2, \dots, x_n]$ and the invariant functions are the *symmetric polynomials*:

$$K[x_1, \dots, x_n]^{\mathcal{S}_n} = \{f \mid f(x_{\sigma(1)}, \dots) = f(x_1, \dots) \text{ for all } \sigma \in \mathcal{S}_n\}.$$

It is well known and classical that every symmetric function can be expressed uniquely as a polynomial in the elementary symmetric functions $\sigma_1, \sigma_2, \dots, \sigma_n$ defined by

$$\begin{aligned}
\sigma_1 &:= x_1 + x_2 + \cdots + x_n, \\
\sigma_2 &:= x_1x_2 + x_1x_3 + \cdots x_{n-1}x_n, \\
&\ \vdots \\
\sigma_k &:= \sum_{i_1 < i_2 < \cdots < i_k} x_{i_1}x_{i_2}\cdots x_{i_k}, \\
&\ \vdots \\
\sigma_n &:= x_1x_2\cdots x_n.
\end{aligned}$$

We will give a proof of this below.

Proposition. *The elementary symmetric functions $\sigma_1, \sigma_2, \dots, \sigma_n$ are algebraically independent and generate the algebra of symmetric functions:*

$$K[x_1, x_2, \dots, x_n]^{\mathcal{S}_n} = K[\sigma_1, \sigma_2, \dots, \sigma_n].$$

Proof: We proof this by induction on n. Let $\sigma'_1, \sigma'_2, \dots, \sigma'_{n-1}$ denote the elementary symmetric functions in the variables $x_1, x_2, \dots, x_{n-1}$. Then

$$\begin{aligned} \sigma_1 &= \sigma'_1 + x_n, \\ \sigma_2 &= \sigma'_2 + x_n\sigma'_1, \\ &\vdots \\ \sigma_{n-1} &= \sigma'_{n-1} + x_n\sigma'_{n-2}, \\ \sigma_n &= x_n\sigma'_{n-1}, \end{aligned}$$

hence $\sigma_i \in K[\sigma'_1, \dots, \sigma'_{n-1}, x_n]$. Assume that the σ_i's are algebraically dependent and let $F(\sigma_1, \sigma_2, \dots, \sigma_n) = 0$ be an algebraic relation of minimal degree. Setting $x_n = 0$ we obtain the relation $F(\sigma'_1, \sigma'_2, \dots, \sigma'_{n-1}, 0) = 0$ between the σ'_i, hence $F(z_1, \dots, z_{n-1}, 0) = 0$ by induction. This implies that F is divisible by x_n which contradicts the minimality.

Now let $f \in K[x_1, \dots, x_n]$ be a symmetric polynomial. Since every homogeneous component of f is symmetric, too, we can assume that f is homogeneous of some degree N. If we write f in the form $f = \sum_i f_i(x_1, \dots, x_{n-1})x_n^i$ then all f_i are symmetric in $x_1, \dots, x_{n-1}$ and so, by induction,

$$f_i \in K[\sigma'_1, \dots, \sigma'_{n-1}] \subset K[\sigma_1, \dots, \sigma_n, x_n].$$

Thus f has the form $f = p(\sigma_1, \dots, \sigma_n) + x_n h(\sigma_1, \dots, \sigma_n, x_n)$ with two polynomials p and h. Again we can assume that $p(\sigma_1, \dots, \sigma_n)$ and $h(\sigma_1, \dots, \sigma_n, x_n)$ are both homogeneous, of degree N and $N-1$, respectively. It follows that $f-p$ is again homogeneous and is divisible by x_n. Since it is symmetric, it is divisible by the product $x_1x_2\cdots x_n$, i.e. $f - p = \sigma_n\bar{f}$ with a symmetric polynomial $\bar{f}$ of degree at most $N - n$. Now the claim follows by induction on the degree of f. □

Exercises

8. Consider the following symmetric functions $n_j := x_1^j + x_2^j + \cdots + x_n^j$ called *power sums* or Newton *functions*.

(a) Prove the following formulas due to Newton:

$$(-1)^{j+1}j\sigma_j = n_j - \sigma_1 n_{j-1} + \sigma_2 n_{j-2} - \cdots + (-1)^{j-1}\sigma_{j-1}n_1$$

for all $j = 1, \dots, n$.

(Hint: The case $j = n$ is easy: Consider $f(t) := \prod_i(t - x_i)$ and calculate $\sum_i f(x_i)$ which is equal to 0. For $j < n$, the right hand side is a symmetric function of degree $\leq j$, hence can be expressed as a polynomial in $\sigma_1, \dots, \sigma_j$. Now put $x_{j+1} = \ldots = x_n = 0$ and use induction on n. Another proof can be found in [Wey46].)

(b) Show that in characteristic 0 the power sums $n_1, n_2, \dots, n_n$ generate the symmetric functions.

9. From the natural representation of $\mathrm{GL}_2(K)$ on $W := K^2$ we get a linear action of $\mathrm{GL}_2(K)$ on the coordinate ring $K[W] = K[x, y]$ (cf. Exercise 5). If $\operatorname{char} K = 0$ then the representations of SL_2 on the homogeneous components $V_n := K[x, y]_n$ $(n = 0, 1, \dots)$ are all irreducible.
The V_n are the classical *binary forms* of degree n.

10. Show that a representation $\rho\colon G \to \mathrm{GL}(W)$ is selfdual if and only if there exists a G-invariant non-degenerate bilinear form $B\colon W \times W \to K$. Use this to give another solution to Exercise 3.

11. Let $f = \sum_{i=0}^{n} a_i x^i y^{n-i}$ and $h = \sum_{j=0}^{n} b_j x^j y^{n-j}$ be two binary forms of degree n (see Exercise 9). Show that for suitable $\gamma_0, \dots, \gamma_n \in \mathbb{Q}$ the bilinear form $B(f, h) := \gamma_0\, a_0 b_n + \gamma_1\, a_1 b_{n-1} + \cdots + \gamma_n\, a_n b_0$ on $V_n \times V_n$ is SL_2-invariant and non-degenerate. In particular, all the V_n are selfdual representations of SL_2.

1.3 Zariski-dense subsets. For many purposes the following notion of “density” turns out to be very useful.

Definition. A subset X of a finite dimensional vector space W is called Zariski-*dense* if every function $f \in K[W]$ which vanishes on X is the zero function. More generally, a subset $X \subset Y$ $(\subset W)$ is called Zariski-*dense* in Y if every function $f \in K[W]$ which vanishes on X also vanishes on Y.

In other words every polynomial function $f \in K[W]$ is completely determined by its restriction $f|_X$ to a Zariski-dense subset $X \subset W$. Denote by $I(X)$ the ideal of functions vanishing on $X \subset W$:

$$I(X) := \{f \in K[W] \mid f(a) = 0 \text{ for all } a \in X\}.$$

$I(X)$ is called the *ideal of* X. Clearly, we have $I(X) = \bigcap_{a \in X} \mathfrak{m}_a$ where $\mathfrak{m}_a = I(\{a\})$ is the the maximal ideal of functions vanishing in a, i.e., the kernel of the *evaluation homomorphism* $\varepsilon_a\colon K[W] \to K$, $f \mapsto f(a)$. It is called the *maximal ideal of* a. (Choosing coordinates we get $\mathfrak{m}_a = (x_1 - a_1, x_2 - a_2, \dots, x_n - a_n)$.) It is clear from the definition above that a subset $X \subset Y$ $(\subset W)$ is Zariski-dense in Y if and only if $I(X) = I(Y)$.

Remark. Let $X \subset Y$ be two subsets of the K-vector space W. For every field extension L/K we have $W \subset W_L := L \otimes_K W$. If X is Zariski-dense in Y it is not a priori clear that this is also the case if X and Y are considered as subsets of W_L. In order to prove this it suffices to show that $L \otimes_K I_K(X) = I_L(X)$. This is obvious if X is a point, hence $I_K(X) = \mathfrak{m}_a$. The general case follows from the description of $I(X)$ above as an intersection of maximal ideals (using Exercise 12).

Exercises

12. Let V be a K-vector space, not necessarily finite dimensional, let L/K be a field extension and $U_i \subset V$ $(i \in I)$ a family of subspaces. Then

$$L \otimes_K (\bigcap_{i \in I} U_i) = \bigcap_{i \in I} (L \otimes_K U_i) \subset L \otimes_K V.$$

13. Let $k \subset K$ be an infinite subfield. Then k^n is ZARISKI-dense in K^n.

A subset $X \subset W$ defined by polynomial equations is characterized by the property that it is not ZARISKI-dense in any strictly larger subset. We call such a subset ZARISKI-*closed* and its complement ZARISKI-*open*.

Lemma. *Let $h \in K[W]$ be a non-zero function and define $W_h := \{w \in W \mid h(w) \neq 0\}$. Then W_h is* ZARISKI-*dense in W.*

PROOF: In fact, if f vanishes on W_h then fh vanishes on W, hence $fh = 0$. Since h is nonzero we must have $f = 0$. □

A typical example of a ZARISKI-dense subset is $\mathrm{GL}_n(K) = \mathrm{M}_n(K)_{\det} \subset \mathrm{M}_n(K)$. It was used in the Example 2 of the previous section 1.2.

Exercise 14. Let $k \subset K$ be an infinite subfield. Then $\mathrm{GL}_n(k)$ is ZARISKI-dense in $\mathrm{GL}_n(K)$. Moreover, $\mathrm{SL}_n(k)$ is ZARISKI-dense in $\mathrm{SL}_n(K)$.
(Hint: The first statement follows from Exercise 13 and the Remark above. For the second statement use the map $A \mapsto \begin{pmatrix} \det A & & \\ & 1 & \\ & & \ddots \end{pmatrix}^{-1} A$.)

Example. Let W be a G-module and assume that G has a ZARISKI-dense orbit in W. Then every invariant function is constant: $K[W]^G = K$. A typical example is the natural representation of $\mathrm{SL}(V)$ on

$$V^r := \underbrace{V \oplus V \oplus \cdots \oplus V}_{r \text{ times}}$$

for $r < \dim V = n$. In fact, using coordinates this corresponds to left multiplication of the $n \times r$-matrices by SL_n.

In 1.2 Example 2 we have seen that for $r = n$ the invariants are generated by the determinant function det. This implies again that there are no non-constant invariants for $r < n$ because the restriction of det to the $n \times r$-matrices vanishes for $r < n$. (We use here the fact that every invariant of $r < n$ copies is a restriction an invariant on n copies; cf. Exercise 27 in 1.5).

The general problem of describing the invariants for arbitrary r is solved by the *First Fundamental Theorem for* SL_n. We will discuss this in §7 and §8 (see

7.5 and 8.4).

Exercises

15. Let $X \subset Y \subset W$ and assume that X is Zariski-dense in Y. Then the linear spans $\langle X \rangle$ and $\langle Y \rangle$ are equal.

16. Let $H \subset G \subset \mathrm{GL}(W)$ be subgroups and assume that H is Zariski-dense in G. Then a linear subspace $U \subset W$ is H-stable if and only if it is G-stable. Moreover, we have $K[W]^H = K[W]^G$.

17. Show that $K[V \oplus V^*]^{\mathrm{GL}(V)} = K[q]$ where the bilinear form q is defined by $q(v, \zeta) := \zeta(v)$ (cf. Exercise 7).
(Hint: The subset $Z := \{(v, \zeta) \mid \zeta(v) \neq 0\}$ of $V \oplus V^*$ is Zariski-dense. Fix a pair (v_0, ζ_0) such that $\zeta_0(v_0) = 1$. Then for every $(v, \zeta) \in Z$ there is a $g \in \mathrm{GL}(V)$ such that $g(v, \zeta) = (v_0, \lambda\zeta_0)$ where $\lambda = \zeta(v)$.)

1.4 Covariants. Let W, V be two (finite dimensional) vector spaces over K. A map $\varphi\colon W \to V$ is called *polynomial* or a *morphism* if the coordinate functions of φ with respect to some basis of V are polynomial functions on W. It is obvious that this does not depend on the choice of the basis. As an example consider the canonical map from W to the r-fold tensor product $W \otimes W \otimes \cdots \otimes W$ given by $w \mapsto w \otimes w \otimes \cdots \otimes w$ or the canonical map from W to the rth symmetric power $\mathrm{S}^r W$ given by $w \mapsto w^r$.

Given a morphism $\varphi\colon W \to V$ every polynomial function f on V determines—by composition—a polynomial function $\varphi^*(f) := f \circ \varphi$ on W. Thus, we obtain an algebra homomorphism $\varphi^*\colon K[V] \to K[W]$ called the *comorphism* of φ which completely determines the morphism φ: Choosing coordinates in V we can identify V with K^n and $K[V]$ with $K[y_1, y_2, \ldots, y_n]$ and then the ith component of $\varphi\colon W \to K^n$ is given by $\varphi_i = \varphi^*(y_i)$.

Exercises

18. The map $\varphi \mapsto \varphi^*$ defines a bijection between the set of morphisms $W \to V$ and the set of algebra homomorphisms $K[V] \to K[W]$. Moreover, the following holds:

(a) If $\psi\colon V \to U$ is another morphism then $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$.

(b) φ^* is injective if and only if the image $\varphi(W)$ is Zariski-dense in V.

(c) If φ^* is surjective then φ is injective.

19. Let $\varphi\colon V \to W$ be a morphism of vector spaces and let $X \subset Y \subset V$ be subsets where X is Zariski-dense in Y. Then $\varphi(X)$ is Zariski-dense in $\varphi(Y)$.

20. A morphism $\varphi\colon W \to V$ is called *homogeneous of degree d* if $\varphi(\lambda w) = \lambda^d \varphi(w)$ for all $w \in W$, $\lambda \in K$ (cf. 1.1).

(a) φ is homogeneous of degree d if and only if $\varphi^*(V^*) \subset K[W]_d$. Equivalently, all components of φ with respect to any basis of V are homogeneous functions of the same degree.

(b) Every morphism can be uniquely written as a sum of homogeneous morphisms. (The summands are called the *homogeneous components* of φ.)

21. Denote by $K[\mathrm{SL}_2]$ the algebra of functions $f|_{\mathrm{SL}_2}$ where f is a polynomial function on M_2.

(a) Show that the kernel of the restriction map res: $K[\mathrm{M}_2] \to K[\mathrm{SL}_2]$ is the ideal generated by the determinant det:

$$K[\mathrm{SL}_2] = K[\bar{a}, \bar{b}, \bar{c}, \bar{d}] \xleftarrow{\sim} K[a, b, c, d]/(ad - bc - 1).$$

(Hint: Assume first that K is algebraically closed. The general case follows by showing that $\mathrm{SL}_2(K)$ is Zariski-dense in $\mathrm{SL}_2(\bar{K})$; see Exercise 14.)

(b) The subgroup $U \subset \mathrm{SL}_2$ of upper triangular unipotent matrices (Exercise 4) acts on SL_2 by left multiplication. Show that $K[\mathrm{SL}_2]^U = K[c, d]$.

22. The orbit map $\mathrm{SL}_2 \to K^2$, $g \mapsto ge_1$ (see Exercise 4) identifies the (regular) functions on K^2 with the U-invariant functions on SL_2 where U acts on SL_2 by right multiplication: $(u, g) \mapsto gu^{-1}$.

Now assume that W and V are both G-modules. Generalizing the notion of invariant functions we introduce the concept of *equivariant morphisms*, i.e., polynomial maps $\varphi\colon W \to V$ satisfying $\varphi(gw) = g\varphi(w)$ for all $g \in G, w \in W$, which leads to the classical concept of *covariants* (or *concomittants*).

Definition. Let W and V be two G-modules. A *covariant of W of type V* is a G-equivariant polynomial map $\varphi\colon W \to V$.

Examples 1. Both examples mentioned above, $W \to W \otimes W \otimes \cdots \otimes W$ and $W \to \mathrm{S}^r W$ are covariants with respect to the group $\mathrm{GL}(W)$.
Another example arises from matrix multiplication: Consider the action of $\mathrm{GL}_n(K)$ on the $n \times n$-matrices $\mathrm{M}_n := \mathrm{M}_n(K)$ by conjugation. Then the power maps $A \mapsto A^i$ are covariants of type M_n.

Working with covariants rather than only with invariants offers a number of interesting new constructions. For instance, covariants can be composed (see Exercise 18). In particular, if $\varphi\colon W \to V$ is a covariant and $f \in K[V]^G$ an invariant then the composition $f \circ \varphi$ is an invariant of W. Another constructions arises in connection with tensor products. Let V_1, V_2 be two G-modules and let $p\colon V_1 \otimes V_2 \to U$ be a linear projection onto some other G-module U. For example, assume that the tensor product $V_1 \otimes V_2$ is completely reducible and that U is a direct summand. If φ_1, φ_2 are covariants of type V_1 and V_2, respectively, then we obtain a covariant $(\varphi_1, \varphi_2)_U$ of type U by composing in the following way:

$$(\varphi_1, \varphi_2)_U : W \xrightarrow{(\varphi_1,\varphi_2)} V_1 \times V_2 \to V_1 \otimes V_2 \xrightarrow{p} U$$

This construction is classically called *transvection* (in German: *Überschiebung*). It was an important tool in the 19th century invariant theory of binary forms. We will discuss this in detail in the last chapter.

Example 2. Assume $\operatorname{char} K \neq 2$ and consider the standard representation of SL_2 on $V = K^2$. Then $V \otimes V = \mathrm{S}^2 V \oplus K$ and the projection $p \colon V \otimes V \to K$ is given by $(x_1, x_2) \otimes (y_1, y_2) \mapsto x_1 y_2 - y_1 x_2$. Given two covariants φ, ψ of type V we thus obtain an invariant by transvection: $(\varphi, \psi)_K = \varphi_1 \psi_2 - \varphi_2 \psi_1$.

A special case of this construction is the multiplication of a covariant $\varphi \colon W \to V$ with an invariant $f \in K[W]$ which is again a covariant of type V, denoted by $f\varphi$. In this way we see that the covariants of a fixed type form *a module over the ring of invariants*.

Example 3. For the group $\mathrm{GL}_n(K)$ an interesting example arises from matrix multiplication (which can be considered as a linear projection $p \colon \mathrm{M}_n \otimes \mathrm{M}_n \to \mathrm{M}_n$): For any GL_n-module W two covariants of type M_n can be multiplied. Thus the covariants of type M_n form even a (non-commutative) algebra over the ring of invariants.

Given a covariant $\varphi \colon W \to V$ of type V the comorphism φ^* defines—by restriction—a G-homomorphism $V^* \to K[W] \colon \lambda \mapsto \lambda \circ \varphi$, which we also denote by φ^* . Clearly, the comorphism and hence φ is completely determined by this linear map.

Proposition. *Let W and V be G-modules. The covariants of W of type V are in bijective correspondence with the G-homomorphisms $V^* \to K[W]$.*

Proof: By standard properties of the polynomial ring we know that the algebra homomophisms $K[V] \to K[W]$ are in 1-1 correspondence with the K-linear maps $V^* \to K[W]$. Thus, there is a natural bijection between the morphisms $W \to V$ and the linear maps $V^* \to K[W]$ (see Exercise 18) which clearly induces a bijection between the subset of G-equivariant morphisms and the subset of G-homomorphisms. □

This proposition shows that the study of covariants of a given G-module W which was an important task in the classical literature corresponds in our modern language to the determination of the G-module structure of the coordinate ring $K[W]$.

Exercises

23. Let φ be a covariant of type V. Then every homogeneous component is again a covariant of type V (see Exercise 20).

24. Let G be $\mathrm{GL}(W)$ or $\mathrm{SL}(W)$. Show that the only covariants of W of type W are the scalar multiplications $t \cdot \mathrm{id}\colon W \to W$, $t \in K$.

25. Let W be a representation of SL_2 and let U the subgroup of upper triangular unipotent matrices. There is an isomorphism $K[W \oplus K^2]^{\mathrm{SL}_2} \xrightarrow{\sim} K[W]^U$ given by $f \mapsto \bar{f}$ where $\bar{f}(w) := f(w,(1,0))$.

(Hint: The inverse map is constructed in the following way: For $h \in K[W]^U$ define $F\colon W \times \mathrm{SL}_2 \to K$ by $F(w,g) := h(g^{-1}w)$. Then (a) $F(w,gu) = F(w,g)$ for all $u \in U$ and (b) $F(hw,hg) = F(w,g)$ for all $h \in \mathrm{SL}_2$. It follows from (a) that F defines a function $\bar{F}$ on $W \times K^2$ (see Exercise 22) which is SL_2-invariant by (b).)

1.5 Classical Invariant Theory. One of the fundamental problems in Classical Invariant Theory (shortly CIT) is the following:

Problem. *Describe generators and relations for the ring of invariants* $K[W]^G$.

This question goes back to the 19th century and a number of well-known mathematicians of that time have made important contributions: Boole, Sylvester, Cayley, Hermite, Clebsch, Gordan, Capelli, Hilbert. We refer the reader to some classical books on invariant theory, like Gordan's "Vorlesungen" [Gor87], "The algebra of Invariants" by Grace and Young [GrY03], Weitzenböck's "Invariantentheorie" [Wei23] and of course the famous "Classical Groups" of Weyl [Wey46] (see also [Gur64] and [Sch68]). Many important notions from modern algebra, like Noetherian Theory, Hilbert's Syzygies, the "Basissatz" and the "Nullstellensatz", were introduced in connection with this problem. Modern treatments of the subject can be found in the books of Dieudonné-Carrell [DiC70], Fogarty [Fog69], Kraft [Kra85] and Springer [Spr77].

One of the main general results in this context is the so called *First Fundamental Theorem* (shortly FFT). It says that the simultaneous invariants of a large number of copies of a given representation W can all be obtained from n copies by "polarization" where $n = \dim W$. We will discuss this in detail in the following paragraphs. There is also a *Second Fundamental Theorem* (SFT) which makes some general statements about the relations. We will say more about this in the second and the third chapter.

Exercises 26. Let the group $\mathbb{Z}_2 = \{\mathrm{id}, \sigma\}$ act on the finite dimensional

K-vector space V by $\sigma(v) = -v$ ($\operatorname{char} K \neq 2$). Determine a system of generators for the ring of invariants $K[V]^\sigma = K[V]^{\mathbb{Z}_2}$.
(Hint: The invariants are the polynomials of even degree.)

27. Let W be a G-module and $V \subset W$ a G-stable subspace. Assume that V has a G-stable complement: $W = V \oplus V'$. Then every invariant function $f \in K[V]^G$ is the restriction $\tilde{f}|_V$ of an invariant function $\tilde{f} \in K[W]$.
Use the subgroup $U \subset \mathrm{SL}_2$ (Exercise 4) to show that the claim above does not hold for every stable subspace V.

28. Consider the two finite subgroups

$$C_n := \left\{ \left(\begin{smallmatrix} \zeta & 0 \\ 0 & \zeta^{-1} \end{smallmatrix} \right) \,\middle|\, \zeta^n = 1 \right\} \text{ and } \tilde{D}_{2n} := C_{2n} \cup \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right) \cdot C_{2n}$$

of $\mathrm{SL}_2(\mathbb{C})$. Find generators and relations for the invariant rings $\mathbb{C}[x,y]^{C_n}$ and $\mathbb{C}[x,y]^{\tilde{D}_{2n}}$.

29. Let L/K be a field extension. For any K-vector space V we put $V_L := V \otimes_K L$. If G is a group and V a G-module then V_L is also a G-module.

(a) Show that ${V_L}^G = (V^G)_L$.

(b) For the invariants we have $L[V]^G = L \otimes_K K[V]^G$. In particular, a subset $S \subset K[V]^G$ is a system of generators if and only if it generates $L[V]^G$.

(c) If $U \subset V$ is a G-submodule and if U_L has a G-stable complement in V_L then U has a G-stable complement in V.
(Hint: Consider the natural map $\mathrm{Hom}(V,U) \to \mathrm{Hom}(U,U)$ and use that $\mathrm{Hom}(V_L, W_L) = \mathrm{Hom}(V,W)_L$.)

(d) If the representation of G on V_L is completely reducible then so is the representation on V.

Remark. In general, we work over an arbitrary (infinite) field K. But sometimes it is convenient to replace K by its algebraic closure $\bar{K}$ and to use geometric arguments. We have already seen above that for a representation $\rho\colon G \to \mathrm{GL}(W)$ on a K-vector space W we always have

$$\bar{K} \otimes_K K[W]^G = \bar{K}[W_{\bar{K}}]^G \quad \text{where } W_{\bar{K}} := \bar{K} \otimes_K W$$

(Exercise 29). On the right hand side of the equation we can even replace G by a larger group $\tilde{G}$ with a representation $\tilde{\rho}\colon \tilde{G} \to \mathrm{GL}(W_{\bar{K}})$ provided that $\rho(G)$ is ZARISKI-dense in $\tilde{\rho}(\tilde{G})$ (see 1.3 Exercise 16; cf. Remark 1.3). A typical example is $G = \mathrm{GL}_n(K)$ (or $\mathrm{SL}_n(K)$) and $\tilde{G} = \mathrm{GL}_n(\bar{K})$ (or $\mathrm{SL}_n(\bar{K})$) (cf. 1.3 Exercise 14).

1.6 Some Finiteness Theorems. One of the highlights of the 19th century invariant theory was GORDAN's famous Theorem showing that the invariants (and covariants) of binary forms (under SL_2) are finitely generated ([Gor68]).

His proof is rather involved and can be roughly described as follows: He gives a general inductive method to construct all invariants, and then he shows that after a certain number of steps the construction does not produce any new invariant. Thus, the finite number of invariants constructed so far form a system of generators.

It was already clear at that time that it will be very difficult to generalize GORDAN's method to other groups than SL_2. So it came as a big surprise when HILBERT presented in 1890 his general finiteness result for invariants, using completely new ideas and techniques ([Hil90], [Hil93]). In the following we give a modern formulation of his result, but the basic ideas of the proof are HILBERT's.

Theorem 1 (HILBERT). *Let W be a G-module and assume that the representation of G on the coordinate ring $K[W]$ is completely reducible. Then the invariant ring $K[W]^G$ is finitely generated.*

OUTLINE OF PROOF: Since the representation of G on $K[W]$ is completely reducible there is a canonical G-equivariant linear projection $R\colon K[W] \to K[W]^G$ which is the identity on $K[W]^G$. This projection is usually called REYNOLDS *operator* (see Exercise 30 below). It is easy to see that $R(hf) = hR(f)$ for $h \in K[W]^G$, $f \in K[W]$. It follows that for every ideal $\mathfrak{a} \subset K[W]^G$ we have $R(K[W]\mathfrak{a}) = K[W]\mathfrak{a} \cap K[W]^G = \mathfrak{a}$. Now we start with the homogeneous maximal ideal $\mathfrak{m}_0 := \oplus_{d>0} K[W]^G_d$ of $K[W]^G$. By HILBERT's Basis Theorem (see [Art91]) the ideal $K[W]\mathfrak{m}_0$ of $K[W]$ is finitely generated, i.e., there are homogeneous polynomials $f_1, \dots, f_s \in \mathfrak{m}_0$ such that $K[W]\mathfrak{m}_0 = (f_1, \dots, f_s)$. But then $\mathfrak{m}_0 = R(K[W]\mathfrak{m}_0)$, as an ideal of $K[W]^G$, is also generated by $f_1, \dots, f_s$. Now it is not difficult to show that any homogeneous system of generators of $\mathfrak{m}_0$ is also a system of generators for the invariant ring $K[W]^G$ (see Exercise 31 below). □

The proof above shows that the invariant ring $K[W]^G$ is *Noetherian*, i.e., every ascending chain of ideals becomes stationary, or equivalently, every ideal is finitely generated.

Exercises

30. Let A be a (commutative) algebra and let G be a group of algebra automorphisms of A. Assume that the representation of G on A is completely reducible. Then the subalgebra A^G of invariants has a canonical G-stable complement and the corresponding G-equivariant projection $p\colon A \to A^G$ satisfies the relation $p(hf) = h\,p(f)$ for $h \in A^G, f \in A$.

31. Let $A = \oplus_{i \geq 0} A_i$ be a graded K-algebra, i.e., $A_i A_j \subset A_{i+j}$. Assume that the ideal $A^+ := \oplus_{i>0} A_i$ is finitely generated. Then the algebra A is finitely generated as an algebra over A_0. More precisely, if the ideal A^+ is generated by the homogeneous elements $a_1, a_2, \dots, a_n$ then $A = A_0[a_1, a_2, \dots, a_n]$.

This result of Hilbert can be applied to the case of finite groups G as long as the characterstic of K is prime to the order of G (Theorem of Maschke, see [Art91] Chap. 9, Corollary 4.9). In characteristic zero there is a more precise result due to Emmy Noether ([Noe16], 1916) which gives an explicit bound for the degrees of the generators.

Theorem 2 (E. Noether)**.** *Assume* $\operatorname{char} K = 0$. *For any representation W of a finite group G the ring of invariants $K[W]^G$ is generated by the invariants of degree less or equal to the order of G.*

Proof: Choose a basis in W and identify $K[W]$ with the polynomial ring $K[x_1, x_2, \dots, x_n]$. For any n-tuple $\mu = (\mu_1, \mu_2, \dots, \mu_n)$ of non-negative integers define the following invariant:

$$j_\mu := \sum_{g \in G} g(x_1^{\mu_1} x_2^{\mu_2} \cdots x_n^{\mu_n}).$$

Now let $f = \sum_\mu a_\mu x_1^{\mu_1} x_2^{\mu_2} \cdots x_n^{\mu_n}$ be any invariant. Then $|G| \cdot f = \sum_{g \in G} gf = \sum_\mu a_\mu j_\mu$. Thus we have to show that $K[W]^G$ is generated by the j_μ's where $|\mu| := \mu_1 + \cdots + \mu_n \leq |G|$. For that purpose consider the following polynomials p_j where $j \in \mathbb{N}$:

$$p_j(x_1, \dots, x_n, z_1, \dots, z_n) := \sum_{g \in G} (gx_1 \cdot z_1 + gx_2 \cdot z_2 + \cdots + gx_n \cdot z_n)^j.$$

Clearly, we have $p_j = \sum_{|\rho| = j} j_\rho \cdot z_1^{\rho_1} z_2^{\rho_2} \cdots z_n^{\rho_n}$. By Exercise 8 we see that each p_j with $j > |G|$ can be expressed as a polynomial in the p_i's for $i \leq |G|$. This implies that the invariants j_ρ for $|\rho| > |G|$ can be written as polynomials in the j_μ's where $\mu \leq |G|$, and the claim follows. □

In connection with this result Schmid introduced in [Sch89, Sch91] a numerical invariant $\beta(G)$ for every finite group G. It is defined to be the minimal number m such that for every representation W of G the invariant ring $K[W]^G$ is generated by the invariants of degree less or equal to m. By Noether's Theorem above we have $\beta(G) \leq |G|$. Schmid shows (loc. cit.) that $\beta(G) = |G|$ *if and only if G is cyclic.* In general, it is rather difficult to calculate $\beta(G)$, except for small groups. For example,

$$\beta(\mathbb{Z}/2 \times \mathbb{Z}/2) = 3, \quad \beta(\mathcal{S}_3) = 4, \quad \beta(\mathcal{S}_4) = 10. \quad \beta(D_{2n}) = n + 1$$

where D_{2n} denotes the dihedral group of order $2n$. For the symmertic group $\mathcal{S}_n$ we can find a lower bound by looking at large cyclic subgroups. Denote by $\gamma(n)$ the maximal order of an element of $\mathcal{S}_n$. Then we have

$$\beta(\mathcal{S}_n) \geq \gamma(n) \quad \text{and} \quad \ln \gamma(n) \sim \sqrt{n \ln n}$$

where $f(n) \sim g(n)$ means that $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$ (see [Mil87]). In particular, $\beta(\mathcal{S}_n)$ grows more rapidly than any power of n.

For finite *abelian* groups G the invariant $\beta(G)$ coincides with the so-called Davenport constant ([GeS92], see Exercise 32 below).

Exercises (The results of the following two exercises are due to Schmid [Sch89, Sch91].)

32. Let G be a finite abelian group (written additively). Define the Davenport constant $\delta(G)$ to be the length m of the longest non-shortable expression

$$0 = g_1 + g_2 + \cdots + g_m, \quad g_i \in G.$$

("Non-shortable" means that no strict non-empty subset of the g_i's has sum zero.)
(a) Show that $\delta(G) = \beta(G)$.
(b) Show that $\delta(G) = |G|$ if and only if G is cyclic.
(c) Calculate $\delta((\mathbb{Z}/2)^n)$.

33. Let $H \subset G$ be a subgroup. Then $\beta(G) \leq [G : H]\beta(H)$. If H is normal then $\beta(G) \leq \beta(H)\beta(G/H)$.

§ 2 First Fundamental Theorem for the General Linear Group

We discuss the so-called *First Fundamental Theorem for* GL_n which describes a minimal system of generators for the "invariants of p vectors and q covectors", i.e., the invariant polynomial functions on p copies of the vector space V and q copies of the dual space V^* with respect to the natural linear action of $\mathrm{GL}(V)$. This result has an interesting geometric interpretation.

Then we study the invariants of several copies of the endomorphism ring $\mathrm{End}(V)$ under simultaneous conjugation and describe a set of generators. This gives the *First Fundamental Theorem for Matrices*. We will see in §4 that the two First Fundamental Theorems are strongly related.

2.1 Invariants of vectors and covectors. Let V be a finite dimensional K-vector space. Consider the representation of $\mathrm{GL}(V)$ on the vector space

$$W := \underbrace{V \oplus \cdots \oplus V}_{p \text{ times}} \oplus \underbrace{V^* \oplus \cdots \oplus V^*}_{q \text{ times}} =: V^p \oplus V^{*q},$$

consisting of p copies of V and q copies of its dual space V^*, given by

$$g(v_1, \dots, v_p, \varphi_1, \dots, \varphi_q) := (gv_1, \dots, gv_p, g\varphi_1, \dots, g\varphi_q)$$

where $g\varphi_i$ is defined by $(g\varphi_i)(v) := \varphi_i(g^{-1}v)$. This representation on V^* is the *dual* or *contragredient* representation of the standard representation of $\mathrm{GL}(V)$ on V (cf. 1.2 Exercise 2). The elements of V are classically called *vectors*, those of the dual space V^* *covectors*. We want to describe the invariants of $V^p \oplus V^{*q}$ under this action. (The easy case $p = q = 1$ was treated in 1.3 Exercise 17.) For every pair (i, j), $i = 1, \dots, p$, $j = 1, \dots, q$, we define the bilinear function $(i \mid j)$ on $V^p \oplus V^{*q}$ by

$$(i \mid j)\colon (v_1, \dots, v_p, \varphi_1, \dots, \varphi_q) \mapsto (v_i \mid \varphi_j) := \varphi_j(v_i).$$

These functions are usually called *contractions*. They are clearly invariant:

$$(i \mid j)(g(v, \varphi)) = (g\varphi_j)(gv_i) = \varphi_j(g^{-1}gv_i) = (i \mid j)(v, \varphi).$$

Now the First Fundamental Theorem (shortly FFT) states that these functions generate the ring of invariants. The proof will be given in 4.7 after some preparation in §3 and §4.

First Fundamental Theorem for $\mathrm{GL}(V)$**.** *The ring of invariants for the action of* $\mathrm{GL}(V)$ *on* $V^p \oplus V^{*q}$ *is generated by the invariants* $(i \mid j)$*:*

$$K[V^p \oplus V^{*q}]^{\mathrm{GL}(V)} = K[(i \mid j) \mid i = 1, \dots, p,\ j = 1, \dots, q].$$

Using coordinates this amounts to the following. Fix a basis in V and the dual basis in V^* and write $v_i \in V$ as a column vector and $\varphi_j \in V^*$ as a row vector:

$$v_i = \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix}, \qquad \varphi_j = (\varphi_{j1}, \dots, \varphi_{jn}).$$

Then $X := (v_1, \dots, v_p)$ is a $n \times p$-matrix and $Y := \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_q \end{pmatrix}$ is a $q \times n$-matrix, and we obtain a canonical identification:

$$V^p \oplus V^{*q} = \mathrm{M}_{n\times p}(K) \oplus \mathrm{M}_{q\times n}(K).$$

The corresponding action of $g \in \mathrm{GL}_n(K)$ on the matrices is given by $g(X,Y) = (gX, Yg^{-1})$. Now consider the map

$$\Psi\colon \mathrm{M}_{n\times p} \times \mathrm{M}_{q\times n} \to \mathrm{M}_{q\times p}, \quad (X,Y) \mapsto YX.$$

Then the (i,j)-component of Ψ is

$$\Psi_{ij}(X,Y) = \sum_{\nu=1}^{n} \varphi_{i\nu} v_{\nu j} = (v_j \mid \varphi_i), \quad \text{i.e., } \Psi_{ij} = (j \mid i).$$

Moreover, the map Ψ is constant on orbits: $\Psi(g(X,Y)) = \Psi(gX, Yg^{-1}) = Yg^{-1}gX = YX = \Psi(X,Y)$. Thus we see again that $(i \mid j)$ is an invariant.

2.2 Geometric interpretation. We can give a geometric formulation of the FFT using the language of algebraic geometry. First we remark that the image of the map Ψ is the subset $V^n_{q\times p} \subset \mathrm{M}_{q\times p}$ of matrices of rank $\leq n$ (see Exercise 1 below). In fact, $V^n_{q\times p}$ is even a *closed subvariety* which means that it is the zero set of a family of polynomials (see below and Exercise 2). Now the FFT says that the map

$$\Psi\colon \mathrm{M}_{n\times p} \times \mathrm{M}_{q\times n} \to V^n_{q\times p}$$

is "universal" in the sense that any morphism $\Phi\colon \mathrm{M}_{n\times p} \times \mathrm{M}_{q\times n} \to Z$ into an affine variety Z which is constant on orbits factors through Ψ, i.e., there is a unique morphism $\bar\Phi\colon V^n_{q\times p} \to Z$ such that $\Phi = \overline{\Phi}\Psi$:

$$\begin{array}{ccc} \mathrm{M}_{n\times p} \times \mathrm{M}_{q\times n} & \xrightarrow{\Psi} & V^n_{q\times p} \\ \big\downarrow \Phi & & \big\downarrow \bar\Phi \\ Z & = & Z \end{array}$$

Thus we see that $V^n_{q\times p}$ is an algebraic analogue to the *orbit space* of the action which is usually denoted by $(\mathrm{M}_{n\times p} \times \mathrm{M}_{q\times n})/\,\mathrm{GL}_n$. In our situation we say that

$\Psi\colon \mathrm{M}_{n\times p}\times\mathrm{M}_{q\times n}\to V^n_{q\times p}$ is an *algebraic quotient* (with respect to the action of GL_n) and use the notation $V^n_{q\times p} = (\mathrm{M}_{n\times p}\times\mathrm{M}_{q\times n})/\!\!/\,\mathrm{GL}_n$. By construction, the *quotient map* Ψ induces an isomorphism

$$\Psi^*\colon K[V^n_{q\times p}] \xrightarrow{\sim} K[\mathrm{M}_{n\times p}\times\mathrm{M}_{q\times n}]^{\mathrm{GL}_n},$$

where $K[V^n_{q\times p}]$ is the coordinate ring of $V^n_{q\times p}$ i.e., the K-algebra of all restrictions $f|_{V^n_{q\times p}}$, $f\in K[\mathrm{M}_{q\times p}]$.

The subvariety $V^n_{q\times p}\subset\mathrm{M}_{q\times p}$ is called a *determinantal variety* because it is defined by the vanishing of all $(n+1)\times(n+1)$-minors (Exercise 2). Since the invariant ring is clearly an integrally closed domain it follows from the above that $V^n_{q\times p}$ is a *normal* variety. (See the following exercises.)

Exercises

1. Show that every matrix $C\in\mathrm{M}_{q\times p}$ of rank $\le n$ can be written as a product $C=AB$ with a $(q\times n)$-matrix A and a $(n\times p)$-matrix B.

2. Show that for any n the set of $p\times q$-matrices of rank $\le n$ forms a closed subvariety of $\mathrm{M}_{p\times q}$, i.e., it is the set of zeroes of some polynomials.
(Consider the $n+1\times n+1$-minors.)

3. Let $\rho\colon G\to\mathrm{GL}(W)$ be a finite dimensional representation of a group G. Then the ring of invariants $K[W]^G$ is normal, i.e., integrally closed in its field of fractions.

2.3 Invariants of conjugacy classes. Consider the linear action of $\mathrm{GL}(V)$ on $\mathrm{End}(V)$ by conjugation. The orbits of this action are the *conjugacy classes* of matrices. For an element $A\in\mathrm{End}(V)$ we write its *characteristic polynomial* in the form

$$P_A(t)=\det(tE-A)=t^n+\sum_{i=1}^{n}(-1)^i s_i(A)t^{n-i}$$

where $n:=\dim V$ and $E\in\mathrm{End}(V)$ is the identity. This shows that the s_i are invariant polynomial functions on $\mathrm{End}(V)$.

Proposition. *The ring of invariants for the conjugation action of* $\mathrm{GL}(V)$ *on* $\mathrm{End}(V)$ *is generated by* $s_1,s_2,\dots,s_n$*:*

$$K[\mathrm{End}(V)]^{\mathrm{GL}(V)}=K[s_1,s_2,\dots,s_n].$$

Moreover, the s_i *are algebraically independent.*

It is a well known fact from linear algebra that $s_i(A)$ is the ith elementary symmetric function of the eigenvalues of A. In particular, if we choose a bases of V and identify $\mathrm{End}(V)$ with the $n\times n$-matrices $\mathrm{M}_n(K)$, the restrictions of the s_i's to the diagonal matrices $D\subset\mathrm{M}_n(K)$ are exactly the elementary symmetric functions σ_i on $D = K^n$. By the main theorem about symmetric functions (Proposition 1.2) they are algebraically independent and generate the algebra of symmetric functions. This already proves the second part of the theorem.

Assume that we know that every invariant function f on M_n is completely determined by its restriction to D (see Exercise 4 below). Then we can finish the proof in the following way: The restriction $f|_D$ of an invariant function f is clearly symmetric, hence of the form $f|_D = p(\sigma_1, \ldots, \sigma_n)$ with a polynomial p in n variables. But then $f - p(s_1, \ldots, s_n)$ is an invariant function on $\mathrm{M}_n(K)$ which vanishes on D, and so by the assumption above $f = p(s_1, \ldots, s_n) \in K[s_1, \ldots, s_n]$.

Following is another proof which does not make use of the theory of symmetric functions.

Proof of Proposition: Define

$$S := \{ \begin{pmatrix} 0 & & & & a_n \\ 1 & 0 & & & \\ & \ddots & \ddots & & \vdots \\ & & 1 & 0 & a_2 \\ & & & 1 & a_1 \end{pmatrix} \mid a_1, \ldots, a_n \in K\} \subset \mathrm{M}_n(K)$$

and let $X := \{A \in \mathrm{M}_n(K) \mid A \text{ is conjugate to a matrix in } S\}$. We claim that X is Zariski-dense in $\mathrm{M}_n(K)$. In fact, a matrix A belongs to X if and only if there is a vector $v \in K^n$ such that $v, Av, \ldots, A^{n-1}v$ are linearly independent. Now consider the polynomial function h on $\mathrm{M}_n(K) \times K^n$ given by

$$h(A, v) := \det(v, Av, A^2v, \ldots, A^{n-1}v).$$

It follows from Lemma 1.3 that the subset

$$\begin{aligned} Y &:= \{(A, v) \mid v, Av, \ldots, A^{n-1}v \text{ linearly independent}\} \\ &= (\mathrm{M}_n(K) \times K^n)_h \end{aligned}$$

is Zariski-dense in $\mathrm{M}_n(K) \times K^n$. Its projection onto $\mathrm{M}_n(K)$ is X which is therefore Zariski-dense, too (see 1.4 Exercise 19). This implies that every invariant function f on $\mathrm{M}_n(K)$ is completely determined by its restriction to S. An elementary calculation shows that for a matrix $A = A(a_1, \ldots, a_n)$ from the set S the characteristic polynomial is given by $P_A(t) = t^n - \sum_{i=1}^n a_i t^{n-i}$. Now $f(A) = q(a_1, \ldots, a_n)$ with a polynomial q in n variables, and we have $a_j = (-1)^{j+1} s_j(A)$. Hence, the function

$$f - q(s_1, -s_2, s_3, \ldots, (-1)^{n+1} s_n)$$

is invariant and vanishes on S, and so $f = q(s_1, -s_2, \dots) \in K[s_1, \dots, s_n]$. □

Exercise 4. The set of diagonalizable matrices is ZARISKI-dense in $\mathrm{M}_n(K)$. In particular, an invariant function on $\mathrm{M}_n(K)$ is completely determined by its restriction to the diagonal matrices.

(Hint: For an algebraically closed field K this is a consequence of the JORDAN decomposition. For the general case use 1.3 Exercise 13 and Remark 1.3.)

2.4 Traces of powers. There is another well-known series of invariant functions on $\mathrm{End}(V)$, namely the traces of the powers of an endomorphism:

$$\mathrm{Tr}_k \colon \mathrm{End}(V) \to K, \quad A \mapsto \mathrm{Tr}\, A^k, \quad k = 1, 2, \dots .$$

There are recursive formulas for expressing Tr_k in terms of the functions s_i:

$$\mathrm{Tr}_k = f_k(s_1, \dots, s_{k-1}) - (-1)^k k s_k \quad \text{for } k \leq n.$$

In fact, we have the same relations between the Tr_k's and the s_j's as those which hold for the power sums $n_k(x) := \sum_{i=1}^n x_i^k$ and the elementary symmetric functions σ_j (see the following Exercise 6). Hence, if $\mathrm{char}\, K > n$, the s_j can be expressed in terms of the Tr_k, $k = 1, 2, \dots, n$, and we get the following result:

Corollary. *If* $\mathrm{char}\, K = 0$ *then the functions* $\mathrm{Tr}_1, \mathrm{Tr}_2, \dots, \mathrm{Tr}_n$ *generate the invariant ring* $K[\mathrm{End}(V)]^{\mathrm{GL}(V)}$.

It is easy to see that the corollary does not hold if $0 < \mathrm{char}\, K \leq n$.

Exercises

5. Consider the polynomial

$$\psi(t) = \prod_{i=1}^{n} (1 - t x_i) = 1 - \sigma_1 t + \sigma_2 t^2 - \dots + (-1)^n \sigma_n t^n$$

where the σ_i are the elementary symmetric functions. Determine its logarithmic derivative $-\frac{\psi'(t)}{\psi(t)} = \sum_{i=1}^n \frac{x_i}{1 - t x_i}$ as a formal power series and deduce the NEWTON formulas:

$$(-1)^{j+1} j\, \sigma_j = n_j - \sigma_1 n_{j-1} + \sigma_2 n_{j-2} - \dots + (-1)^{j-1} \sigma_{j-1} n_1$$

for all $j = 1, \dots, n$.
(This approach is due to WEYL; see [Wey46] Chap. II.A.3. Another proof is suggested in 1.2 Exercise 8.)

6. Show that the same relations as above hold for the functions Tr_k and s_i:

$$(-1)^{j+1} j\, s_j = \mathrm{Tr}_j - s_1 \mathrm{Tr}_{j-1} + s_2 \mathrm{Tr}_{j-2} - \dots + (-1)^{j-1} s_{j-1} \mathrm{Tr}_1$$

for all $j = 1, \dots, n$.

Example (Pairs of 2×2 matrices). *Assume* $\operatorname{char} K \neq 2$. *The invariants of pairs of* $2{\times}2$ *matrices* $(A, B) \in \mathrm{M}_2(K){\times}\mathrm{M}_2(K)$ *under simultaneous conjugation are generated by the following functions:*

$$\operatorname{Tr} A, \quad \operatorname{Tr} A^2, \quad \operatorname{Tr} B, \quad , \operatorname{Tr} B^2, \quad , \operatorname{Tr} AB.$$

Moreover, these five invariants are algebraically independent.

Proof: The last statement is easy. Also, we can assume that K is algebraically closed (see 1.5 Remark 1.5). Moreover, it suffices to consider the traceless matrices ${\mathrm{M}_2}'$ since we have the direct sum decomposition $\mathrm{M}_2 = K \oplus {\mathrm{M}_2}'$.
There is a Zariski-dense set $U \subset {\mathrm{M}_2}' \times {\mathrm{M}_2}'$ where every pair $(A, B) \in U$ is equivalent to one of the form

$$(\begin{pmatrix} t & \\ & -t \end{pmatrix}, \begin{pmatrix} a & 1 \\ c & -a \end{pmatrix}), \quad t, c \neq 0.$$

In addition, such a pair is equivalent to the pair where t and a are replaced by $-t$ and $-a$. Thus, an invariant function f restricted to these pairs depends only on t^2, a^2, at and c. But

$$t^2 = \frac{1}{2} \operatorname{Tr} A^2, \quad a^2 + c = \frac{1}{2} \operatorname{Tr} B^2, \quad at = \frac{1}{2} \operatorname{Tr} AB$$

and so

$$c = \frac{1}{2} \operatorname{Tr} B^2 - \frac{(\operatorname{Tr} AB)^2}{\operatorname{Tr} A^2}.$$

It follows that f can be written as a rational function in $\operatorname{Tr} A^2, \operatorname{Tr} B^2$ and $\operatorname{Tr} AB$. Since f is a polynomial function on ${\mathrm{M}_2}' \times {\mathrm{M}_2}'$ and the given invariants are algebraically independent, it follows that f must be a polynomial function in these invariants. □

2.5 Invariants under simultaneous conjugation. Consider the linear action of $\mathrm{GL}(V)$ on $\operatorname{End}(V)^m := \operatorname{End}(V) \oplus \cdots \oplus \operatorname{End}(V)$ by simultaneous conjugation:

$$g(A_1, \dots, A_m) := (gA_1g^{-1}, gA_2g^{-1}, \dots, gA_mg^{-1}).$$

We want to describe the invariants under this action. For every finite sequence $i_1, i_2, \dots, i_k$ of numbers $1 \leq i_\nu \leq m$ we define a function

$$\operatorname{Tr}_{i_1 \dots i_k} \colon \operatorname{End}(V)^m \to K \qquad (A_1, \dots, A_m) \mapsto \operatorname{Tr}(A_{i_1} A_{i_2} \cdots A_{i_k}).$$

These *generalized traces* are clearly invariant functions.

First Fundamental Theorem for Matrices. *If* $\operatorname{char} K = 0$ *the ring of* *functions on* $\mathrm{End}(V)^m$ *which are invariant under simultaneous conjugation is generated by the invariants* $\mathrm{Tr}_{i_1 \ldots i_k}$*:*

$$K[\mathrm{End}(V)^m]^{\mathrm{GL}(V)} = K[\mathrm{Tr}_{i_1 \ldots i_k} \mid k \in \mathbb{N},\ 1 \leq i_1, \ldots, i_k \leq m].$$

The proof will be given in 4.7 after some preparation in §3 and §4. In fact, we will see that in characteristic 0 it is equivalent to the FFT for $\mathrm{GL}(V)$ (2.1). We have already remarked in 2.4 that the theorem does not hold if $\mathrm{char}\, K$ is positive and $\leq n$.

Remark. The theorem as stated gives an infinite set of generators. We will show in Chapter II ?? that the traces $\mathrm{Tr}_{i_1 \ldots i_k}$ of degree $k \leq n^2$ already generate the invariant ring. It is conjectured that $k \leq \binom{n+1}{2}$ suffices, but this is proved only for $\dim V \leq 3$ (see [For87]).

§ 3 Endomorphisms of Tensors

In this paragraph we study the m-fold tensor product $V^{\otimes m}$ as an $\mathcal{S}_m \times \mathrm{GL}(V)$-module. Although seemingly unrelated to our earlier considerations it will turn out that there is a strong connection with the First Fundamental Theorems discussed in the previous paragraph. In fact, the results here will provide a first proof of the FFTs valid in characteristic zero (see §4). Moreover, we obtain a beautiful correspondence between irreducible representations of the general linear group GL_n and irreducible representations of the symmetric group $\mathcal{S}_m$ which is due to Schur.

3.1 Centralizers and endomorphism rings. Let us consider the *m-fold tensor product*

$$V^{\otimes m} := \underbrace{V \otimes \cdots \otimes V}_{m \text{ times}}$$

of a (finite dimensional) vector space V and the usual linear action of $\mathrm{GL}(V)$ given by

$$g(v_1 \otimes \cdots \otimes v_m) := gv_1 \otimes \cdots \otimes gv_m.$$

The symmetric group $\mathcal{S}_m$ of m letters also acts on $V^{\otimes m}$:

$$\sigma(v_1 \otimes \cdots \otimes v_m) := v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(m)}.$$

It is obvious that these two actions commute. Let us denote by $\langle \mathrm{GL}(V) \rangle$ the linear subspace of $\mathrm{End}(V^{\otimes m})$ spanned by the linear operators coming from $\mathrm{GL}(V)$, i.e., by the image of $\mathrm{GL}(V)$ in $\mathrm{End}(V^{\otimes m})$ under the representation considered above. Similarly, we define the subspace $\langle \mathcal{S}_m \rangle \subset \mathrm{End}(V^{\otimes m})$. Both are subalgebras and they centralize each other:

$$ab = ba \text{ for all } a \in \langle \mathrm{GL}(V) \rangle \text{ and all } b \in \langle \mathcal{S}_m \rangle.$$

For any subalgebra $A \subset \mathrm{End}(W)$ where W is an arbitrary vector space the *centralizer* (or *commutant*) of A is the subalgebra consisting of those elements of $\mathrm{End}(W)$ which commute with all element of A. It will be denoted by A':

$$A' := \{b \in \mathrm{End}(W) \mid ab = ba \text{ for all } a \in A\}.$$

Equivalently, A' is the algebra of A-linear endomorphisms of W considered as an A-module:

$$A' = \mathrm{End}_A(W).$$

The next result claims that the two subalgebras $\langle \mathrm{GL}(V) \rangle$ and $\langle \mathcal{S}_m \rangle$ introduced above are the centralizers of each other:

Theorem. *Consider the usual linear actions of* $\mathrm{GL}(V)$ *and* $\mathcal{S}_m$ *on* $V^{\otimes m}$ *and denote by* $\langle \mathrm{GL}(V) \rangle$ *and* $\langle \mathcal{S}_m \rangle$ *the subalgebras of* $\mathrm{End}(V^{\otimes m})$ *spanned by the* *linear operators from* $\mathrm{GL}(V)$ *and* $\mathcal{S}_m$*, respectively. Then*

(a) $\mathrm{End}_{\mathcal{S}_m}(V^{\otimes m}) = \langle \mathrm{GL}(V) \rangle$.
(b) *If* $\mathrm{char}\, K = 0$ *then* $\mathrm{End}_{\mathrm{GL}(V)}(V^{\otimes m}) = \langle \mathcal{S}_m \rangle$.

The proof of the theorem is divided up into several steps. First we prove assertion (a). Then we show that in characteristic zero assertion (b) follows from (a) by the "Double Centralizer Theorem" (3.2). Finally, we will give a more general formulation of the result in (3.3) including a description of $V^{\otimes m}$ as an $\mathcal{S}_m \times \mathrm{GL}(V)$-module.

Exercise 1. Give a direct proof of the theorem in case $m = 2$.

Proof of (a): We use the natural isomorphism $\gamma\colon \mathrm{End}(V)^{\otimes m} \xrightarrow{\sim} \mathrm{End}(V^{\otimes m})$ given by $\gamma(A_1 \otimes \cdots \otimes A_m)(v_1 \otimes \cdots \otimes v_m) = A_1 v_1 \otimes \cdots \otimes A_m v_m$. Then the corresponding representation $\mathrm{GL}(V) \to \mathrm{End}(V)^{\otimes m}$ is $g \mapsto g \otimes \cdots \otimes g$. We claim that the corresponding action of $\mathcal{S}_m$ on $\mathrm{End}(V)^{\otimes m}$ is the obvious one:

$$\sigma(A_1 \otimes \cdots \otimes A_m) = A_{\sigma^{-1}(1)} \otimes \cdots \otimes A_{\sigma^{-1}(m)}.$$

In fact,

$$\begin{aligned} &\sigma(\gamma(A_1 \otimes \cdots \otimes A_m)(\sigma^{-1}(v_1 \otimes \cdots \otimes v_m))) \\ &\quad = \sigma(A_1 v_{\sigma(1)} \otimes \cdots \otimes A_m v_{\sigma(m)}) \\ &\quad = A_{\sigma^{-1}(1)} v_1 \otimes \cdots \otimes A_{\sigma^{-1}(m)} v_m \\ &\quad = \gamma(A_{\sigma^{-1}(1)} \otimes \cdots \otimes A_{\sigma^{-1}(m)})(v_1 \otimes \cdots \otimes v_m). \end{aligned}$$

This implies that γ induces an isomorphism between the *symmetric tensors* in $\mathrm{End}(V)^{\otimes m}$ and the subalgebra $\mathrm{End}_{\mathcal{S}_m}(V^{\otimes m})$ of $\mathrm{End}(V^{\otimes m})$. The claim now follows from the next lemma applied to $X := \mathrm{GL}(V) \subset W := \mathrm{End}(V)$. □

Lemma. *Let* W *be a finite dimensional vector space and* $X \subset W$ *a* Zariski*-dense subset. Then the linear span of the tensors* $x \otimes \cdots \otimes x$*,* $x \in X$*, is the subspace* $\Sigma_m \subset W^{\otimes m}$ *of all symmetric tensors.*

Recall that a subset $X \subset W$ is Zariski-dense if every function $f \in K[W]$ vanishing on X is the zero function (see 1.3).

Proof: Let $w_1 \ldots w_N$ be a basis of W. Then $B := \{w_{i_1} \otimes \cdots \otimes w_{i_m}\}_{i_1,\ldots,i_m}$ is a basis of $W^{\otimes m}$ which is stable under the action of $\mathcal{S}_m$. Two elements $w_{i_1} \otimes \cdots \otimes w_{i_m}$ and $w_{j_1} \otimes \cdots \otimes w_{j_m}$ belong to the same orbit under $\mathcal{S}_m$ if an only if each w_i appears the same number of times in both expressions. In particular, every orbit has a unique representative of the form $w_1^{\otimes h_1} \otimes w_2^{\otimes h_2} \otimes \cdots \otimes w_N^{\otimes h_N}$ where $h_1 + h_2 + \cdots + h_N = m$. Let us denote by $r_{h_1 \ldots h_N} \in W^{\otimes m}$ the sum of the elements in this orbit. Clearly, $\{r_{h_1\dots h_N} \mid \sum h_i = m\}$ is a basis of the symmetric tensors $\Sigma_m \subset W^{\otimes m}$. In order to prove the lemma we will show that every linear function $\lambda\colon \Sigma_m \to K$ which vanishes on all $x \otimes \cdots \otimes x$, $x \in X$, is the zero function.

Write $x = \sum_{i=1}^{N} x_i w_i$. Then

$$x \otimes \cdots \otimes x = \sum_{\sum h_i = m} x_1^{h_1} \cdots x_N^{h_N} r_{h_1\dots h_N}$$

and so

$$\lambda(x^{\otimes m}) = \sum a_{h_1\dots h_N} x_1^{h_1} \cdots x_N^{h_N}$$

where $a_{h_1\dots h_N} := \lambda(r_{h_1\dots h_N}) \in K$. This is a polynomial in $x_1, \dots, x_N$ which vanishes on X, by assumption. Hence, it is the zero polynomial and so all $a_{h_1\dots h_N}$ are zero, i.e. $\lambda = 0$. □

Exercises

2. Let $F_d := K[x_1, \dots, x_n]_d$ denote the vector space of homogeneous forms of degree d and assume that $\operatorname{char} K = 0$ or $> d$. Show that F_d is linearly spanned by the dth powers of the linear forms.
(Why is here the assumption about the characteristic of K necessary whereas the lemma above holds in any characteristic?)

3. Let $\rho\colon G \to \mathrm{GL}(V)$ be an irreducible representation and assume that $\mathrm{End}_G(V) = K$. Denote by V^n the G-module $V \oplus V \oplus \cdots \oplus V$, n-times, which we identify with $V \otimes K^n$.

(a) We have $\mathrm{End}_G(V^n) = \mathrm{M}_n(K)$ in a canonical way.

(b) Every G-submodule of V^n is of the form $V \otimes U$ with a subspace $U \subset K^n$.

(c) If $\mu\colon H \to \mathrm{GL}(W)$ is an irreducible representation of a group H then $V \otimes W$ is a simple $G \times H$-module.

(d) We have $\langle G\rangle = \mathrm{End}(V)$.
(Hint: $\langle G\rangle$ is a $G \times G$ submodule of $\mathrm{End}(V) = V^* \otimes V$.)

(e) For every field extension L/K the representation of G on $V_L := V \otimes_K L$ is irreducible.
(Hint: $\langle G\rangle_L = \mathrm{End}_L(V_L)$.)

3.2 Double Centralizer Theorem. If $\operatorname{char} K = 0$ the theorem of Maschke (see [Art91] Chap. 9, Corollary 4.9) tells us that the *group algebra* $K[\mathcal{S}_m]$ is *semisimple*, i.e., every representation of $\mathcal{S}_m$ is *completely reducible*. As a consequence, the homomorphic image $\langle \mathcal{S}_m\rangle$ of $K[\mathcal{S}_m]$ is a semisimple subalgebra of $\mathrm{End}(V^{\otimes m})$. In this situation we have the following general result (which holds for any field K):

Proposition. *Let* $A \subset \mathrm{End}(W)$ *be a semisimple subalgebra and* $A' := \{b \in \mathrm{End}(W) \mid ab = ba$ *for all* $a \in A\}$ *its centralizer. Then:*

(a) A' *is semisimple and* $(A')' = A$.

(b) W *has a unique decomposition* $W = W_1 \oplus \cdots \oplus W_r$ *into simple, non-isomorphic* $A \otimes A'$*-modules* W_i. *In addition, this is the isotypic decomposition as an* A*-module and as an* A'*-module.*

(c) *Each simple factor* W_i *is of the form* $U_i \otimes_{D_i} U_i'$ *where* U_i *is a simple* A*-module,* U_i' *a simple* A'*-module, and* D_i *is the division algebra* $\mathrm{End}_A(U_i)^{\mathrm{op}} = \mathrm{End}_{A'}(U_i')^{\mathrm{op}}$.

Recall that for a G-module W and a simple module U the *isotypic component* of W of type U is the sum of all submodules of W isomorphic to U. The isotypic components form a direct sum which is all of W if and only if W is semisimple. In that case it is called the *isotypic decomposition.*

Proof: Let $W = W_1 \oplus \cdots \oplus W_r$ be the isotypic decomposition of W as an A-module, $W_i \xrightarrow{\sim} U_i^{s_i}$ with simple A-modules U_i which are pairwise non-isomorphic. Corresponding to this the algebra A decomposes in the form

$$A = \prod_{i=1}^{r} A_i, \quad A_i \xrightarrow{\sim} \mathrm{M}_{n_i}(D_i)$$

with a division algebra $D_i \supset K$. Furthermore, $U_i \stackrel{\sim}{=} D_i^{n_i}$ as an A-module where the module structure on $D_i^{n_i}$ is given by $A \xrightarrow{\mathrm{pr}} A_i \xrightarrow{\sim} \mathrm{M}_{n_i}(D_i)$. It follows that

$$\begin{aligned} A' &:= \mathrm{End}_A(W) = \prod \mathrm{End}_A(W_i), \qquad \text{and} \\ A_i' &:= \mathrm{End}_A(W_i) = \mathrm{End}_{A_i}(W_i) \stackrel{\sim}{=} \mathrm{M}_{s_i}(D_i') \end{aligned}$$

where $D_i' := \mathrm{End}_{A_i}(U_i) = D_i^{\mathrm{op}}$. In particular, A' is semisimple and

$$\begin{aligned} \dim A_i \dim A_i' &= n_i^2 \dim D_i \; s_i^2 \dim D_i' = (n_i s_i \dim D_i)^2 \\ &= (\dim W_i)^2 = \dim \mathrm{End}(W_i). \qquad (*) \end{aligned}$$

This implies that the canonical algebra homomorphism $A_i \otimes A_i' \xrightarrow{\sim} \mathrm{End}(W_i)$ is an isomorphism. (Recall that $A_i \otimes A_i'$ is simple.) In particular, W_i is a simple $A \otimes A'$-module. More precisely, consider $U_i = D_i^{n_i}$ as a right D_i-module and $U_i' := (D_i^{\mathrm{op}})^{s_i}$ as a left D_i-module. Then these structures commute with the A- resp. A'-module structure, hence $U_i \otimes_{D_i} U_i'$ is a $A \otimes A'$-module and we have an isomorphism $U_i \otimes_{D_i} U_i' \xrightarrow{\sim} U_i^{s_i} \xrightarrow{\sim} W_i$.

It remains to show that $(A')' = A$. For this we apply the same reasoning as above to A' and see from $(*)$ that $\dim A_i' \dim A_i'' = \dim \mathrm{End}(W_i) =$

$\dim A_i \dim A_i'$. Hence $\dim A'' = \dim A$ and since $A'' \supset A$ we are done. □

Exercise 4. Let V be an irreducible finite dimensional representation of a group G where $\mathrm{End}_G(V) = K$, and let W be an arbitrary finite dimensional representation of G.

(a) The linear map $\gamma\colon \mathrm{Hom}_G(V, W) \otimes V \to W$, $\alpha \otimes v \mapsto \alpha(v)$ is injective and G-equivariant and its image is the *isotypic submodule* of W of type V (i.e., the sum of all simple submodules of W isomorphic to V).

(b) If there is another group H acting on W and commuting with G then γ is H-equivariant.

(c) Assume that K is algebraically closed. Then every simple $G \times H$-module is of the form $V \otimes U$ with a simple G-module V and a simple H-module U.

3.3 Decomposition of $V^{\otimes m}$. Now we are ready to finish the proof of Theorem 3.1. This is contained in the following result which gives a more precise description of the $\mathcal{S}_m \times \mathrm{GL}(V)$-module structure of $V^{\otimes m}$. In addition, we obtain a beautiful correspondence between irreducible representations of the general linear group $\mathrm{GL}(V)$ and of the symmetric group $\mathcal{S}_m$ which was dicovered by SCHUR in his dissertation (Berlin, 1901).

Decomposition Theorem. *Assume* $\operatorname{char} K = 0$.

(a) *The two subalgebras* $\langle \mathcal{S}_m \rangle$ *and* $\langle \mathrm{GL}(V) \rangle$ *are both semisimple and are the centralizers of each other.*

(b) *There is a canonical decomposition of* $V^{\otimes m}$ *as an* $\mathcal{S}_m \times \mathrm{GL}(V)$*-module into simple non-isomorphic modules* V_λ*:*

$$V^{\otimes m} = \bigoplus_\lambda V_\lambda.$$

(c) *Each simple factor* V_λ *is of the form* $\mathrm{M}_\lambda \otimes L_\lambda$ *where* M_λ *is a simple* $\mathcal{S}_m$*-module and* L_λ *a simple* $\mathrm{GL}(V)$*-module. The modules* M_λ *(resp.* L_λ*) are all non-isomorphic.*

PROOF: In 3.1 we have already proved that $\langle \mathcal{S}_m \rangle' = \langle \mathrm{GL}(V) \rangle$, and so (a) and (b) follow from the Double Centralizer Theorem 3.2. For the last statement it remains to show that the endomorphism ring of every simple $\mathcal{S}_m$-modules M_λ is the base field K. This is clear if K is algebraically closed. For arbitrary K of characteristic zero it will be proved in 5.7 Corollary 2. □

One easily shows that every irreducible representation of $\mathcal{S}_m$ occurs in $V^{\otimes m}$ provided we have $\dim V \geq m$ (see Exercise 7 below). Let us fix such a representation M_λ. Then we obtain in a natural way

$$L_\lambda = L_\lambda(V) = \mathrm{Hom}_{\mathcal{S}_m}(M_\lambda, V^{\otimes m})$$

as a consequence of the fact that $\mathrm{End}_{\mathcal{S}_m}(M_\lambda) = K$ (Lemma of Schur; cf. Exercise 9 and Remark below). This shows that $L_\lambda(V)$ depends *functorially* on V. (This simply means that a linear map $\varphi\colon V \to W$ determines a linear map $L_\lambda(\varphi)\colon L_\lambda(V) \to L_\lambda(W)$, with $L_\lambda(\varphi \circ \psi) = L_\lambda(\varphi) \circ L_\lambda(\psi)$ and $L_\lambda(\mathrm{id}_V) = \mathrm{id}_{L_\lambda(V)}$.) As an example, if M_0 denotes the trivial representation and M_{sgn} the signum representation of $\mathcal{S}_m$ then we find the classical functors

$$L_0(V) = \mathrm{S}^m(V) \quad \text{and} \quad L_{\mathrm{sgn}}(V) = \textstyle\bigwedge^m V$$

(see Exercise 6). The *functor* L_λ is usually called Schur *functor or* Weyl *module*. We will discuss this again in §5 (see 5.9 Remark 2).

Exercises

5. Give a direct proof of the theorem above in case $m = 2$.

6. Show that the isotypic component of $V^{\otimes m}$ of the trivial representation of $\mathcal{S}_m$ is the symmetric power $\mathrm{S}^m V$ and the one of the signum representation is the exterior power $\bigwedge^m V$.

7. If $\dim V \geq m$ then every irreducible representation of $\mathcal{S}_m$ occurs in $V^{\otimes m}$. (In fact, the regular representation of $\mathcal{S}_m$ occurs as a subrepresentation.)

Remark. The fact that the endomorphism rings of the simple modules L_λ and M_λ are the base field K implies that these representations are "defined over $\mathbb{Q}$," i.e., $M_\lambda = M_\lambda^\circ \otimes_{\mathbb{Q}} K$ and $L_\lambda = L_\lambda^\circ \otimes_{\mathbb{Q}} K$, where M_λ° is a simple $\mathbb{Q}[\mathcal{S}_m]$-module L_λ° a simple $\mathrm{GL}(\mathbb{Q})$-module .

In fact, by the Theorem above we have for $K = \mathbb{Q}$ a decomposition $V^{\otimes m} = \bigoplus_\lambda M_\lambda^\circ \otimes L_\lambda^\circ$ and so $V_K^{\otimes m} = \bigoplus_\lambda (M_\lambda^\circ \otimes K) \otimes_K (L_\lambda^\circ \otimes K)$ which shows that $M_\lambda = M_\lambda^\circ \otimes_{\mathbb{Q}} K$ and $L_\lambda = L_\lambda^\circ \otimes_{\mathbb{Q}} K$.

Exercises

8. Let $\rho\colon G \to \mathrm{GL}(V)$ be a completely reducible representation. For any field extension K'/K the representation of G on $V \otimes_K K'$ is completely reducible, too.
(Hint: The subalgebra $A := \langle \rho(G) \rangle \subset \mathrm{End}(V)$ is semisimple. This implies that $A \otimes_K K'$ is semisimple, too.)

9. Let V be an irreducible finite dimensional K-representation of a group G. Then $\mathrm{End}_G(V) = K$ if and only if $V \otimes_K K'$ is irreducible for every field extension K'/K.
(Hint: Show that $\mathrm{End}_G(V \otimes_K K') = \mathrm{End}_G(V) \otimes_K K'$.)

The following corollary is clear. For the second statement one simply remarks that the scalars $t \in K^* \subset \mathrm{GL}(V)$ act by $t^m \cdot \mathrm{id}$ on $V^{\otimes m}$.

Corollary. *The representation* $V^{\otimes m}$ *of* $\mathrm{GL}(V)$ *is completely reducible. For* $m \neq m'$, $V^{\otimes m}$ *and* $V^{\otimes m'}$ *do not contain isomorphic submodules.*

Exercises

10. There is a canonical isomorphism $\varphi\colon \operatorname{End}(V) \xrightarrow{\sim} \operatorname{End}(V)^*$. It is induced by the bilinear form $(A, B) \mapsto \operatorname{Tr}(AB)$ and is $\mathrm{GL}(V)$-equivariant. Moreover, $\varphi^* = \varphi$.

11. There is a natural $\mathrm{GL}(V)$-equivariant isomorphism $V \otimes V^* \xrightarrow{\sim} \operatorname{End}(V)$. Which element of $V \otimes V^*$ corresponds to $\mathrm{id} \in \operatorname{End}(V)$ and which elements of $\operatorname{End}(V)$ correspond to the "pure" tensors $v \otimes \lambda$?

§ 4 Polarization and Restitution

In this paragraph we study the *multilinear* invariants of vectors and covectors and of matrices. We prove the multilinear versions of the corresponding First Fundamental Theorems from §2 whose proofs have been set aside. In fact, we show that these multilinear versions are both equivalent to our result from the previous paragraph claiming that the $\mathrm{GL}(V)$-equivariant endomorphisms of $V^{\otimes m}$ are linearly generated by the permutations (Theorem 3.1b).

Then, using *polarization* and *restitution* we will be able to reduce (in characteristic zero) the general versions of the First Fundamental Theorems to the multilinear case. Thus, we obtain a first proof of the FFT's. A completely different proof based on the theory of CAPELLI will be presented in §8.

4.1 Multihomogeneous invariants. Consider a direct sum $V = V_1 \oplus \cdots \oplus V_r$ of finite dimensional vector spaces. A function $f \in K[V_1 \oplus \cdots \oplus V_r]$ is called *multihomogeneous* of degree $h = (h_1, \ldots, h_r)$ if f is homogeneous of degree h_i in V_i, i.e., for all $v_1, \ldots, v_r \in V$, $t_1, \ldots, t_r \in K$ we have

$$f(t_1v_1, t_2v_2, \ldots, t_rv_r) = t_1^{h_1} \cdots t_r^{h_r} f(v_1, \ldots, v_r).$$

Every polynomial function f is in a unique way a sum of multihomogeneous functions: $f = \sum f_h$; the f_h are usually called the *multihomogeneous components* of f. This gives rise to a decomposition

$$K[V_1 \oplus \cdots \oplus V_r] = \bigoplus_{h \in \mathbb{N}^r} K[V_1 \oplus \cdots \oplus V_r]_h,$$

where $K[V_1 \oplus \cdots \oplus V_r]_h$ is the subspace of all multihomogeneous functions of degree h. We remark that this is a *graduation* of the algebra in the sense that

$$K[V_1 \oplus \cdots \oplus V_r]_h \cdot K[V_1 \oplus \cdots \oplus V_r]_k = K[V_1 \oplus \cdots \oplus V_r]_{h+k}.$$

It is also clear that each $K[V_1 \oplus \cdots \oplus V_r]_h$ is stable under the action of $\mathrm{GL}(V_1) \times \cdots \times \mathrm{GL}(V_r)$. In particular, if the V_i are representations of a group G we find a corresponding decomposition for the ring of invariants:

$$K[V_1 \oplus \cdots \oplus V_r]^G = \bigoplus_{h \in \mathbb{N}^r} K[V_1 \oplus \cdots \oplus V_r]_h^G.$$

In other words, *if f is an invariant then every multihomogeneous component is also an invariant.* This will enable us to reduce many questions about invariants to the multihomogeneous case.

> **Exercise 1.** Let W, U be two G-modules where G is an arbitrary group. Show that the homogeneous covariants $\varphi\colon W \to U$ of degree d are in one-to-one correspondence with the bihomogenous invariants of $W \oplus U^*$ of degree $(d, 1)$.

4.2 Multilinear invariants of vectors and covectors. Consider again the natural action of $\mathrm{GL}(V)$ on $V^p \oplus V^{*q}$ as in §2 and let $f\colon V^p \oplus V^{*q} \to K$ be a *multilinear* invariant. If $f \neq 0$, then we must have $p = q$. In fact, if we apply a scalar $\lambda \in K^* \subset \mathrm{GL}(V)$ to $(v, \varphi) = (v_1, \dots, v_p, \varphi_1, \dots, \varphi_q)$ we obtain $(\lambda v_1, \dots, \lambda v_p, \lambda^{-1}\varphi_1, \dots, \lambda^{-1}\varphi_q)$, hence $f(\lambda(v,\varphi)) = \lambda^{p-q} f((v,\varphi))$.

Now the FFT for $\mathrm{GL}(V)$ (2.1) claims that the invariants are generated by the contractions $(i \mid j)$ defined by $(i \mid j)(v,\varphi) = \varphi_j(v_i)$. Therefore, a multilinear invariant of $V^p \oplus V^{*p}$ is a linear combination of products of the form

$$(1 \mid i_1)(2 \mid i_2) \cdots (p \mid i_p)$$

where $(i_1, i_2, \dots, i_p)$ is a permutation of $(1, 2, \dots, p)$. Thus the following is a special case of the FFT for $\mathrm{GL}(V)$.

Theorem (Multilinear FFT for $\mathrm{GL}(V)$). *Assume* $\operatorname{char} K = 0$. *Multilinear invariants of $V^p \oplus V^{*q}$ exist only for $p = q$. They are linearly generated by the functions*

$$f_\sigma := (1 \mid \sigma(1)) \cdots (p \mid \sigma(p)), \quad \sigma \in \mathcal{S}_p\,.$$

This theorem holds in arbitrary characteristic by the fundamental work of De Concini and Procesi [DeP76], but our proof will only work in characteristic zero since it is based on the double centralizer theorem.

Proof: We will prove the following more general statement:

Claim. *The theorem above is equivalent to Theorem 3.1(b) stating that*

$$\operatorname{End}_{\mathrm{GL}(V)} V^{\otimes m} = \langle \mathcal{S}_m \rangle.$$

Let us denote by M the multilinear functions on $V^m \oplus V^{*m}$. Then we have in a canonical way

$$M = (V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*)^* = (W \otimes W^*)^*$$

where $W := V^{\otimes m}$. Now there is a canonical isomorphism

$$\alpha\colon \operatorname{End}(W) \xrightarrow{\sim} (W \otimes W^*)^*$$

given by $\alpha(A)(w \otimes \psi) = \psi(Aw)$ which is clearly $\mathrm{GL}(W)$-equivariant. Hence, we get a $\mathrm{GL}(V)$-equivariant isomorphism $\operatorname{End}(V^{\otimes m}) \xrightarrow{\sim} (V^{\otimes m} \otimes V^{*\otimes m})^* = M$ which induces an isomorphism

$$\operatorname{End}_{\mathrm{GL}(V)}(V^{\otimes m}) \xrightarrow{\sim} M^{\mathrm{GL}(V)} = \{\text{invariant multilinear functions}\}.$$

Let us calculate the image of $\sigma \in \operatorname{End}(V^{\otimes m})$ under α:

$$\begin{aligned}
&\alpha(\sigma)(v_1 \otimes \cdots \otimes v_m \otimes \varphi_1 \otimes \cdots \otimes \varphi_m) \\
&\quad = (\varphi_1 \otimes \cdots \otimes \varphi_m)(\sigma(v_1 \otimes \cdots \otimes v_m)) \\
&\quad = (\varphi_1 \otimes \cdots \otimes \varphi_m)(v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(m)}) \\
&\quad = (\varphi_1 \mid v_{\sigma^{-1}(1)})(\varphi_2 \mid v_{\sigma^{-1}(2)}) \cdots (\varphi_m \mid v_{\sigma^{-1}(m)}) \\
&\quad = f_{\sigma^{-1}}(v_1 \otimes \cdots \otimes v_m \otimes \varphi_1 \otimes \cdots \otimes \varphi_m).
\end{aligned}$$

Thus, $\alpha\langle \mathcal{S}_m \rangle = \langle f_\sigma \mid \sigma \in \mathcal{S}_p \rangle$ and the claim follows. □

4.3 Multilinear invariants of matrices. Now we look for multilinear invariants on m copies of $\operatorname{End}(V)$. Let $\sigma \in \mathcal{S}_m$ and write σ as a product of *disjoint cycles* (including all cycles of length one):

$$\sigma = (i_1, \ldots, i_k)(j_1, \ldots, j_r) \cdots (l_1, \ldots, l_s).$$

Define a function $\operatorname{Tr}_\sigma \colon \operatorname{End}(V)^m \to K$ by

$$\begin{aligned}
&\operatorname{Tr}_\sigma(A_1, \ldots, A_m) := \\
&\qquad \operatorname{Tr}(A_{i_1} \cdots A_{i_k}) \operatorname{Tr}(A_{j_1} \cdots A_{j_r}) \cdots \operatorname{Tr}(A_{l_1} \cdots A_{l_s}).
\end{aligned}$$

Clearly, Tr_σ is a multilinear invariant. It is easy to see that it does not depend on the presentation of σ as a product of disjoint cycles (cf. Exercise 2 below). It is now obvious that the following theorem is a special case of the FFT for matrices (2.5).

Theorem (Multilinear FFT for matrices). *Assume* $\operatorname{char} K = 0$. *The multilinear invariants on* $\operatorname{End}(V)^m$ *are linearly generated by the functions* Tr_σ, $\sigma \in \mathcal{S}_m$.

Again this holds in arbitrary characteristic by the fundamental work of De Concini and Procesi [DeP76], but our proof only works in characteristic zero.

Proof: This result follows from the multilinear FFT for $\operatorname{GL}(V)$ which we have just proved. In fact, we have again the following more precise statement:

Claim. *The theorem above is equivalent to Theorem 4.2, the multilinear version of the FFT for* $\operatorname{GL}(V)$.

The multilinear functions on $\operatorname{End}(V)^m$ can be identified with $(\operatorname{End}(V)^{\otimes m})^*$. This time we want to use the following canonical isomorphism. (Recall that $\operatorname{End}(V) \xrightarrow{\sim} \operatorname{End}(V)^*$ in a canonical way, see 3.3 Exercise 10.)

$$\beta \colon V \otimes V^* \xrightarrow{\sim} \operatorname{End} V, \qquad \beta(v \otimes \varphi)(u) = \varphi(u)v.$$

In other words, $\beta(v \otimes \varphi)$ is a rank one linear endomorphism of V with image Kv and kernel $\ker \varphi$. The following two statements are easily verified:

(a) $\mathrm{Tr}(\beta(v \otimes \varphi)) = \varphi(v)$,
(b) $\beta(v \otimes \varphi) \circ \beta(w \otimes \psi) = \beta(v \otimes \varphi(w)\psi)$.

Now β induces a $\mathrm{GL}(V)$-equivariant isomorphism

$$\widetilde{\beta}\colon V^{\otimes m} \otimes V^{*\otimes m} \xrightarrow{\sim} \mathrm{End}(V)^{\otimes m}.$$

Hence, the dual map $\widetilde{\beta}^*$ identifies the multilinear invariants of $\mathrm{End}(V)^m$ with those of $V^m \otimes V^{*m}$. The claim follows once we have shown that $\widetilde{\beta}^*(\mathrm{Tr}_\sigma) = f_\sigma$. Let $\sigma = (i_1, \dots, i_k)(j_1, \dots, j_r) \cdots (l_1, \dots, l_s)$ be the decomposition into disjoint cycles. Then, using (b) and (a) above we find

$$\begin{aligned}
\mathrm{Tr}_\sigma \widetilde{\beta}(v_1 \otimes \cdots &\otimes v_m \otimes \varphi_1 \otimes \cdots \otimes \varphi_m) \\
&= \mathrm{Tr}_\sigma(\beta(v_1 \otimes \varphi_1)\beta(v_2 \otimes \varphi_2) \cdots) \\
&= \mathrm{Tr}(\beta(v_{i_1} \otimes \varphi_{i_1})\beta(v_{i_2} \otimes \varphi_{i_2}) \cdots \beta(v_{i_k} \otimes \varphi_{i_k})) \cdots \\
&= \mathrm{Tr}(\beta(v_{i_1} \otimes \varphi_{i_1}(v_{i_2})\varphi_{i_2}(v_{i_3}) \cdots \varphi_{i_{k-1}}(v_{i_k})\varphi_{i_k}) \cdots \\
&= \varphi_{i_1}(v_{i_2})\varphi_{i_2}(v_{i_3}) \cdots \varphi_{i_k}(v_{i_1}) \cdots
\end{aligned}$$

But $\sigma(i_\nu) = i_{\nu+1}$ for $\nu < k$ and $\sigma(i_k) = i_1$, etc., and so the last product equals

$$\prod_{i=1}^{m} \varphi_i(v_{\sigma(i)}) = f_\sigma(v_1 \otimes \cdots \otimes v_m \otimes \varphi_1 \otimes \cdots \otimes \varphi_m)$$

which proves the claim. □

Remark. Although our proofs of the two theorems 4.2 and 4.3 above are only valid in characteristic zero we have shown more generally that, independently of the characteristic of the field K, they are both equivalent to the statement that $\mathrm{End}_{\mathrm{GL}(V)} V^{\otimes m} = \langle \mathcal{S}_m \rangle$.

Exercises

2. Show that $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB)$ for all $A, B, C \in \mathrm{End}(V)$.

3. Let R be a finite dimensional algebra. The multiplication of R corresponds to a tensor $\mu \in R^* \otimes R^* \otimes R$ and the automorphism group of R is equal to the stabilizer of μ in $\mathrm{GL}(R)$.

4. Let V, W be two finite dimensional vector spaces and let $\gamma\colon V^* \otimes W \xrightarrow{\sim} \mathrm{Hom}(V, W)$ be the canonical isomorphism. Then a tensor t corresponds to a homomorphism of rank $\leq r$ if and only if t can be written as a sum of at most r pure tensors $\lambda_i \otimes w_i$.
(Cf. 3.3 Exercise 11)

4.4 Polarization. Let $f \in K[V]$ be a homogeneous function of degree d. Calculating f on a vector of the form $v = \sum_{i=1}^{d} t_i v_i$, $t_i \in K$, $v_i \in V$, we obtain

$$f(t_1 v_1 + \cdots + t_d v_d) = \sum_{s_1+\cdots+s_d=d} t_1^{s_1} \cdots t_d^{s_d} f_{s_1 \ldots s_d}(v_1, \ldots, v_d) \qquad (*)$$

where the polynomials $f_{s_1 \ldots s_d} \in K[V^d]$ are well defined and are multihomogeneous of degree $(s_1, \ldots, s_d)$.

Definition. The multilinear polynomial $f_{1\,1 \ldots 1} \in K[V^d]$ is called the *(full) polarization* of f. It will be denoted by $\mathcal{P}f$.

Lemma. *The linear operator* $\mathcal{P}$: $K[V]_d \to K[V^d]_{(1,1,\ldots,1)}$ *has the following properties:*

(a) $\mathcal{P}$ *is* $\mathrm{GL}(V)$*-equivariant;*
(b) $\mathcal{P}f$ *is symmetric;*
(c) $\mathcal{P}f(v, v, \ldots, v) = d!\, f(v)$.

Proof: The first two statements (a) and (b) are easily verified. For (c) we replace in $(*)$ every v_i by v and obtain on the left hand side:

$$f((\sum_i t_i)v) = (\sum_i t_i)^d f(v) = (t_1^d + \cdots + d!\, t_1 t_2 \cdots t_d) f(v),$$

and the claim follows. □

4.5 Restitution. Next we define the inverse operator to polarization.

Definition. For a multilinear $F \in K[V^d]$ the homogeneous polynomial $\mathcal{R}\, F$ defined by $\mathcal{R}\, F(v) := F(v, v, \ldots, v)$ is called the *(full) restitution* of F.

Again, $\mathcal{R}$: $K[V^d]_{(1,\ldots,1)} \to K[V]_d$ is a linear $\mathrm{GL}(V)$-equivariant operator and we have $\mathcal{R}\,\mathcal{P}f = d!\, f$ by the property (c) of the lemma above. As a consequence, we get the following result.

Proposition. *Assume* $\operatorname{char} K = 0$ *and let* V *be a finite dimensional representation of a group* G*. Then every homogeneous invariant* $f \in K[V]^G$ *of degree* d *is the full restitution of a multilinear invariant* $F \in K[V^d]^G$*.*

In fact, f is the full restitution of $\frac{1}{d!}\mathcal{P}f$, which is a multihomogeneous invariant by Lemma 4.4 above.

Exercises

5. Show that $\mathcal{P}\mathcal{R}\, F$ is the *symmetrization* of F defined by

$$\operatorname{sym} F(v_1, \ldots, v_d) := \sum_{\sigma \in \mathcal{S}_d} F(v_{\sigma(1)}, \ldots, v_{\sigma(d)}).$$

6. Let $f \in K[V]$ be homogeneous of degree d and write

$$f(sv + tw) = \sum_{i=0}^{d} s^i t^{d-i} f_i(v, w), \quad s, t \in K,\ v, w \in V.$$

Then the polynomials f_i are bihomogeneous of degree $(i, d-i)$ and the linear operators $f \mapsto f_i$ are $\mathrm{GL}(V)$-equivariant. Moreover, $f_i(v, v) = \binom{d}{i} f(v)$. In particular, if G is any subgroup of $\mathrm{GL}(V)$ then f is an invariant under G if and only if all f_i are G-invariant.
The f_i's are sometimes called *partial polarizations* of f.

4.6 Generalization to several representations. For some applications we have in mind we need a slight generalization of the results above. The proofs are obvious and left to the reader.

Let $f \in K[V_1 \oplus \cdots \oplus V_r]$ be a multihomogeneous polynomial of degree $d = (d_1, \ldots, d_r)$. Denote by $\mathcal{P}_i$ the linear operator "full polarization with respect to the variable $v_i \in V_i$." Then

$$\mathcal{P}f := \mathcal{P}_r \mathcal{P}_{r-1} \cdots \mathcal{P}_1 f \in K[V_1^{d_1} \oplus V_2^{d_2} \oplus \cdots \oplus V_r^{d_r}]$$

is a multilinear polynomial which we call again the *polarization* of f. Similarly, we define the *restitution* $\mathcal{R}F$ of a multilinear $F \in K[V_1^{d_1} \oplus \cdots \oplus V_r^{d_r}]$ by

$$\mathcal{R}F(v_1, \ldots, v_d) := F(\underbrace{v_1, \ldots, v_1}_{d_1}, \underbrace{v_2, \ldots, v_2}_{d_2}, \ldots, \underbrace{v_r, \ldots, v_r}_{d_r}).$$

Lemma. *The two linear operators*

$$\mathcal{P}\colon K[V_1 \oplus \cdots \oplus V_r]_{(d_1, \ldots, v_r)} \to K[V_1^{d_1} \oplus \cdots \oplus V_r^{d_r}]_{\text{multlin}}$$

$$\mathcal{R}\colon K[V_1^{d_1} \oplus \cdots \oplus V_r^{d_r}]_{\text{multlin}} \to K[V_1 \oplus \cdots \oplus V_r]_{(d_1, \ldots, d_r)}$$

are $\mathrm{GL}(V_1) \times \cdots \times \mathrm{GL}(V_r)$*-equivariant and satisfy the following conditions:*

(a) $\mathcal{P}f$ *is multisymmetric* (i.e., symmetric with respect to the variables in V_i for each i)*;*

(b) $\mathcal{R}\mathcal{P}f = d_1!\, d_2! \cdots d_r!\, f$*;*

(c) $\mathcal{P}\mathcal{R}F$ *is the multisymmetrization of* F*:*

$$\mathcal{P}\mathcal{R}F(v_1^{(1)}, \ldots, v_1^{(d_1)}, \ldots) = \\ = \sum_{(\sigma_1, \ldots, \sigma_r) \in \mathcal{S}_{d_1} \times \cdots \times \mathcal{S}_{d_r}} F(v_1^{(\sigma_1(1))}, \ldots, v_1^{(\sigma_1(d_1))}, \ldots).$$

As before we obtain the following important consequence about multihomogeneous invariants:

Proposition. *Assume* $\operatorname{char} K = 0$ *and let* $V_1, V_2, \dots, V_r$ *be representations of a group* G. *Then every multihomogeneous invariant* $f \in K[V_1 \oplus \cdots \oplus V_r]^G$ *of degree* $d = (d_1, \dots, d_r)$ *is the restitution of a multilinear invariant* $F \in K[V_1^{d_1} \oplus \cdots \oplus V_r^{d_r}]^G$.

4.7 Proof of the First Fundamental Theorems. Now we are ready to complete the proof for the FFTs. Our Theorem 4.2 states that every multilinear invariant F of $V^p \oplus V^{*p}$ is a linear combination of the invariants

$$f_\sigma = (1 \mid \sigma(1)) \cdots (p \mid \sigma(p)),\ \sigma \in \mathcal{S}_p \,.$$

Any (partial) restitution of such an f_σ is a monomial in the $(i \mid j)$. By Proposition 4.5 above this shows that every invariant of $V^p \oplus V^{*q}$ is a polynomial in the $(i \mid j)$. Or,

> *If* $\operatorname{char} K = 0$ *the FFT 2.1 for vectors and covectors is a consequence of its multilinear version stated in Theorem 4.2.*

In case of matrices a similar argument can be used: The restitution of the invariant Tr_σ is a product of functions of the form $\operatorname{Tr}(i_1, \dots, i_k)$. As above, this implies the following result:

> *If* $\operatorname{char} K = 0$ *the FFT 2.5 for matrices is a consequence of its multilinear version stated in Theorem 4.3.*

Thus, we have completed a first proof of the two versions of the First Fundamental Theorem. The argument is rather indirect and only valid in characteristic zero. We will give a completely different approach later in §8 which is based on the Capelli-Deruyts expansion. There it will follow from the FFT for SL_n (see Remark 8.5).

4.8 Example: Invariants of forms and vectors. Let F_d denote the space of homogeneous forms of degree d on the vector space V: $F_d := K[V]_d$.

Question. *What are the invariants of* $F_d \oplus V$ *under* $\mathrm{GL}(V)$?

There is an obvious invariant ε given by "evaluation": $\varepsilon(f, v) := f(v)$. We claim that this is a generator of the invariant ring:

$$K[F_d \oplus V]^{\mathrm{GL}(V)} = K[\varepsilon].$$

Proof: Let $h\colon F_d \oplus V \to K$ be a bihomogeneous invariant of degree (r, s) and let $\tilde{h}$ be the full polarization of h with respect to the first variable. Then

$$\tilde{h}\colon F_d^r \oplus V \to K$$

is a multihomogeneous invariant of degree $(1, 1, \ldots, 1, s)$. Composing $\tilde{h}$ with the dth power map $\zeta \mapsto \zeta^d \colon V^* \to F_d$ we obtain a multi-homogeneous invariant

$$H\colon V^{*r} \oplus V \to K$$

of degree $(d, d, \ldots, d, s)$. Now it follows from the FFT for $\mathrm{GL}(V)$ (2.1) that $rd = s$ and that H is a scalar multiple of the invariant $(1 \mid 1)^d (2 \mid 1)^d \cdots (r \mid 1)^d$.

On the other hand, starting with $h = \varepsilon$ we find $\tilde{h} = \varepsilon$ and $H(\zeta, v) = \zeta(v)^d$, hence $H = (1 \mid 1)^d$. Since F_d is linearly spanned by the dth powers of linear forms we see that $\tilde{h}$ is completely determined by H and therefore h is a scalar multiple of ε^r. □

Exercises

7. Show that the invariants of a form f of degree d and two vectors v, w are generated by the following invariants ε_i, $i = 0, 1, \ldots, d$:

$$K[F_d \oplus V^2]^{\mathrm{GL}(V)} = K[\varepsilon_0, \ldots, \varepsilon_d], \quad \varepsilon_i(f, v, w) := f_i(v, w)$$

where the f_i are the partial polarizations of f as defined in Exercise 6.

8. The homogeneous covariants $\mathrm{M}_n(K) \to \mathrm{M}_n(K)$ of degree d are linearly spanned by the maps $A \mapsto h_i(A)A^{d-i}$, $i = 0, 1, \ldots, d$, where h_i is a homogeneous invariant of $\mathrm{M}_n(K)$ of degree i. It follows that the covariants $\mathrm{M}_n(K) \to \mathrm{M}_n(K)$ form a free module over the invariants with basis $\mu_i \colon A \mapsto A^i$, $i = 0, 1, \ldots, n-1$.
(Hint: Use Exercise 1.)

§ 5 Representation Theory of GL_n

In this paragraph we show that in characteristic zero every rational representation of GL_n is completely reducible. This will be deduced from the general fact that every polynomial representation of GL_n occurs as a subrepresentation of a direct sum of suitable tensor powers $(K^n)^{\otimes m}$ which are all completely reducible by our results from §3. Then we develop the theory of weights and describe the irreducible representations of GL_n by their highest weights. Finally, we will have a new look at the Decomposition Theorem 3.3.

5.1 Polynomial and rational representations. Let V be a finite dimensional vector space over K. We want to extend the notion of a regular function on a vector space (1.1) to more general varieties.

Definition. A function $f\colon \mathrm{GL}(V) \to K$ is called *polynomial* if it is the restriction of a polynomial function $\tilde{f} \in K[\mathrm{End}(V)]$. It is called *regular* if $\det^r \cdot f$ is polynomial for some $r \in \mathbb{N}$. The ring of regular functions is called the *coordinate ring* of $\mathrm{GL}(V)$ and will be denoted by $K[\mathrm{GL}(V)]$.

Since $\mathrm{GL}(V)$ is Zariski-dense in $\mathrm{End}(V)$ (1.3) we have in a canonical way:

$$K[\mathrm{End}(V)] \subset K[\mathrm{GL}(V)] = K[\mathrm{End}(V)][\det^{-1}].$$

A representation $\rho\colon \mathrm{GL}(V) \to \mathrm{GL}(W)$ is called *polynomial* if the entries $\rho_{ij}(g)$ of the matrix $\rho(g)$ with respect to a basis of W are polynomial functions on $\mathrm{GL}(V)$. It is called *rational* if they are regular functions on $\mathrm{GL}(V)$. We also say that W is a *polynomial* or a *rational* $\mathrm{GL}(V)$*-module*. We leave it to the reader to check that this does not depend on the choice of a basis in W.

> **Exercise 1.** Let $\rho\colon \mathrm{GL}(V) \to \mathrm{GL}_N(K)$ an irreducible polynomial representation. Then the matrix entries $\rho_{ij} \in K[\mathrm{GL}(V)]$ are homogeneous polynomials of the same degree. More generally, every polynomial representation ρ is a direct sum of representations $\rho^{(i)}$ whose matrix entries are homogeneous polynomials of degree i.
> (Hint: Let $\rho(tE) = \sum_i t^i A_i$. Then the A_i form a system of central idempotents: $A_i A_j = \delta_{ij} A_i$.)

It is obvious how to generalize the notion of polynomial and rational functions to products of the form $\mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \times \cdots$. Thus we can also talk about polynomial and rational representations of such groups.

5.2 Construction of representations and examples. The next lemma follows immediately from the definitions.

Lemma. *Let $\rho\colon \mathrm{GL}(V) \to \mathrm{GL}(W)$ be a rational representation.*

(a) *There is an $n \in \mathbb{N}$ such that $\det^n \otimes \rho$ is polynomial.*

(b) *ρ is polynomial if and only if ρ extends to a polynomial map*

$$\tilde{\rho}\colon \operatorname{End}(V) \to \operatorname{End}(W),$$

and this extension is unique.

Let $\rho\colon \mathrm{GL}_n(K) \to \mathrm{GL}(W)$ be a rational representation where W is a K-vector space. *For every field extension K'/K we get a rational representation*

$$\rho_{K'}\colon \mathrm{GL}_n(K') \to \mathrm{GL}(K' \otimes_K W)$$

and $\rho_{K'}$ is polynomial if and only if ρ is polynomial.

In fact, by the lemma above we may assume that ρ is polynomial and therefore defines a polynomial map $\tilde{\rho}\colon \mathrm{M}_n(K) \to \mathrm{M}_m(K)$ which satisfies the following identity:

$$\tilde{\rho}_{ij}(\dots, \sum_l a_{rl}b_{ls}, \dots) = \sum_k \tilde{\rho}_{ik}(A)\tilde{\rho}_{kj}(B)$$

for $A = (a_{rs}), B = (b_{pq}) \in \mathrm{GL}_n(K)$. Since $\mathrm{GL}_n(K)$ is Zariski-dense in $\mathrm{M}_n(K)$ (Lemma 1.3) this is an identity of polynomials and therefore holds for every field extension K'/K and all $A = (a_{rs}), B = (b_{pq}) \in \mathrm{GL}_n(K')$.

Exercises

2. If $\rho\colon \mathrm{GL}_n(K) \to \mathrm{GL}(W)$ is an irreducible rational representation and assume that $\operatorname{End}_{\mathrm{GL}_n}(W) = K$. Then $\rho_{K'}\colon \mathrm{GL}_n(K') \to \mathrm{GL}(K' \otimes_K W)$ is irreducible, too.
(Cf. 3.1 Exercise 3)

3. Let $\rho\colon \mathrm{GL}_n(K) \to \mathrm{GL}(W)$ be a rational representation, W a finite dimensional K-vector space, and let $k \subset K$ be an arbitrary subfield. Let ${}_kW$ denote the space W considered as a (possibly infinite dimensional) k-vector space. Then the linear action of $\mathrm{GL}_n(k)$ on ${}_kW$ is *locally finite* and rational. (This means that every finite dimensional k-subspace of ${}_kW$ is contained in a finite dimensional $\mathrm{GL}_n(k)$-stable subspace which carries a rational representation of $\mathrm{GL}_n(k)$.)

Examples. We leave it as an exercise to prove the following statements.

(1) The natural representation of $\mathrm{GL}(V)$ on $V^{\otimes m}$ is polynomial.

(2) $\det^n\colon \mathrm{GL}(V) \to \mathrm{GL}_1$ is a rational representation for every $n \in \mathbb{Z}$; it is polynomial for $n \geq 0$.

(3) If W_1, W_2 are polynomial (resp. rational) representations of $\mathrm{GL}(V)$ then so are the *direct sum* $W_1 \oplus W_2$ and the *tensor product* $W_1 \otimes W_2$.

(4) If W is a rational representation then the *dual representation* W^* is rational, too.

(5) *Subrepresentations* and *quotient representations* of polynomial (resp. rational) representations are again polynomial (resp. rational).

(6) If W is a polynomial (resp. rational) representation, then so are the *symmetric powers* $\mathrm{S}^n W$ and the *exterior powers* $\bigwedge^n W$.

Some properties of the exterior powers $\bigwedge^j$ are collected in the following exercises.

Exercises

4. For every linear map $f\colon V \to W$ we have linear maps $\bigwedge^j f\colon \bigwedge^j V \to \bigwedge^j W$ for all j. This defines a regular map

$$\textstyle\bigwedge^j\colon \mathrm{Hom}(V,W) \to \mathrm{Hom}(\bigwedge^j V, \bigwedge^j W)$$

of degree j. If V, W are G-modules then $\bigwedge^j$ is G-equivariant.

5. For every $j = 0, \dots, \dim V$ there are canonical $\mathrm{GL}(V)$-equivariant isomorphisms $\bigwedge^j V^* \xrightarrow{\sim} (\bigwedge^j V)^*$ given by

$$\lambda_1 \wedge \dots \wedge \lambda_j \colon v_1 \wedge \dots \wedge v_j \mapsto \sum_{\sigma \in \mathcal{S}_j} \operatorname{sgn}\sigma\, \lambda_1(v_{\sigma(1)}) \cdots \lambda_j(v_{\sigma(j)}).$$

.

6. Choose a basis $(e_1, \dots, e_n)$ of V and define the linear map

$$\mu\colon \textstyle\bigwedge^{n-1} V \to V^* \quad \text{by } e_1 \wedge \dots \wedge \widehat{e_i} \wedge \cdots e_n \mapsto \varepsilon_i$$

where $(\varepsilon_1, \dots, \varepsilon_n)$ is the dual basis of V^*. Show that

(a) μ is an isomorphism and $\mu(g\omega) = \det g \cdot \mu(\omega)$ for $g \in \mathrm{GL}(V)$.

(b) μ is independent of the choice of a basis, up to a scalar.

(For a generalization see the next exercise.)

7. For $j = 0, 1, \dots, \dim V$ there is a non-degenerate $\mathrm{GL}(V)$-equivariant pairing $\bigwedge^j V \times \bigwedge^{n-j} V \to \bigwedge^n V \simeq K$ given by $(\omega, \mu) \mapsto \omega \wedge \mu$. In particular, we have isomorphisms

$$(\textstyle\bigwedge^j V)^* = \bigwedge^j V^* \simeq \det^{-1} \bigwedge^{n-j} V$$

of $\mathrm{GL}(V)$-modules.

Let us recall that the coordinate ring $K[W]$ can be identified with the symmetric algebra of the dual space (1.1):

$$K[W] = \mathrm{S}(W^*) = \oplus_{n \ge 0} \mathrm{S}^n W^*.$$

So we see that for every rational representation W of $\mathrm{GL}(V)$ the coordinate ring $K[W]$ is a direct sum of rational representations. In particular, every finite dimensional subspace of $K[W]$ is contained in a finite dimensional rational representation. This is expressed by saying that the action of $\mathrm{GL}(V)$ on the coordinate ring $K[W]$ given by

$$(gf)(w) := f(g^{-1}w) \quad \text{for } g \in \mathrm{GL}(V),\ w \in W$$

is *locally finite and rational.*

Exercises

8. The one-dimensional rational representations of $K^* = \mathrm{GL}_1(K)$ are of the form $t \mapsto t^r$ where $r \in \mathbb{Z}$.

9. Every one-dimensional rational representation of $\mathrm{GL}(V)$ is of the form $\det^r\colon \mathrm{GL}(V) \to \mathrm{GL}_1$, $r \in \mathbb{Z}$.
(Hint: Assume K algebraically closed. Choose a basis of V and restrict ρ to the diagonal matrices. Then $\rho(\mathrm{diag}(a_1, \dots, a_n)) = a_1^{r_1} a_2^{r_2} \cdots a_n^{r_n}$ by the previous exercise. Since the matrices $\mathrm{diag}(a_1, \dots, a_n)$, $\mathrm{diag}(a_{\sigma(1)}, \dots, a_{\sigma(n)})$ are both conjugate we get $\rho(\mathrm{diag}(a_1, \dots, a_n)) = (a_1 \cdots a_n)^r$. Now use the fact that the diagonalizable matrices are Zariski-dense in GL_n (2.3 Exercise 4).)

10. If ρ is an irreducible rational representation of $\mathrm{GL}(V)$ then the restriction $\rho|_{\mathrm{SL}(V)}$ is also irreducible.

5.3 Complete reducibility in characteristic zero. The next result holds in any characteristic.

Proposition. *Every polynomial representation of* $\mathrm{GL}(V)$ *is isomorphic to a subrepresentation of a direct sum of the form* $\bigoplus_i V^{\otimes m_i}$.

By Corollary 3.3 this implies that in characteristic zero every polynomial representation of $\mathrm{GL}(V)$ is completely reducible. With Lemma 5.2(a) this extends immediately to every rational representation:

Corollary 1. *If* $\mathrm{char}\, K = 0$ *every rational representation of* $\mathrm{GL}(V)$ *is completely reducible.*

Another consequence is that every irreducible polynomial representation W of $\mathrm{GL}(V)$ occurs in some $V^{\otimes m}$. Moreover, the integer m is uniquely determined by W since we have

$$tw = t^m \cdot w \quad \text{for all } t \in K^* \subset \mathrm{GL}(V) \text{ and } w \in W.$$

Corollary 2. *Every irreducible polynomial representation* W *of* $\mathrm{GL}(V)$ *occurs in a unique* $V^{\otimes m}$.

Definition. The number m is called the *degree* of the representation W.

For the *multiplicative group* $K^* := \mathrm{GL}_1(K)$ we get the following result which holds in any characteristic. It follows immediately from the proposition above and Lemma 5.2(a) (see Exercise 8).

Corollary 3. *Every rational representation of K^* is diagonalizable and the irreducible representations of K^* are of the form $t \mapsto t^m$, $m \in \mathbb{Z}$.*

Proof of Proposition: Let $\rho\colon \mathrm{GL}(V) \to \mathrm{GL}(W)$ be a polynomial representation and $\tilde{\rho}\colon \mathrm{End}(V) \to \mathrm{End}(W)$ its extension (Lemma 6.2(b)). On $\mathrm{End}(V)$ we consider the linear action of $\mathrm{GL}(V)$ given by right multiplication:

$$g\,A := A \cdot g^{-1} \quad \text{for } g \in \mathrm{GL}(V),\ A \in \mathrm{End}(V).$$

For every $\lambda \in W^*$ we define a linear map

$$\varphi_\lambda\colon W \to K[\mathrm{End}(V)] \quad \text{by} \quad \varphi_\lambda(w)(A) := \lambda(\tilde{\rho}(A)w)$$

where $w \in W$ and $A \in \mathrm{End}(V)$. This map φ_λ is $\mathrm{GL}(V)$-equivariant:

$$\begin{aligned}\varphi_\lambda(gw)(A) &= \lambda(\tilde{\rho}(A)gw) = \lambda(\tilde{\rho}(A\cdot g)w) = \varphi_\lambda(w)(A\cdot g)\\ &= (g\varphi_\lambda(w))(A).\end{aligned}$$

Furthermore, $\varphi_\lambda(w)(\mathrm{id}) = \lambda(w)$. Choosing a basis $\lambda_1, \dots, \lambda_m$ of W^* this implies that the linear map

$$\varphi : W \to K[\mathrm{End}(V)]^m, \quad w \mapsto (\varphi_{\lambda_1}(w), \dots, \varphi_{\lambda_m}(w))$$

is injective and $\mathrm{GL}(V)$-equivariant. Thus, every m-dimensional polynomial representation occurs in $K[\mathrm{End}(V)]^m$.

It remains to show that every finite dimensional subrepresentation of $K[\mathrm{End}(V)]$ can be $\mathrm{GL}(V)$-equivariantly embedded in a direct sum of tensor powers $V^{\otimes j}$. By definition of the $\mathrm{GL}(V)$-action on $\mathrm{End}(V)$ we have $\mathrm{End}(V) \simeq (V^*)^n$, hence

$$K[\mathrm{End}(V)] \simeq K[(V^*)^n] \simeq \mathrm{S}(V^n) \simeq \mathrm{S}(V) \otimes \cdots \otimes \mathrm{S}(V)$$

as a $\mathrm{GL}(V)$-module. Now $\mathrm{S}(V) = \bigoplus_{m \geq 0} \mathrm{S}^m V$, and for each m we have a $\mathrm{GL}(V)$-linear embedding

$$\mathrm{S}^m V \hookrightarrow V^{\otimes m}, \quad v_1 v_2 \dots v_m \mapsto \sum_{\sigma \in \mathcal{S}_m} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(m)}$$

which identifies $\mathrm{S}^m V$ with the symmetric tensors in $V^{\otimes m}$. This finishes the proof of the proposition. □

Remark. There are two basic steps in the proof above. First we show that every polynomial representation of $\mathrm{GL}(V)$ occurs in $K[\mathrm{End}(V)]^m$ for some m where the linear action on $K[\mathrm{End}(V)]$ comes from right multiplication of $\mathrm{GL}(V)$ on $\mathrm{End}(V)$. Then we show that every finite dimensional subrepresentation of $\mathrm{End}(V)$ is contained in a direct sum $\oplus_i V^{\otimes n_i}$. We will use this again in a more general situation in the following section.

5.4 Generalization to linear groups and SL_n. We want to extend the previous considerations to arbitrary subgroups $G \subset \mathrm{GL}(V)$, the so-called *linear groups*. A K-valued function $f: G \to K$ is called *regular* (respectively *polynomial*) if it is the restriction of a regular (resp. polynomial) function on $\mathrm{GL}(V)$. We denote by $K[G]$ the algebra of regular functions on G and by $K[G]_{\mathrm{pol}}$ the subalgebra of polynomial functions. Clearly, the two coincide if and only if $G \subset \mathrm{SL}(V)$. As above, this allows to define *polynomial* and *rational* representations $\rho: G \to \mathrm{GL}(W)$ meaning that the matrix entries ρ_{ij} with respect to any basis of W are polynomial, respectively regular functions on G.

It is obvious from the definition that the restriction of a polynomial (resp. rational) representation of $\mathrm{GL}(V)$ to G is again polynomial (resp. rational). Moreover, the standard constructions forming direct sums, tensor products, symmetric and exterior powers, sub- and quotient representations lead again to polynomial (resp. rational) representations (see Examples 5.2).

Exercises

11. Embed $\mathrm{GL}_n(K) \times \mathrm{GL}_m(K)$ into $\mathrm{GL}_{m+n}(K)$ in the usual way:

$$(A, B) \mapsto \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

(a) $\mathrm{GL}_n(K) \times \mathrm{GL}_m(K)$ is ZARISKI-closed in $\mathrm{GL}_{m+n}(K)$.

(b) There is a canonical isomorphism

$$K[\mathrm{GL}_n] \otimes_K K[\mathrm{GL}_m] \xrightarrow{\sim} K[\mathrm{GL}_n \times \mathrm{GL}_m]$$

given by $(f \otimes h)(A, B) := f(A)h(B)$.
(Hint: Consider first the subvector space $\mathrm{M}_n \times \mathrm{M}_m \subset \mathrm{M}_{n+m}$.)

(c) If $G \subset \mathrm{GL}_n(K)$ and $H \subset \mathrm{GL}_m(K)$ are subgroups consider $G \times H$ as a subgroup of $\mathrm{GL}_{m+n}(K)$. Then we have a canonical isomorphism

$$K[G] \otimes_K K[H] \xrightarrow{\sim} K[G \times H]$$

defined as in (b).
(Hint: It follows from (b) that the map exists and is surjective. Now choose a K-basis $(f_i)_{i \in I}$ of $K[G]$ and assume that the function $\sum_i f_i \otimes h_i$ is identically zero on $G \times H$. Then show that $h_i(B) = 0$ for all $B \in H$.)

12. Let $G \subset \mathrm{GL}(V)$ be a subgroup and denote by $\bar{G}$ the ZARISKI-closure:

$$\bar{G} := \{h \in \mathrm{GL}(V) \mid f(h) = 0 \text{ for all } f \in I(G)\}.$$

($I(G)$ is the ideal of G, see 1.3.)

(a) $\bar{G}$ is a subgroup of $\mathrm{GL}(V)$. (Hint: Look at left and right multiplications, first with elements from G, then with those from $\bar{G}$.)

(b) $K[\bar{G}] = K[G]$ and $K[\bar{G}]_{\mathrm{pol}} = K[G]_{\mathrm{pol}}$.

(c) $\bar{G}$ and G have the same rational representations, i.e., $\rho \mapsto \rho|_G$ is an equivalence of categories.

The right multiplication of G on itself determines a linear representation of G on $K[G]$ and on $K[G]_{\mathrm{pol}}$ in the usual way:

$$\rho(g)f\,(h) := f(hg) \quad \text{for } g, h \in G.$$

Lemma. *The representation ρ on $K[G]$ is locally finite and rational. It is polynomial on $K[G]_{\mathrm{pol}}$.*

Proof: By definition, the (linear) restriction maps $K[\mathrm{GL}(V)] \to K[G]$ and $K[\mathrm{End}(V)] \to K[G]_{\mathrm{pol}}$ are surjective and G-equivariant. Thus it suffices to consider the case $G = \mathrm{GL}(V)$. We have already remarked in 5.2 that the representation on $K[\mathrm{End}(V)]$ is locally finite and polynomial. In fact, $\mathrm{End}(V)^* \simeq V \oplus \cdots \oplus V = V^n$ under the given representation and so $K[\mathrm{End}(V)]_m \simeq \mathrm{S}^m(V^n)$. Since $K[\mathrm{GL}(V)] = \bigcup_i \det^{-i} K[\mathrm{End}(V)]$ it follows that the representation on $K[\mathrm{GL}(V)]$ is locally finite and rational. □

Now we can generalize Proposition 5.3.

Proposition 1. *Every polynomial representation of G is isomorphic to a subquotient of a direct sum of the form $\oplus_i V^{\otimes n_i}$.*

(A *subquotient* of a representation W is a subrepresentation of a quotient representation of W. Clearly, a subquotient of a subquotient is again a subquotient.)

Proof: (See Remark at the end of 5.3.) The same argument as in the proof of Proposition 5.3 shows that every polynomial representation W of G occurs in ${K[G]_{\mathrm{pol}}}^m$ for some m. Thus there exists a finite dimensional G-stable subspace $\tilde{W} \subset K[\mathrm{End}(V)]^m$ such that W is a subquotient of $\tilde{W}$. Now we have seen in the proof of Proposition 5.3 that every such $\tilde{W}$ is contained in a direct sum $\oplus_i V^{\otimes n_i}$. □

Corollary. *If the representation of G on $V^{\otimes m}$ is completely reducible for all m, then every rational representation of G is completely reducible. Moreover, every irreducible polynomial representation occurs in some $V^{\otimes m}$.*

As a first application we see that every rational representation of the subgroup $T_n \subset \mathrm{GL}_n(K)$ of *diagonal matrices* is completely reducible. We will discuss this in the next section 5.6.

Proposition 2. *Assume* $\mathrm{char}\, K = 0$. *Then every rational representation ρ of* $\mathrm{SL}(V)$ *is completely reducible. Moreover, ρ is the restriction of a rational representation $\tilde{\rho}$ of* $\mathrm{GL}(V)$, *and ρ is irreducible if and only if $\tilde{\rho}$ is irreducible.*

PROOF: It is easy to see that the restriction of any irreducible representaiton of $\mathrm{GL}(V)$ to $\mathrm{SL}(V)$ is again irreducible (cf. Exercise 10). Thus the representation of $\mathrm{SL}(V)$ on $V^{\otimes m}$ is completely reducible for any m which proves the first claim by the corollary above. Moreover, this implies that every subquotient is isomorphic to a subrepresentation. Hence, every irreducible representation of $\mathrm{SL}(V)$ occurs in some $V^{\otimes m}$ and is therefore the restriction of a polynomial representation of $\mathrm{GL}(V)$. □

Exercise 13. Let ρ, ρ' be two irreducible representations of $\mathrm{GL}(V)$ and assume that $\rho|_{\mathrm{SL}(V)} \simeq \rho'|_{\mathrm{SL}(V)}$. Then $\rho' \simeq \det^r \cdot \rho$ for some $r \in \mathbb{Z}$.
(Hint: $\mathrm{Hom}_{\mathrm{SL}(V)}(W', W)$ is a $\mathrm{GL}(V)$-stable one-dimensional subspace of $\mathrm{Hom}(W', W)$. Now use Exercise 9.)

5.5 FROBENIUS **Reciprocity.** Let G be a finite group and $H \subset G$ a subgroup. If W is a H-module we denote by $\mathrm{Ind}_H^G W$ the *induced G-module* (or *induced representation*) which is usually defined by

$$\mathrm{Ind}_H^G W := KG \otimes_{KH} W.$$

where KG, KH denote the group algebras. On the other hand, every G-module V can be regarded as an H-module by restriction; we will denote it by $V|_H$.

Proposition (FROBENIUS **reciprocity).** *Let V be an irreducible representation of G and W an irreducible representation of H. Assume that* $\mathrm{End}_G(V) = K = \mathrm{End}_H(W)$. *Then*

$$\mathrm{mult}(V, \mathrm{Ind}_H^G W) = \mathrm{mult}(W, V|_H)$$

where $\mathrm{mult}(V, U)$ *denotes the multiplicity of the irreducible representation V in the representation U.*

The proof will follow from a more general result which holds for arbitrary linear groups. It is easy to see that $\mathrm{Ind}_H^G W$ is canonically isomorphic to $\{\eta\colon G \to W \mid \eta(gh^{-1}) = h\eta(g) \text{ for all } h \in H\}$ (see Exercise 14 below).

Definition. Let $H \subset G \subset \mathrm{GL}_n(K)$ be linear groups and let W be a rational H-module (5.4). Then we define the *induced G-module* by

$$\mathrm{Ind}_H^G(W) := \{\eta\colon G \to W \text{ regular} \mid \eta(gh^{-1}) = h\eta(g) \text{ for all } h \in H\}.$$

The right hand side will be shortly denoted by $\mathrm{Mor}_H(G, W)$.

By definition, we have

$$\mathrm{Ind}_H^G(W) = (K[G] \otimes W)^H$$

where the H-action is given by $h(f \otimes w) := f^h \otimes hw$ and $f^h(g) := f(gh)$

$(h \in H, g \in G)$. This shows that $\mathrm{Ind}_H^G(W)$ is a *locally finite and rational G-module.*

Exercise 14. For finite groups $H \subset G$ and any H-module W there is a canonical G-isomorphism

$$\mathrm{Map}_H(G, W) \xrightarrow{\sim} K[G] \otimes_{K[H]} W$$

associating to an H-equivariant map $\eta \colon G \to W$ the element $\sum_{g \in G} g \otimes \eta(g)$. (The H-action on G is given by right multiplication $(g, h) \mapsto gh^{-1}$.)

Theorem. *Let $H \subset G \subset \mathrm{GL}_n(K)$ be linear groups, V a rational G-module and W be a rational H-module. There is a canonical isomorphism*

$$\mathrm{Hom}_G(V, \mathrm{Ind}_H^G(W)) \xrightarrow{\sim} \mathrm{Hom}_H(V|_H, W)$$

given by $\varphi \mapsto e_W \circ \varphi$ where $e_W \colon \mathrm{Ind}_H^G(W) \to W$ sends α to $\alpha(e)$.

PROOF: The map $\varphi \mapsto e_W \circ \varphi$ is clearly well-defined and linear. The inverse map has the following description: If $\psi \colon V|_H \to W$ is H-linear and $v \in V$ define $\varphi_v \colon G \to W$ by $\varphi_v(g) := \psi(g^{-1}v)$. It is clear that φ_v is H-equivariant and so $\varphi_v \in \mathrm{Ind}_H^G W$. Now it is easy to verify that $\varphi \colon v \mapsto \varphi_v$ is a G-equivariant linear map and that $\psi \mapsto \varphi$ is the inverse map to $\varphi \mapsto e_W \circ \varphi$. □

Now suppose that V is an irreducible (rational) representation of G such that $\mathrm{End}_G(V) = K$. If U is a completely reducible (locally finite and rational) representation then the *multiplicity* of V in U is given by

$$\mathrm{mult}(V, U) = \dim \mathrm{Hom}_G(V, U).$$

This shows that the Theorem above generalizes the FROBENIUS Reciprocity for finite groups given in the Proposition at the beginning of this section.

5.6 Representations of tori and weights. In this section the characteristic of K can be arbitrary. Consider the group

$$T_n := \left\{ t = \begin{pmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{pmatrix} \middle| t_1, \ldots, t_n \in K^* \right\} \subset \mathrm{GL}_n(K)$$

of *diagonal matrices* which can be identified with the product

$$\underbrace{\mathrm{GL}_1 \times \mathrm{GL}_1 \times \cdots \times \mathrm{GL}_1}_{n \text{ times}} = (K^*)^n$$

(see Exercise 11). Such a group is called an *n-dimensional torus*. Its coordinate ring is given by

$$K[T_n] = K[x_1, x_1^{-1}, x_2, x_2^{-1}, \dots , x_n, x_n^{-1}] = \bigoplus_{r_1,\dots,r_n\in\mathbb{Z}} Kx_1^{r_1}x_2^{r_2}\cdots x_n^{r_n}$$

and the linear action of T_n on this algebra (by right multiplication, see 5.4) is the obvious one: $\rho(t)x_i = t_i x_i$.

Proposition. *Every rational representation of T_n is diagonalizable and the one-dimensional representations are of the form*

$$\rho\begin{pmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{pmatrix} = t_1^{r_1}t_2^{r_2}\cdots t_n^{r_n}$$

where $r_1, \dots , r_n \in \mathbb{Z}$.

PROOF: This is an immediate consequence of Corollary 5.4. □

Exercise 15. Let $\rho\colon \mathrm{GL}_n \to \mathrm{GL}(W)$ be a rational representation. Then ρ is polynomial if and only if $\rho|_{T_n}$ is polynomial.
(Hint: Use the fact that the diagonalizable matrices are ZARISKI-dense in M_n; see 2.3 Exercise 4.)

The one-dimensional representations of a torus T form a group (under multiplication), the *character group* of T. It will be denoted by $\mathcal{X}(T)$ and is usually written additively:

$$\mathcal{X}(T_n) = \bigoplus_{i=1}^{n} \mathbb{Z}\varepsilon_i \quad \text{where} \quad \varepsilon_i\begin{pmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{pmatrix} := t_i.$$

In other words, $\chi = r_1\varepsilon_1 + r_2\varepsilon_2 + \cdots r_n\varepsilon_n \in \mathcal{X}(T_n)$ corresponds to the rational (invertible) function $x_1^{r_1}x_2^{r_2}\cdots x_n^{r_n} \in K[T_n]$.

Given a rational representation $\rho\colon \mathrm{GL}_n \to \mathrm{GL}(W)$ we can decompose W with respect to T_n:

$$W = \bigoplus_{\lambda\in\mathcal{X}(T)} W_\lambda, \quad W_\lambda := \{w \in W \mid \rho(t)w = \lambda(t)\cdot w \text{ for all } t \in T_n\}.$$

The characters $\lambda \in \mathcal{X}(T)$ such that $W_\lambda \neq 0$ are called the *weights* of W, the eigenspaces W_λ are the corresponding *weight space*, and the non-zero elements of W_λ are the *weight vectors.*

For example, the weights of $\bigwedge^k K^n$ are $\{\varepsilon_{i_1} + \cdots + \varepsilon_{i_k} \mid i_1 < i_2 < \cdots < i_k\}$ and the weights of $\mathrm{S}^k K^n$ are $\{\varepsilon_{j_1} + \cdots + \varepsilon_{j_k} \mid j_1 \leq j_2 \leq \cdots \leq j_k\}$. In both cases the weight spaces are one-dimensional.

Exercises

16. Determine the weights and the weight spaces of the adjoint representation

$$\mathrm{Ad}\colon \mathrm{GL}_n \to \mathrm{GL}(M_n), \quad \mathrm{Ad}(g)A := g \cdot A \cdot g^{-1}.$$

17. Identify the symmetric group $\mathcal{S}_n$ with the subgroup of permutation matrices of GL_n, i.e., $\sigma \in \mathcal{S}_n$ corresponds to the linear map which sends e_i to $e_{\sigma(i)}$, $i = 1, \dots, n$.

(a) $\mathcal{S}_n$ normalizes T_n, and $\mathcal{S}_n T_n = T_n \mathcal{S}_n$ is the normalizer of T_n in GL_n.

(b) Let $\rho\colon \mathrm{GL}_n \to \mathrm{GL}(W)$ be a rational representation. For any weight space W_λ and any $\sigma \in \mathcal{S}_n$ the automorphism $\rho(\sigma)$ of W induces an isomorphism $W_\lambda \simeq W_{\sigma(\lambda)}$ where the weight $\sigma(\lambda)$ is defined by $\sigma(\lambda)(t) := \lambda(\sigma^{-1}\, t\, \sigma)$.

5.7 Fixed vectors under U_n and highest weights. For $i \neq j$ and $s \in K$ define the following elements

$$u_{ij}(s) = E + sE_{ij} \in \mathrm{GL}_n(K),$$

where $E_{ij} \in M_n$ is the matrix with entry 1 in position (i,j) and 0's elsewhere. Clearly, $u_{ij}(s) \cdot u_{ij}(s') = u_{ij}(s+s')$ and so $U_{ij} := \{u_{ij}(s) \mid s \in K\}$ is a subgroup of GL_n isomorphic to the additive group K^+. Furthermore, U_{ij} is normalized by T_n:

$$t\, u_{ij}(s)\, t^{-1} = u_{ij}(t_i t_j^{-1} s) \quad \text{for } t = \begin{pmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{pmatrix} \in T_n,\ s \in K.$$

It is well known that the elements $u_{ij}(s)$ where $i < j$ and $s \in K$ generate the subgroup U_n of upper triangular unipotent matrices:

$$U_n := \left\{ \begin{pmatrix} 1 & * & * \\ & \ddots & * \\ & & 1 \end{pmatrix} \right\} = \langle u_{ij}(s) \mid i < j,\ s \in K \rangle.$$

This group U_n is usually called the *standard unipotent subgroup* of $\mathrm{GL}_n(K)$.

Exercise 18. Give a proof of the last statement.
(Hint: Left multiplication of a matrix A by $u_{ij}(s)$ corresponds to the elementary row operation "adding s times the jth row to the ith row".)

The following is the basic result in the theory of highest weights.

Lemma. *Let λ be a weight of W and $w \in W_\lambda$ a weight vector. There are elements $w_k \in W_{\lambda + k(\varepsilon_i - \varepsilon_j)}$, $k \in \mathbb{N}$ where $w_0 = w$ such that*

$$u_{ij}(s)w = \sum_{k\geq 0} s^k \cdot w_k, \quad s \in K.$$

Proof: Consider the application $\varphi\colon s \mapsto u_{ij}(s)w$. Since W is a rational GL_n-module and $\det u_{ij}(s) = 1$ it follows that $\varphi\colon K \to W$ is a polynomial map, hence of the form $\varphi(s) = \sum_{k\geq 0} s^k \cdot w_k$ for suitable $w_k \in W$. For $t \in T_n$ we get

$$\begin{aligned} t\,\varphi(s) &= t\,u_{ij}(s)w = (tu_{ij}(s)t^{-1})(tw) \\ &= u_{ij}(t_it_j^{-1}s)(tw) = u_{ij}(t_it_j^{-1}s)(\lambda(t)\cdot w) \\ &= \sum_{k\geq 0} \lambda(t)(t_it_j^{-1}s)^k \cdot w_k \end{aligned}$$

for all $s \in K$. Thus we get $t\,w_k = \lambda(t)(t_it_j^{-1})^k \cdot w_k$ which shows that $w_k \in W_{\lambda+k(\varepsilon_i-\varepsilon_j)}$. □

Definition 1. The weights of the form $\sum_{i<j} n_{ij}(\varepsilon_i - \varepsilon_j)$ where $n_{ij} \geq 0$ are called *positive weights*. Equivalently, $\lambda = \sum_i m_i\varepsilon_i$ is positive if and only if $m_1 + m_2 + \cdots + m_n = 0$ and $m_1, m_1 + m_2, m_1 + m_2 + m_3, \ldots \geq 0$. We define a *partial ordering* on the weights by setting

$\lambda \succeq \mu$ if and only if $\lambda - \mu$ is a positive weight.

Now we can prove the main result about weights of GL_n-modules.

Proposition. *Let W be a non-trivial GL_n-module. Then we have $W^{U_n} \neq 0$. Moreover, let $w \in W^{U_n}$ be a weight vector of weight λ and let $W' := \langle \mathrm{GL}_n\, w\rangle \subset W$ be the submodule generated by w. Then the weight space W'_λ is equal to Kw, and the other weights of W' are all $\prec \lambda$.*

Proof: The lemma above implies that for every weight vector $w \in W_\lambda$ and $i < j$ we have $u_{ij}(s)w \in w + \sum_{\mu\succ\lambda} W_\mu$. So, if λ is a maximal weight with respect to the partial ordering then w is fixed under $\langle u_{ij}(s) \mid i < j, s \in K\rangle = U_n$ which proves the first claim.

For the second we need the fact that $U_n^- T_n U_n$ is Zariski-dense in GL_n where $U_n^- := \{u^t \mid u \in U_n\}$ are the lower triangular unipotent matrices (see Exercise 19 below). Thus, for any weight vector $w \in W^{U_n}$ the set $U_n^- T_n U_n K w = U_n^- K w$ is Zariski-dense in $\mathrm{GL}_n\, Kw$. It follows from the lemma above that $U_n^- w \subset w + \sum_{\mu\prec\lambda} W_\mu$ and so $U_n^- Kw \subset Kw \oplus \sum_{\mu\prec\lambda} W_\mu$. Hence, $\mathrm{GL}_n\, w \subset Kw \oplus \sum_{\mu\prec\lambda} W_\mu$ and the claim follows. □

It is easy to see that with the notation of the proof above $W' = \langle \mathrm{GL}_n\, w\rangle = \langle U_n^- w\rangle$ (see 1.3 Exercise 15).

Exercises

19. For a matrix $A = (a_{ij}) \in \mathrm{M}_n(K)$ the determinants $\det A_r$ of the submatrices of the form $A_r := (a_{ij})_{i,j=1}^r$ are called the *principal minors*. Show that $U_n^- T_n U_n$ is the set of those matrices whose principal minors are all $\neq 0$. In particular, $U_n^- T_n U_n$ is ZARISKI-dense in $\mathrm{M}_n(K)$.
$U_n^- T_n U_n$ is called the *open cell* of GL_n.

20. Show that the module W' generated by a weight vector $w \in W^{U_n}$ as in the proposition above is *indecomposable*, i.e., W' cannot be written as a direct sum of two non-zero submodules.

Remark. The torus T_n normalizes the standard unipotent subgroup U_n of GL_n. It follows that for any GL_n-module W the subspace W^{U_n} is stable under T_n and therefore a direct sum of weight spaces.

Corollary 1. *Assume* $\operatorname{char} K = 0$ *and let* W *be a* GL_n*-module. Then* W *is simple if and only if* $\dim W^{U_n} = 1$. *In this case* W^{U_n} *is a one-dimensional weight space* W_λ *and all other weights of* W *are* $\prec \lambda$.

This is clear from the proposition above and the complete reducibility in characteristic zero (3.3 Corollary 1).

Definition 2. A simple module W as in the corollary above is called a *module of highest weight* λ.

Exercise 21. Assume that $\operatorname{char} K = 2$ and consider the representation on $W := \mathrm{S}^2 V$ where $V = K^2$ is the standard representation.

(a) $W^{U_2} = Ke_1^2$, and $\langle \mathrm{GL}_2 \, e_1^2 \rangle = Ke_1^2 \oplus Ke_2^2$ is isomorphic to V.

(b) $(W^*)^{U_2} = Kx_1x_2 \oplus Kx_2^2$, Kx_1x_2 is the determinant representation and $\langle \mathrm{GL}_2 \, x_2^2 \rangle = W^*$.

Corollary 2. *Assume* $\operatorname{char} K = 0$.

(a) *If* W *is a simple* $\mathrm{GL}_n(K)$*-module then* $\mathrm{End}_{\mathrm{GL}_n}(W) = K$.

(b) *Every rational representation of* $\mathrm{GL}_n(K)$ *is defined over* $\mathbb{Q}$.

Recall that this completes the proof of the Decomposition Theorem 3.3.

PROOF: We first remark that a simple GL_n-module W remains simple under base field extensions which implies (a) (cf. 3.3 Exercise 9). In fact, it follows from the weight space decomposition that $(K' \otimes_K W)_\lambda = K' \otimes_K W_\lambda$ and so $(K' \otimes_K W)^{U_n} = K' \otimes_K W^{U_n}$.

We have already seen in Remark 3.3 that assertion (a) is equivalent to the fact that all simple submodules of $V^{\otimes m}$ are defined over $\mathbb{Q}$. This proves

(b) by Corollary 2 of 5.3. Here is another proof of this result: Let W be a simple $\mathrm{GL}_n(K)$-module and λ its highest weight. Then W considered as a $\mathbb{Q}$-vector space is a locally finite and rational $\mathrm{GL}_n(\mathbb{Q})$-module (see Exercise 3). Now choose $w \in W^{U_n}$. By the Proposition above the $\mathbb{Q}$-vector space $W_{\mathbb{Q}} := \langle \mathrm{GL}_n(\mathbb{Q})w\rangle_{\mathbb{Q}} \subset W$ spanned by $\mathrm{GL}_n(\mathbb{Q})w$ is a simple $\mathrm{GL}_n(\mathbb{Q})$-module of highest weight λ. Thus, $K \otimes_{\mathbb{Q}} W_{\mathbb{Q}}$ is also simple and the canonical homomorphism $K \otimes_{\mathbb{Q}} W_{\mathbb{Q}} \to W$ is an isomorphism. □

Examples. Assume $\mathrm{char}\, K = 0$ and let $V = K^n$.

(1) The GL_n-modules $\bigwedge^j V$, $j = 1, 2, \dots, n$ ($V = K^n$) are simple with highest weight $\varepsilon_1 + \cdots + \varepsilon_j$. The GL_n-modules $\mathrm{S}^k V$, $k = 1, 2, \dots$ are simple with highest weight $k\varepsilon_1$.

(In fact, one easily sees that $(\bigwedge^j V)^{U_n} = K(e_1 \wedge \ldots \wedge e_j)$ and that $(\mathrm{S}^k V)^{U_n} = K{e_1}^k$.)

(2) If W is a simple GL_n-module of highest weight $\lambda = p_1\varepsilon_1 + \cdots + p_n\varepsilon_n$ then the dual module W^* is simple of highest weight $\lambda^* := -p_n\varepsilon_1 - p_{n-1}\varepsilon_2 - \cdots - p_1\varepsilon_n$.

(In fact, $p_n\varepsilon_1 + p_{n-1}\varepsilon_2 + \cdots p_1\varepsilon_n = \sigma_0\lambda$ is the lowest weight of W where σ_0 is the order reversing permutation $\sigma_0(i) = n + 1 - i$; see Exercise 17 or the following section 5.8. Now the claim follows since the weights of W^* are $\{-\mu \mid \mu \text{ a weight of } W\}$.)

Exercises

22. ($\mathrm{char}\, K = 0$) Every rational representation of $\mathrm{GL}(V) \times \mathrm{GL}(V')$ is completely reducible. Every irreducible representation is of the form $W \otimes W'$ with irreducible representations W and W' of $\mathrm{GL}(V)$ and $\mathrm{GL}(V')$, respectively. (Cf. 3.3 Exercise 4.)

23. Show that the *traceless* matrices $\mathrm{M}'_n := \{A \in \mathrm{M}_n(K) \mid \mathrm{Tr}\, A = 0\}$ form a simple GL_n-module (with respect to conjugation: $(g, A) \mapsto gAg^{-1}$) of highest weight $\varepsilon_1 - \varepsilon_n$.

24. Consider two rational representations ρ and ρ' of the torus T_n. Show that ρ and ρ' are equivalent if and only if $\mathrm{Tr}\, \rho(t) = \mathrm{Tr}\, \rho'(t)$ for all $t \in T_n$.

5.8 Highest weight modules. Assume that $\mathrm{char}\, K = 0$. To complete the picture we describe the weights which occur as a highest weight of a simple GL_n-module. Moreover, we show that two simple GL_n-modules with the same highest weight are isomorphic.

First we remark that the group $\mathcal{S}_n$ of permutation matrices in GL_n normalizes the torus T_n:

$$\sigma^{-1}\begin{pmatrix} t_1 & & & \\ & t_2 & & \\ & & \ddots & \\ & & & t_n \end{pmatrix}\sigma = \begin{pmatrix} t_{\sigma(1)} & & & \\ & t_{\sigma(2)} & & \\ & & \ddots & \\ & & & t_{\sigma(n)} \end{pmatrix}.$$

Therefore, we get an action of $\mathcal{S}_n$ on the character group $\mathcal{X}(T_n)$ defined by $\sigma(\chi(t)) := \chi(\sigma^{-1}t\sigma)$, i.e. $\sigma(\varepsilon_i) = \varepsilon_{\sigma(i)}$. It is easy to see (cf. Exercise 17 (b)) that for a rational representation $\rho\colon \mathrm{GL}_n \to \mathrm{GL}(W)$ the linear map $\rho(\sigma)$ induces an isomorphism $W_\lambda \xrightarrow{\sim} W_{\sigma(\lambda)}$. Thus, the weights of W are invariant under the action of $\mathcal{S}_n$.

Proposition 1. *The element $\lambda = \sum_{i=1}^n p_i\varepsilon_i$ is a highest weight of a simple GL_n-module if and only if $p_1 \geq p_2 \geq \cdots \geq p_n$. The module is polynomial if and only if $p_n \geq 0$.*

Proof: It is obvious from the above that for every $\mu \in \mathcal{X}(T_n)$ there is a $\sigma \in \mathcal{S}_n$ such that $\sigma(\mu) = \sum_{i=1}^n p_i\varepsilon_i$ satisfies $p_1 \geq p_2 \geq \cdots \geq p_n$. But then $\sigma(\mu) \succeq \mu$ and so a highest weight has to satisfy this condition. In order to construct a simple module with this highest weight we write $\lambda = \sum_{i=1}^n p_i\varepsilon_i$ in the form $\lambda = m_1\omega_1 + m_2\omega_2 + \cdots + m_n\omega_n$ where $\omega_j := \varepsilon_1 + \cdots + \varepsilon_j$ is the highest weight of $\bigwedge^j K^n$ (5.7 Example (2)). By assumption we have $m_1, \ldots, m_{n-1} \geq 0$, and we see that the element

$$\begin{aligned} w &:= e_1^{\otimes m_1} \otimes (e_1 \wedge e_2)^{\otimes m_2} \otimes \cdots \otimes (e_1 \wedge \ldots \wedge e_{n-1})^{\otimes m_{n-1}} \\ &\in V^{\otimes m_1} \otimes (\textstyle\bigwedge^2 V)^{\otimes m_2} \otimes \cdots \otimes (\textstyle\bigwedge^{n-1} V)^{\otimes m_{n-1}} \end{aligned}$$

is fixed under U_n and has weight $\lambda' := \sum_{i=1}^{n-1} p_i\omega_i$. It follows from Proposition 6.6 that the submodule $W' := \langle \mathrm{GL}_n\, w\rangle$ is simple of highest weight λ'. Hence $W := \det^{m_n} \otimes W'$ is simple of highest weight $\lambda = \lambda' + m_n\omega_n$. In addition, W' is polynomial and so W is polynomial, too, in case $m_n = p_n \geq 0$. Conversely, it is clear that for a polynomial representation every weight $\mu = \sum q_i\varepsilon_i$ satisfies $q_i \geq 0$ for all i. □

Definition. A weight of the form $\lambda = \sum_{i=1}^n p_i\varepsilon_i$ where $p_1 \geq p_2 \geq \cdots \geq p_n$ is called a *dominant weight*. The dominant weights form a monoid $\Lambda = \Lambda_{\mathrm{GL}_n}$ generated by the *fundamental weights*

$$\omega_1 := \varepsilon_1, \quad \omega_2 := \varepsilon_1 + \varepsilon_2, \quad \ldots, \quad \omega_n := \varepsilon_1 + \cdots + \varepsilon_n$$

and by $-\omega_n$:

$$\Lambda_{\mathrm{GL}_n} = \mathbb{N}\omega_1 + \mathbb{N}\omega_2 + \cdots + \mathbb{N}\omega_{n-1} + \mathbb{Z}\omega_n.$$

The fundamental weights ω_i are the highest weights of the irreducible representations $\bigwedge^i V$, $i = 1, \ldots, n$ (see 5.7 Example 1 and the proof above).

Remark. Let W be a GL_n-module of highest weight $\lambda = \sum_{i=1}^n p_i \varepsilon_i$. Then W occurs in $V^{\otimes m}$ if and only if $\sum_{i=1}^n p_i = m$ and $p_n \geq 0$. In this case we have $p_i \geq 0$ for all i and $p_j = 0$ for $j > m$. This follows immediately from the construction given in the proof above.

So we see that the number $|\lambda| := \sum_{i=1}^n p_i$ coincides with the degree of W (see 5.3 Corollary 2); we call it the *degree* of λ. The maximal k such that $p_k \neq 0$ is called the *height* of λ. We will come back to this notion in in the next paragraph (Definition 6.2) where we will show among other things that $\mathrm{ht}(\lambda)$ is the minimal k such that W occurs as a subrepresentation of $\mathrm{S}(V^k)$ (6.6 Corollary 2).

> **Exercise 25.** Let W be a GL_n-module of highest weight λ. Then the weight spaces $W_{\det^k}$ are stable under $\mathcal{S}_n \subset \mathrm{GL}_n$, and $W_{\det^k} \neq 0$ if and only if $n = k \cdot |\lambda|$.

Proposition 2. *Two simple* GL_n*-modules are isomorphic if and only if they have the same highest weight.*

Proof: Let W_1 and W_2 be two simple GL_n-modules of the same highest weight λ and let $w_1 \in W_1$ and $w_2 \in W_2$ be highest weight vectors. Then $w := (w_1, w_2) \in W_1 \oplus W_2$ is U_n-invariant and of weight λ. Thus, the submodule $W' := \langle \mathrm{GL}_n w \rangle \subset W_1 \oplus W_2$ is simple. If follows that the two projections $\mathrm{pr}_i \colon W_1 \oplus W_2 \to W_i$ induce non-trivial homomorphisms $W' \to W_i$, hence isomorphisms $W' \xrightarrow{\sim} W_1$ and $W' \xrightarrow{\sim} W_2$. □

> **Exercise 26.** Denote by τ the automorphism $A \mapsto (A^{-1})^t$ of GL_n. If $\rho\colon \mathrm{GL}_n \to \mathrm{GL}(W)$ is a rational representation then $\rho \circ \tau$ is isomorphic to the dual representation ρ^*.

5.9 The decomposition of $V^{\otimes m}$ revisited. Using the results of the previous sections we can reformulate Theorem 3.5 which describes the decomposition of $V^{\otimes m}$ as a $\mathcal{S}_m \times \mathrm{GL}(V)$-module. In §7 we will use a different approach and calculate the characters of the irreducible components in this decomposition.

Proposition 1. *Let $V = K^n$. The $\mathcal{S}_m \times \mathrm{GL}_n$-module $V^{\otimes m}$ admits an isotypic decomposition of the following form*

$$V^{\otimes m} = \bigoplus_\lambda V_\lambda(n) = \bigoplus_\lambda M_\lambda \otimes L_\lambda(n)$$

where λ runs through the set $\{\sum p_i \varepsilon_i \mid p_1 \geq p_2 \geq \cdots \geq p_n \geq 0, \sum p_i = m\}$, $L_\lambda(n)$ is a GL_n-module of highest weight λ and M_λ a simple $\mathcal{S}_m$-module.

Remark 1. In the notation above we use an index n to emphasize the dependence on GL_n. In fact, an element $\lambda = \sum_{i>0} p_i\varepsilon_i$ $(p_1 \geq p_2 \geq \cdots)$ can be seen as a character of T_n as soon as $p_j = 0$ for $j > n$. Hence, there exists a GL_n-module of highest weight λ for every $n \geq \mathrm{ht}(\lambda)$. Moreover, using the usual embedding $\mathrm{GL}_n \subset \mathrm{GL}_{n+1}$ we get a canonical inclusion

$$L_\lambda(n) \subset L_\lambda(n+1) = \langle L_\lambda(n)\rangle_{\mathrm{GL}_{n+1}},$$

and it follows that $\langle V_\lambda(n)\rangle_{\mathrm{GL}_{n+1}} = V_\lambda(n+1)$. Thus, the $\mathcal{S}_n$-modules M_λ do not depend on n.

Remark 2. If we do not want to specify a basis of V we write $L_\lambda(V)$ for the corresponding simple $\mathrm{GL}(V)$-module. In fact, we have in a canonical way

$$L_\lambda(V) = \mathrm{Hom}_{\mathcal{S}_m}(M_\lambda, V^{\otimes m})$$

because $\mathrm{End}_{\mathcal{S}_m}(M_\lambda) = K$. This shows that $V \mapsto L_\lambda(V)$ can be regarded as a *functor*. This means that a linear map $\varphi\colon V \to W$ determines a linear map $L_\lambda(\varphi)\colon L_\lambda(V) \to L_\lambda(W)$, with $L_\lambda(\varphi \circ \psi) = L_\lambda(\varphi) \circ L_\lambda(\psi)$ and $L_\lambda(\mathrm{id}_V) = \mathrm{id}_{L_\lambda(V)}$. In particular, if G is any group and V a G-module then $L_\lambda(V)$ is again a G-module:

$$L_\lambda\colon \mathrm{Mod}_G \to \mathrm{Mod}_G$$

where Mod_G denotes the category of G-modules. As an example, $L_{(d)}(V) = \mathrm{S}^d(V)$ and $L_{(1^r)}(V) = \bigwedge^r V$.

> **Exercise 27.** Show that $L_\lambda(V^*) = L_\lambda(V)^*$ in a canonical way and that this module is isomorphic to $L_{\lambda^*}(V)$.

We can say a bit more about the modules M_λ. Let W be any rational GL_n-module. Then the weight space $W_{\det}$ is a $\mathcal{S}_n$-module because the character det is fixed under $\mathcal{S}_n \subset \mathrm{GL}_n$ (see 5.6 Exercise 17).

Lemma. *Let λ be a highest weight and put $n = |\lambda|$. Then the $\mathcal{S}_n$-module $L_\lambda(n)_{\det}$ is isomorphic to M_λ.*

Proof: We have

$$(V^{\otimes n})_{\det} = \bigoplus_{\sigma\in\mathcal{S}_n} Ke_\sigma \quad \text{where } e_\sigma := e_{\sigma(1)} \otimes e_{\sigma(2)} \otimes \cdots \otimes e_{\sigma(n)}.$$

This is an $\mathcal{S}_n \times \mathcal{S}_n$-module where the action of the second factor comes from the inclusion $\mathcal{S}_n \subset \mathrm{GL}_n$: $(\tau,\nu)e_\sigma = e_{\nu\sigma\tau^{-1}}$ for $(\tau,\nu) \in \mathcal{S}_n \times \mathcal{S}_n$. Clearly, this is the regular representation, i.e., $(V^{\otimes n})_{\det} \simeq \bigoplus_\lambda M_\lambda \otimes M_\lambda$. Thus

$$V_n(\lambda)_{\det} \simeq M_\lambda \otimes L_\lambda(n)_{\det} \simeq M_\lambda \otimes M_\lambda$$

and the claim follows. □

Examples. $(\mathrm{S}^n V)_{\det} = K e_1 e_2 \cdots e_n$ is the trivial representation of $\mathcal{S}_n$ and $(\bigwedge^n V)_{\det} = \bigwedge^n V$ is the sign representation of $\mathcal{S}_n$.

§ 6 Irreducible Characters of $\mathrm{GL}(V)$ and $\mathcal{S}_m$

Following WEYL's *Classical Groups* [Wey46] we describe the characters of the irreducible polynomial representations of GL_n, the so-called SCHUR *polynomials*, and relate them to the characters of the symmetric group. As a consequence, we obtain some classical decomposition formulas (CAUCHY's formula and PIERI's formula).
In this paragraph we assume $\operatorname{char} K = 0$.

6.1 Characters. As before let $T_n \subset \mathrm{GL}_n$ denote the subgroup of diagonal matrices. If $\rho\colon \mathrm{GL}_n \to \mathrm{GL}(W)$ is a rational representation, the function

$$\chi_\rho\colon (x_1, \dots, x_n) \mapsto \operatorname{Tr} \rho(\begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix})$$

is called the *character* of ρ. It will also be denoted by χ_W.

Lemma. *Let ρ, ρ' be rational representations of* GL_n.

(a) $\chi_\rho \in \mathbb{Z}[x_1, {x_1}^{-1}, \dots]$, *and* $\chi_\rho \in \mathbb{Z}[x_1, \dots, x_n]$ *in case ρ is polynomial.*
(b) χ_ρ *is a symmetric function.*
(c) *If ρ and ρ' are equivalent representations of* GL_n *then* $\chi_\rho = \chi_{\rho'}$.

PROOF: (a) follows from the definitions, see 5.1. For (b) we use the action of $\mathcal{S}_n \subset \mathrm{GL}_n$ on T_n by conjugation which we discussed in 5.8. The last statement is clear by definition. □

We have shown in the last paragraph that an irreducible representation ρ of GL_n is determined (up to equivalence) by its highest weight (5.8 Proposition 2). On the other hand, it is easy to see that the character χ_ρ determines all the weights of ρ and even their *multiplicities* (i.e., the dimensions of the weight spaces; see the following Exercise 1). It follows that *two representations with the same character are equivalent.* We will see this again in 6.6 as a consequence of the character theory.

> **Exercise 1.** Show that the character χ_W determines the weights of W and their multiplicities (i.e., the dimensions of the corresponding weight spaces). (Use 5.7 Exercise 24.)

The next proposition collects some simple facts about characters. The proofs are easy and will be left to the reader.

Proposition.

(a) *Let W_1, W_2 be two rational representations of* GL_n. *Then* $\chi_{W_1 \oplus W_2} = \chi_{W_1} + \chi_{W_2}$ *and* $\chi_{W_1 \otimes W_2} = \chi_{W_1} \cdot \chi_{W_2}$.

(b) *If W is an irreducible polynomial representation of degree m* (see 5.3 Corollary 2) *then χ_W is a homogeneous polynomial of degree m.*

(c) *The character of the dual representation W^* is given by*

$$\chi_{W^*}(x_1, \ldots, x_n) = \chi_W(x_1^{-1}, \ldots, x_n^{-1}).$$

Examples. Let $V = K^n$.

(1) $\chi_{V^{\otimes m}} = (x_1 + \cdots + x_n)^m$.

(2) $\chi_{\mathrm{S}^2 V} = \sum_{i \le j} x_i x_j$, $\chi_{\wedge^2 V} = \sum_{i<j} x_i x_j$.

(3) $\chi_{\det} = x_1 x_2 \cdots x_n$, $\chi_{V^*} = x_1^{-1} + \cdots + x_n^{-1}$.

(4) $\chi_{\wedge^{n-1} V} = \sum_{i=1}^{n} x_1 \cdots \widehat{x_i} \cdots x_n = (x_1 \cdots x_n)(x_1^{-1} + \cdots + x_n^{-1})$.

The characters of the symmetric powers $\mathrm{S}^j V$ (i.e., the sum over all monomials of degree j) are called the *complete symmetric polynomials*:

$$h_j(x_1, \ldots, x_n) := \chi_{\mathrm{S}^j V} = \sum_{i_1 \le i_2 \le \cdots} x_{i_1} x_{i_2} \cdots x_{i_j}.$$

We obviously have the following *generating function*:

$$\prod_{i=1}^{n} \frac{1}{1 - x_i t} = \sum_{j=0}^{\infty} h_j \cdot t^j.$$

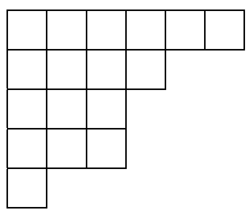
These polynomials h_j are special cases of SCHUR polynomials which we are going to introduce and discuss in the next section.

> **Exercise 2.** The representation $\rho\colon \mathrm{GL}_n \to \mathrm{GL}(W)$ is polynomial if and only if its character χ_ρ is a polynomial.
> (Hint: Use 5.6 Exercise 15.)

6.2 SCHUR **polynomials.** Let $\mathcal{P}$ denote the set of decreasing finite sequences of natural numbers:

$$\mathcal{P} := \{\lambda = (\lambda_1, \lambda_2, \ldots) \mid \lambda_i \in \mathbb{N},\ \lambda_1 \ge \lambda_2 \ge \ldots,\ \lambda_i = 0 \text{ for large } i\}.$$

The elements of $\mathcal{P}$ are called *partitions* and are geometrically represented by their YOUNG *diagram* with rows consisting of $\lambda_1, \lambda_2, \ldots$ boxes, respectively. E.g., the partition $(6, 4, 3, 3, 1)$ is represented by



Mostly, we will identify the partition λ with its YOUNG diagram; if we want to emphasize the difference we use the notion YD(λ) for the YOUNG diagram associated to the partition λ.

By Proposition 1 of 5.8 the highest weights $\sum_i p_i\varepsilon_i$ of irreducible polynomial representations of GL_n can be identified with the partitions $(p_1, p_2, \dots)$ subject to the condition that $p_i = 0$ for $i > n$. This leads to the following definition.

Definition. For $\lambda \in \mathcal{P}$ we define the *height* (or the *length*) of λ by

$$\operatorname{ht}\lambda := \max\{i \mid \lambda_i \neq 0\} = \text{length of the first column of } \mathrm{YD}(\lambda)$$

and the *degree* of λ by

$$|\lambda| := \sum_i \lambda_i = \#\text{ boxes in } \mathrm{YD}(\lambda).$$

For example, for the partition $\lambda = (6,4,3,3,1)$ above we find $\operatorname{ht}\lambda = 5$ and $|\lambda| = 17$.

We have $\mathcal{P} = \bigcup_{m\geq 0}\mathcal{P}_m$, where $\mathcal{P}_m := \{\lambda \in \mathcal{P} \mid |\lambda| = m\}$ are the partitions of m, i.e., the YOUNG diagrams with m boxes.

Now we fix a number $n \in \mathbb{N}$. Let $\lambda \in \mathcal{P}$ be of height $\leq n$, $\lambda = (\lambda_1, \dots, \lambda_n)$. Define

$$\begin{aligned} v_\lambda(x_1,\dots,x_n) \quad &:= \quad \det\begin{pmatrix} x_1^{\lambda_1+n-1} & x_1^{\lambda_2+n-2} & \cdots & x_1^{\lambda_n} \\ x_2^{\lambda_1+n-1} & & & \\ \vdots & & & \\ & & & \end{pmatrix} \\ &= \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}\sigma\, x_{\sigma(1)}^{\lambda_1+n-1} x_{\sigma(2)}^{\lambda_2+n-2}\cdots x_{\sigma(n)}^{\lambda_n}. \end{aligned}$$

Clearly, v_λ is an *alternating homogeneous polynomial* $\in \mathbb{Z}[x_1,\dots,x_n]$ of degree $|\lambda| + \binom{n}{2}$. Moreover,

$$v_{(0,\dots,0)}(x_1,\dots,x_n) = \prod_{i<j}(x_i - x_j)$$

is the VANDERMONDE *determinant* which will be denoted by $\Delta(x_1,\dots,x_n)$. Since every v_λ vanishes for $x_i = x_j$ $(i \neq j)$ it is divisible by Δ. The polynomial

$$s_\lambda(x_1,\dots,x_n) := \frac{v_\lambda(x_1,\dots,x_n)}{\Delta(x_1,\dots,x_n)} \in \mathbb{Z}[x_1,\dots,x_n]$$

is called the SCHUR *polynomial* associated to λ. It follows from the definition that s_λ is a *symmetric homogeneous polynomial of degree* $|\lambda|$.

Exercises

3. Show that $s_{(1,1,\dots,1)} = x_1 x_2 \cdots x_n$. More generally, we have $s_{(k,k,\dots,k)} = (x_1 x_2 \cdots x_n)^k$.

4. Let $\operatorname{ht}\lambda = r$. Then we have

$$s_\lambda(x_1^{-1}, \dots, x_n^{-1})\,(x_1 \cdots x_n)^r = s_{\lambda^c}(x_1, \dots, x_n)$$

where λ^c is the *complementary partition* defined by

$$\lambda^c := (n - \lambda_r, \dots, n - \lambda_1).$$

We denote by $\mathbb{Z}[x_1, \dots, x_n]_{\text{sym}}$ the subring of $\mathbb{Z}[x_1, \dots, x_n]$ of symmetric polynomials and by $\mathbb{Z}[x_1, \dots, x_n]_{\text{alt}}$ the subgroup of alternating polynomials.

Lemma. (a) $\{v_\lambda \mid \operatorname{ht}\lambda \le n\}$ *is a* $\mathbb{Z}$*-bases of* $\mathbb{Z}[x_1, \dots, x_n]_{\text{alt}}$.
(b) $\{s_\lambda \mid \operatorname{ht}\lambda \le n\}$ *is a* $\mathbb{Z}$*-bases of* $\mathbb{Z}[x_1, \dots, x_n]_{\text{sym}}$.

PROOF: (a) Let $f \in \mathbb{Z}[x_1, \dots, x_n]_{\text{alt}}$ and let $\alpha x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}$ be the leading term of f with respect to the *lexicographic ordering* of the exponents $(r_1, \dots, r_n)$. Then $r_1 > r_2 > \dots > r_n \ge 0$ because f is alternating. Now we define $\lambda := (r_1 - n + 1, r_2 - n + 2, \dots, r_n) \in \mathcal{P}$. Then v_λ has leading term $x_1^{r_1} \cdots x_n^{r_n}$ and this term cancels in $f_1 = f - \alpha v_\lambda$. The claim now follows by induction.
(b) This is an immediate consequence of (a): Every alternating function is divisible by the VANDERMONDE determinant Δ and so $\mathbb{Z}[x_1, \dots, x_n]_{\text{alt}} = \Delta \cdot \mathbb{Z}[x_1, \dots, x_n]_{\text{sym}}$. □

Exercises

5. The SCHUR polynomial s_λ has leading term $x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n}$ in the lexicographic ordering.

6. Let $m < n$. Then

$$s_\lambda(x_1, \dots, x_m, 0, \dots, 0) = \begin{cases} s_\lambda(x_1, \dots, x_m) & \text{for } \operatorname{ht}\lambda \le m, \\ 0 & \text{for } \operatorname{ht}\lambda > m. \end{cases}$$

(Hint: One has $v_\lambda(x_1, \dots, x_{n-1}, 0) = x_1 x_2 \cdots x_{n-1}\, v_\lambda(x_1, \dots, x_{n-1})$ if $\lambda_n = 0$ and similarly $\prod_{i<j}(x_i - x_j)\big|_{x_n = 0} = x_1 x_2 \cdots x_{n-1} \prod_{i<j<n}(x_i - x_j)$.)

7. PIERI**'s formula.**Let λ be a partition of height $\le n$. Show that

$$s_\lambda \cdot \prod_{i=1}^{n} \frac{1}{1 - x_i} = \sum_\mu s_\mu$$

where the sum is over all partitions μ such that $\mu_1 \ge \lambda_1 \ge \mu_2 \ge \lambda_2 \ge \cdots \ge \mu_n \ge \lambda_n$. This means that $\operatorname{YD}(\mu)$ is obtained from $\operatorname{YD}(\lambda)$ by adding

some boxes to the rows, at most one box to every column. (Both sides of the equation are considered as formal power series in $x_1, \dots, x_n$.)
(Hint: Multiplying with Δ the claim becomes

$$\sum_\sigma \operatorname{sgn}\sigma\, x_{\sigma(1)}^{\ell_1}\cdots x_{\sigma(n)}^{\ell_n}\cdot \sum_{i_1,\dots,i_n} x_1^{i_1}\cdots x_n^{i_n} = $$
$$= \sum_{m_1,\dots,m_n}\sum_\tau \operatorname{sgn}\tau\, x_{\tau(1)}^{m_1}\cdots x_{\tau(n)}^{m_n}$$

where $\ell_i := \lambda_i + n - i$ and the sum is over all $m_1 \geq \ell_1 > m_2 \geq \ell_2 > \cdots > m_n \geq \ell_n$.)

8. Let λ be a partition of height $\leq n$. Show that

$$s_\lambda \cdot \sum_{i_1<\cdots<i_k} x_{i_1}\cdots x_{i_k} = \sum_\mu s_\mu$$

where the sum is over all partitions μ of height $\leq n$ such that YD(μ) is obtained from YD(λ) by adding some boxes, at most one to every row.
(Hint: See previous exercise.)

6.3 Cauchy**'s formula.** *We have the following formula where both sides are considered as elements in the ring* $\mathbb{Z}[\![x_1,\dots,x_n,y_1,\dots,y_n]\!]$ *of formal power series:*

$$\prod_{\substack{i=1,\dots,n\\ j=1,\dots,m}} \frac{1}{1-x_iy_j} = \sum_{\text{ht}\,\lambda\leq\min(n,m)} s_\lambda(x_1,\dots,x_n)s_\lambda(y_1,\dots,y_m).$$

Proof: It is easy to reduce this to the case $n = m$: Assume $m < n$. Then the claim follows by setting $x_{m+1} = \ldots = x_n = 0$ in the formula for $\prod_{i,j=1}^n \frac{1}{1-x_iy_j}$ and using Exercise 6 above. For $n = m$ Cauchy's formula is a consequence of the following two formulas:

$$\det\left(\frac{1}{1-x_iy_j}\right) = \Delta(x_1,\dots,x_n)\cdot\Delta(y_1,\dots,y_n)\cdot\prod_{i,j=1}^n\frac{1}{1-x_iy_j} \qquad (*)$$
$$\det\left(\frac{1}{1-x_iy_j}\right) = \sum_{\text{ht}\,\lambda\leq n} v_\lambda(x_1,\dots,x_n)\cdot v_\lambda(y_1,\dots,y_n) \qquad (**)$$

The first one is obtained by induction on n. Subtracting the first row from all other rows in the matrix $\left(\frac{1}{1-x_iy_j}\right)_{i,j}$ and using

$$\frac{1}{1-x_iy_j} - \frac{1}{1-x_1y_j} = \frac{x_i-x_1}{1-x_1y_j}\cdot\frac{y_j}{1-x_iy_j}$$

we get

$$\det\left(\frac{1}{1-x_iy_j}\right) =$$

$$= \frac{(x_2-x_1)(x_3-x_1)\cdots(x_n-x_1)}{\prod_{j=1}^n(1-x_1y_j)} \cdot \det\begin{pmatrix} 1 & 1 & \cdots \\ \frac{y_1}{1-x_2y_1} & \frac{y_2}{1-x_2y_2} & \cdots \\ \vdots & & \end{pmatrix}.$$

Now we subtract in the matrix on the right hand side the first column from all others and use

$$\frac{y_j}{1-x_iy_j} - \frac{y_1}{1-x_iy_1} = \frac{y_j-y_1}{1-x_iy_1} \cdot \frac{1}{1-x_iy_j}.$$

This gives

$$\det\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \frac{y_1}{1-x_2y_1} & \frac{y_2}{1-x_2y_2} & \cdots & \frac{y_n}{1-x_2y_n} \\ \vdots & & & \vdots \end{pmatrix} =$$

$$= \frac{(y_2-y_1)(y_3-y_1)\cdots(y_n-y_1)}{\prod_{i=2}^n(1-x_iy_1)} \det\left(\frac{1}{1-x_iy_j}\right)_{i,j=2,\dots,n}.$$

Or

$$\det\left(\frac{1}{1-x_iy_j}\right) =$$

$$= \frac{(x_2-x_1)\cdots(x_n-x_1)(y_2-y_1)\cdots(y_n-y_1)}{\prod_{i \text{ or } j=1}(1-x_iy_j)} \cdot \det\left(\frac{1}{1-x_iy_j}\right)_{i,j=2,\dots}$$

from which $(*)$ follows by induction.

For the second equation $(**)$ we put $\frac{1}{1-x_iy_j} = \sum_{\nu\ge 0}(x_iy_j)^\nu$ and obtain

$$\begin{aligned} \det\left(\frac{1}{1-x_iy_j}\right) &= \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}\sigma \prod_{i=1}^n \sum_{\nu\ge 0}(x_iy_{\sigma(i)})^\nu \\ &= \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}\sigma \sum_{\nu_1,\dots,\nu_n} x_1^{\nu_1}\cdots x_n^{\nu_n} y_{\sigma(1)}^{\nu_1}\cdots y_{\sigma(n)}^{\nu_n} \\ &= \sum_{\nu_1,\dots,\nu_n\ge 0} x_1^{\nu_1}\cdots x_n^{\nu_n} \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}\sigma\, y_{\sigma(1)}^{\nu_1}\cdots y_{\sigma(n)}^{\nu_n}. \end{aligned}$$

This shows that in the first summation we can assume that all ν_i are different. Hence, this summation can be replaced by a double summation over all $\nu_1 >$

$\nu_2 > \ldots > \nu_n$ and all permutations:

$$\sum_{\nu_1>\ldots>\nu_n} \sum_{\tau\in\mathcal{S}_n} x_{\tau(1)}^{\nu_1} \cdots x_{\tau(n)}^{\nu_n} \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}\sigma \cdot y_{\tau\sigma(1)}^{\nu_1} \cdots y_{\tau\sigma(n)}^{\nu_n} =$$
$$= \sum_{\nu_1>\ldots>\nu_n} \sum_{\tau\in\mathcal{S}_n} \operatorname{sgn}\tau \cdot x_{\tau(1)}^{\nu_1} \cdots x_{\tau(n)}^{\nu_n} \cdot \sum_{\mu\in\mathcal{S}_n} \operatorname{sgn}\mu \cdot y_{\mu(1)}^{\nu_1} \cdots y_{\mu(n)}^{\nu_n},$$

and the claim follows from the definition of $v_\lambda(x_1, \ldots, x_n)$. □

6.4 NEWTON **polynomials.** Let $n_i(x_1, \ldots, x_n) := x_1^i + x_2^i + \cdots + x_n^i$ denote the *power sum* and define for $\mu \in \mathcal{P}$:

$$n_\mu(x_1, \ldots, x_n) := \prod_{i\geq 1} n_{\mu_i}(x_1, \ldots, x_n).$$

This is a homogeneous symmetric function in $\mathbb{Z}[x_1, \ldots, x_n]$ of degree $|\mu|$. The $n_\mu(x_1, \ldots, x_n)$ are called NEWTON *polynomials*. Since the SCHUR polynomials form a bases of the symmetric functions (Lemma 6.2 (b)) we can express n_μ in terms of the s_λ's:

$$n_\mu(x_1, \ldots, x_n) = \sum_{\operatorname{ht}\lambda\leq n;\ |\lambda|=|\mu|} a_\lambda(\mu)\, s_\lambda(x_1, \ldots, x_n) \tag{1}$$

where $a_\lambda(\mu) \in \mathbb{Z}$. It is not hard to see that the $a_\lambda(\mu)$ do not depend on n, the number of variables (cf. Exercise 6).

Now recall that there is a canonical bijection between the conjugacy classes in $\mathcal{S}_m$ and the partitions $\mathcal{P}_m$ of m: *For $\sigma \in \mathcal{S}_m$ the corresponding partition $\mu(\sigma)$ is given by the lengths of the cycles in a decomposition of σ as a product of disjoint cycles* (see Exercise 10 in the following section 6.5). Using this we will understand the coefficients $a_\lambda(\mu)$ as *class functions on $\mathcal{S}_m$*, $m := |\lambda|$:

$$a_\lambda \colon \mathcal{S}_m \to \mathbb{Z}, \qquad \sigma \mapsto a_\lambda(\mu(\sigma)).$$

We will show in the next section that the a_λ are exactly the *irreducible characters of $\mathcal{S}_m$*.

Exercise 9. FROBENIUS' **formula.**For a polynomial $f \in \mathbb{Z}[x_1, \ldots, x_n]$ we denote by $[f]_{\ell_1\cdots\ell_n}$ the coefficient of the monomial $x_1^{\ell_1} \cdots x_n^{\ell_n}$ in f. Then we have the following formula:

$$a_\lambda(\mu) = [\Delta \cdot n_\mu]_{\ell_1\cdots\ell_n}$$

where $\ell_i := \lambda_i + n - i$.

Remark. The NEWTON polynomials $n_\mu(x_1, \ldots, x_n)$ can also be interpreted as traces. Let $\sigma \in \mathcal{S}_m$ and consider the endomorphism

$$\varphi := \sigma \circ \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}$$

of $V^{\otimes m}$ where $V = K^n$. *If σ has partition μ then*

$$\operatorname{Tr} \varphi = n_\mu(x_1, \dots, x_n).$$

In fact, the lines $K(e_{i_1} \otimes \cdots \otimes e_{i_m}) \subset V^{\otimes m}$ are stable under φ. Hence, $\operatorname{Tr}\varphi$ is the sum over those monomials $x_{i_1} x_{i_2} \cdots x_{i_m}$ for which $e_{i_1} \otimes \cdots \otimes e_{i_m}$ is fixed under σ. We may assume that

$$\sigma = (1, 2, \dots, \mu_1)(\mu_1 + 1, \dots, \mu_1 + \mu_2) \cdots (m - \mu_s + 1, \dots, m).$$

Then a tensor $e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_m}$ is fixed under σ if and only if it is of the form $e_{j_1}^{\otimes \mu_1} \otimes e_{j_2}^{\otimes \mu_2} \otimes \cdots \otimes e_{j_s}^{\otimes \mu_s}$. Hence

$$\begin{aligned} \operatorname{Tr} \varphi &= \sum_{j_1, \dots, j_s} x_{j_1}^{\mu_1} \cdots x_{j_s}^{\mu_s} = (\sum_i x_i^{\mu_1})(\sum_i x_i^{\mu_2}) \cdots (\sum_i x_i^{\mu_s}) \\ &= n_\mu(x_1, \dots, x_n). \end{aligned}$$

6.5 The irreducible characters of $\mathcal{S}_m$. We want to show that the coefficients a_λ in formula (1) of 6.4 are the irreducible characters of $\mathcal{S}_m$. A first step is to prove the following *orthogonality relations*:

Lemma 1. *Let $\lambda, \lambda' \in \mathcal{P}_m$. Then*

$$\sum_{\sigma \in \mathcal{S}_m} a_\lambda(\sigma)\, a_{\lambda'}(\sigma) = \begin{cases} m! & \text{for } \lambda = \lambda', \\ 0 & \text{otherwise.} \end{cases}$$

Proof: We use Cauchy's formula 6.3. Taking the formal logarithm we get

$$\log \left(\prod_{i,j=1}^{n} \frac{1}{1 - x_i y_j} \right) = \sum_{i,j=1}^{n} \sum_{\nu \geq 0} \frac{x_i^\nu y_j^\nu}{\nu} = \sum_{\nu \geq 0} \frac{n_\nu(x) n_\nu(y)}{\nu}.$$

Since $\exp \log = \mathrm{id}$ we can calculate the term of degree m in the expansion of $\prod_{i,j} \frac{1}{1 - x_i y_j}$ in the following way:

$$\begin{aligned} R_m &= \text{term of degree } m \text{ in } \exp \left(\sum_{\nu \geq 0} \frac{n_\nu(x) n_\nu(y)}{\nu} \right) \\ &= \sum_{\substack{r_1 + 2r_2 + \\ \cdots + s r_s = m}} \frac{n_1(x)^{r_1} n_2(x)^{r_2} \cdots n_s(x)^{r_s} n_1(y)^{r_1} \cdots n_s(y)^{r_s}}{1^{r_1} 2^{r_2} \cdots s^{r_s} r_1! r_2! \cdots r_s!} \end{aligned}$$

Now the sequences $r = (r_1, \dots , r_s)$ with $r_1 + 2r_2 + \cdots + sr_s = m$ correspond bijectively to the partitions of m:

$$r = (r_1, \dots , r_s) \leftrightarrow \mu = (\underbrace{s, s, \dots , s}_{r_s}, \underbrace{s-1, \dots , s-1}_{r_{s-1}}, \dots , \underbrace{1, 1, \dots , 1}_{r_1}),$$

and the number

$$\gamma(\mu) := \frac{m!}{1^{r_1}2^{r_2}\cdots s^{r_s}r_1!\cdots r_s!}$$

is equal to the number of elements in the conjugacy class of $\mathcal{S}_m$ corresponding to μ (see Exercise 10 below). Hence

$$\begin{aligned} R_m &= \sum_{\mu\in\mathcal{P}_m} \frac{1}{m!}\gamma(\mu)\, n_\mu(x)n_\mu(y) \\ &= \sum_{\mu\in\mathcal{P}_m} \frac{1}{m!}\gamma(\mu) \sum_{\lambda,\lambda'\in\mathcal{P}_m} a_\lambda(\mu)a_{\lambda'}(\mu)s_\lambda(x)s_{\lambda'}(y) \\ &= \sum_{\lambda,\lambda'\in\mathcal{P}_m} \left(\sum_{\mu\in\mathcal{P}_{\Updownarrow}} \frac{1}{m!}\gamma(\mu)\, a_\lambda(\mu)a_{\lambda'}(\mu) \right) s_\lambda(x)s_{\lambda'}(y). \end{aligned}$$

Looking at the right hand side of CAUCHY's formula 6.3, we get

$$R_m = \sum_{\lambda\in\mathcal{P}_m} s_\lambda(x)s_\lambda(y),$$

and the claim follows. □

Exercise 10. Show that two permutations $\sigma, \tau \in \mathcal{S}_m$ are conjugate if and only if they correspond to the same partition λ (i.e., they have the same cycle lengths in their decomposition into disjoint cycles). Prove that the number of elements in the conjugacy class corresponding to the partition $\lambda = (s^{r_s}, (s-1)^{r_{s-1}}, \dots , 2^{r_2}, 1^{r_1})$ is given by

$$\gamma(\lambda) := \frac{m!}{1^{r_1}2^{r_2}\cdots s^{r_s}r_1!\cdots r_s!}.$$

The next step is to show that the a_λ are *virtual characters*, i.e., $\mathbb{Z}$-linear combinations of characters. Recall that there is another $\mathbb{Z}$-bases $\{r_\lambda \mid \lambda \in \mathcal{P}, \operatorname{ht}\lambda \le n\}$ of the symmetric polynomials $\mathbb{Z}[x_1, \dots , x_n]_{\text{sym}}$ which has already been used in the proof of Lemma 3.1: For $\lambda = (\lambda_1, \lambda_2, \dots , \lambda_n)$ the function $r_\lambda(x_1, \dots , x_n)$ is the sum of the monomials in the orbit of $x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n}$ under $\mathcal{S}_n$. Now we express the NEWTON polynomials in terms of the r_λ's:

$$n_\mu(x_1,\dots,x_n) = \sum_{|\lambda|=|\mu|} b_\lambda(\mu)\, r_\lambda(x_1,\dots,x_n). \tag{2}$$

Lemma 2. *With the notation above we have:*

(a) $a_\lambda \in \sum_\eta \mathbb{Z}\, b_\eta$*;*

(b) $b_\eta \in a_\eta + \sum_{\lambda>\eta} \mathbb{Z}\, a_\lambda$*, where* $>$ *is the* lexicographic ordering*;*

(c) b_λ *is the character of the induced representation* $\operatorname{Ind}_{S_\lambda}^{S_m} K$ *of the trivial representation* K *where* $S_\lambda := S_{\lambda_1} \times \cdots \times S_{\lambda_r} \subset S_m$*,* $m := |\lambda|$.

(Recall that the *induced representation* $\operatorname{Ind}_{S_\lambda}^{S_m} K$ is the permutation representation coming from the action of S_m on the coset space S_m/S_λ, i.e., $\operatorname{Ind}_{S_\lambda}^{S_m} K$ has a basis of the form $e_{\bar\tau}$ ($\bar\tau \in S_m/S_\lambda$) and $\sigma(e_{\bar\tau}) = e_{\overline{\sigma\tau}}$.)

PROOF: (a) To see this we express the r_λ's in (2) in terms of SCHUR polynomials and compare with equation (1) in 6.4.

(b) With respect to the lexicographic ordering of the exponents the function $v_\lambda = \Delta\cdot s_\lambda$ has leading term $x_1^{\lambda_1+n-1} x_2^{\lambda_2+n-2}\cdots x_n^{\lambda_n}$. Since the function $\Delta\cdot r_\lambda = \prod_{i<j}(x_i - x_j)\cdot r_\lambda$ has the same leading term we must have a relation of the form:

$$s_\lambda = r_\lambda + \sum_{\lambda'<\lambda} \gamma_{\lambda\lambda'} r_{\lambda'}, \quad \gamma_{\lambda\lambda'} \in \mathbb{Z}.$$

Putting this into formula (1) of 6.4 and comparing with (2) above we get

$$b_\eta = a_\eta + \sum_{\lambda>\eta} \gamma_{\lambda\eta} a_\lambda.$$

(c) We have to show that $b_\lambda(\mu)$ is the number of fixed points of a permutation $\sigma \in \mathcal{S}_m$ with partition μ acting on the coset space $\mathcal{S}_m \,/\, \mathcal{S}_\lambda$. Equation (2) above has the form

$$\begin{aligned}(x_1^{\mu_1} + x_2^{\mu_1} + \cdots)(x_1^{\mu_2} + x_2^{\mu_2} + \cdots)\cdots &= \\ = \sum_\lambda b_\lambda(\mu)(x_1^{\lambda_1} x_2^{\lambda_2}\cdots x_n^{\lambda_n} + \cdots).\end{aligned}$$

This shows that $b_\lambda(\mu)$ is the number of possibilities to decompose the set $M = \{\mu_1, \mu_2, \dots, \mu_m\}$ into m disjoint subsets $M = M_1 \cup M_2 \cup \ldots \cup M_m$ such that the sum of the μ_j's in M_i is equal to λ_i. We claim that this number equals the number of fixed points of a permutation σ with partition μ acting on $\mathcal{S}_m \,/\, \mathcal{S}_\lambda$. In fact, $\tau\,\mathcal{S}_\lambda$ is fixed under σ if and only if $\tau^{-1}\sigma\tau \in \mathcal{S}_\lambda$ which implies that the set $M = \{\mu_1, \ldots, \mu_m\}$ of the cycle lengths of σ is decomposable in the way described above. It is now easy to see that the fixed points correspond exactly to the different decompositions of M. □

Now we are ready to prove the main result of this section.

Theorem. *For every $\lambda \in \mathcal{P}_m$ there is an irreducible K-linear representation of $\mathcal{S}_m$ with character a_λ. In particular, the a_λ ($\lambda \in \mathcal{P}_m$) are the irreducible characters of $\mathcal{S}_m$.*

We use the notation $M_\lambda(K) = M_\lambda$ for a simple $\mathcal{S}_m$-module (defined over K) with character a_λ. We will see in a moment that this is in accordance with the notation introduced in 3.3 and 5.9.

PROOF: It follows from Lemma 2 (a) and (c) that $a_\lambda = \sum \gamma_i \chi_i$ with irreducible characters χ_i and $\gamma_i \in \mathbb{Z}$. The orthogonality relations (Lemma 1) imply $\sum \gamma_i^2 = 1$. Hence a_λ or $-a_\lambda$ is an irreducible character. Using Lemma 2 (b) (and again (c)) we see that the sign must be $+1$ since the character b_λ is a sum of irreducible characters with positive coefficients. It also follows that the representation $\mathrm{Ind}_{\mathcal{S}_\lambda}^{\mathcal{S}_m} K$ must contain a unique subrepresentation M_λ with character a_λ because a_λ occurs with multiplicity one in b_λ. □

Exercises

11. Let $\lambda = (\lambda_1, \ldots, \lambda_r)$ be a partition of m. Show that

$$\dim M_\lambda = \frac{m!}{\ell_1! \cdots \ell_r!} \prod_{i<j} (\ell_i - \ell_j)$$

where $\ell_i = \lambda_i + m - i$.
(Hint: Use the FROBENIUS formula of Exercise 9:

$$\dim M_\lambda = a_\lambda(1) = [\Delta (x_1 + \cdots + x_r)^r]_{\ell_1 \cdots \ell_r}.$$

12. Hook length formula. To each box B in a YOUNG diagram we associate a *hook* consisting of all boxes below or to the right hand side of B; its length is by definition the number of boxes in the hook. Prove the following formula:

$$\dim M_\lambda = \frac{m!}{\prod \text{hook lengths}}.$$

(Hint: Use the previous exercise.)

The next result is clear from the above (cf. Proposition 3.3 and Corollary 5.7).

Corollary. *For every field extension K'/K we have $M_\lambda(K) \otimes_K K' \cong M_\lambda(K')$ as $\mathcal{S}_m$-modules. In particular, every $\mathcal{S}_m$-module M is defined over $\mathbb{Q}$, and for a simple module M we get $\mathrm{End}_{\mathcal{S}_m}(M) = K$.*

Examples. (1) For $\lambda = (m)$ we have $\mathrm{Ind}_{\mathcal{S}_m}^{\mathcal{S}_m} K = K$, hence $b_{(m)} = a_{(m)}$ is the trivial character and $M_{(m)} = K$, the trivial representation.

(2) For $\lambda = (m-1,1)$ we see that $\mathrm{Ind}_{\mathcal{S}_{m-1}}^{\mathcal{S}_m} K$ is the natural representation of $\mathcal{S}_m$ on K^m and $b_{(m-1,1)} = a_{(m-1,1)} + a_{(m)}$. Thus $M_{(m-1,1)} \simeq K^n/K(1,1,\dots,1)$.

6.6 The irreducible characters of GL_n**.** Now we are in position to prove the main result about the characters of GL_n which is due to Schur.

Theorem. *For every partition λ of height $\leq n$ there is an irreducible polynomial representation L_λ of GL_n whose character is the* Schur *polynomial s_λ. The L_λ represent all isomorphism classes of simple polynomial GL_n-modules. Moreover, L_λ has highest weight $\sum_i \lambda_i \varepsilon_i$.*

Proof: Consider the representation of $\mathcal{S}_m \times \mathrm{GL}_n$ on $V^{\otimes m}$, $V = K^n$ and its character

$$\chi(\sigma; x_1, \dots, x_n) := \mathrm{Tr}(\sigma \circ \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}) \colon \mathcal{S}_m \times T_n \to K.$$

Since the a_λ form a $\mathbb{Z}$-basis of the class functions on $\mathcal{S}_m$ we can write χ in the form

$$\chi = \sum_{\lambda \in \mathcal{P}_m} a_\lambda \cdot \tilde{s}_\lambda \quad \text{with} \quad \tilde{s}_\lambda \in \mathbb{Z}[x_1, \dots, x_n].$$

On the other hand, we have the $\mathcal{S}_m \times \mathrm{GL}(V)$-stable decomposition

$$V^{\otimes m} \cong \bigoplus_{\lambda \in \mathcal{P}'} M_\lambda \otimes L_\lambda$$

with a suitable subset $\mathcal{P}' \subset \mathcal{P}_m$ and certain irreducible polynomial representations L_λ of GL_n (Theorem 3.3). Hence, by uniqueness, $\tilde{s}_\lambda = \chi_{L_\lambda}$ for $\lambda \in \mathcal{P}'$ (and $\tilde{s}_\lambda = 0$ otherwise). Now we have seen in Remark 6.4 that

$$\chi(\sigma, x_1, \dots, x_n) = n_\mu(x_1, \dots, x_n)$$

where $\mu = \mu(\sigma)$ is the partition of σ. Since we know from equation (1) in 6.4 that

$$n_\mu(x_1, \dots, x_n) = \sum_{|\lambda| = m,\ \mathrm{ht}\,\lambda \leq n} a_\lambda(\mu) s_\lambda(x_1, \dots, x_n)$$

we finally get $\chi_{L_\lambda} = s_\lambda$ for $\lambda \in \mathcal{P}' = \{\lambda \in \mathcal{P}_m \mid \mathrm{ht}\,\lambda \leq n\}$.

The last statement of the theorem is clear since s_λ has leading term $x_1^{\lambda_1} \cdots x_n^{\lambda_n}$ in the lexicographic ordering (see Exercise 5). □

Remark. By the last statement of the theorem the notation L_λ is in accordance with the one introduced in 5.9. We also write $L_\lambda(n)$ or $L_\lambda(K)$ in order to emphasize the group GL_n or the base field K. If we do not want to specify a basis in V we use the notation $L_\lambda(V)$ for the corresponding simple polynomial $\mathrm{GL}(V)$-module for which we have the following functorial description (cf. Remark 5.9):

$$L_\lambda(V) = \mathrm{Hom}_{\mathcal{S}_m}(M_\lambda, V^{\otimes m}).$$

Moreover, we get

$$\mathrm{End}_{\mathrm{GL}_n} L_\lambda(K) = K \quad \text{and} \quad L_\lambda(K) \otimes_K K' \cong L_\lambda(K')$$

for any field extension K'/K. Since the SCHUR polynomials s_λ are linearly independent (Lemma 6.2 (b)) we see again that every rational representation of GL_n is determined by its character, up to equivalence (cf. 5.8 Exercise 24).

From the above we get the following decomposition formula for $V^{\otimes m}$ as a $\mathcal{S}_m \times \mathrm{GL}(V)$-module which was considered earlier in 3.3 and again in 5.9. It can now be seen as a representation theoretic interpretation of formula (1) of 6.4 expressing the NEWTON polynomials in terms of the SCHUR polynomials.

Corollary 1. *We have*

$$V^{\otimes m} \cong \bigoplus_{|\lambda|=m,\ \mathrm{ht}\,\lambda \le \dim V} M_\lambda(K) \otimes L_\lambda(V)$$

as $\mathcal{S}_m \times \mathrm{GL}(V)$*-modules.*

We can also give representation theoretic interpretation of CAUCHY's formula 6.3.

Corollary 2. *As a representation of* $\mathrm{GL}(V) \times \mathrm{GL}(W)$ *we have*

$$\mathrm{S}(V \otimes W) \cong \bigoplus_{\mathrm{ht}\,\lambda \le h} L_\lambda(V) \otimes L_\lambda(W)$$

where $h := \min(\dim V, \dim W)$*. More precisely,*

$$\mathrm{S}^d(V \otimes W) \cong \bigoplus_{\mathrm{ht}\,\lambda \le h,\ |\lambda|=d} L_\lambda(V) \otimes L_\lambda(W).$$

In particular, we see that $L_\lambda(V)$ *occurs in* $\mathrm{S}(V^m)$ *if and only if* $\mathrm{ht}\,\lambda \le m$*.*

Proof: It is easy to see that

$$\prod_{i=1\dots n, j=1\dots,m} \frac{1}{1-x_i y_j}$$

is the (formal) character of the representation of $\mathrm{GL}_n \times \mathrm{GL}_m$ on the symmetric algebra $\mathrm{S}(K^n \otimes K^m) = K[x_{ij} \mid i = 1, \dots, n; j = 1, \dots, m]$. Now we can argue as in the proof of the Theorem above to get the required result. □

Finally, from Pieri's formula (Exercises 7 and 8) we obtain the following rules:

Corollary 3. *For any partition λ of height $\leq n = \dim V$ we have*

$$L_\lambda \otimes S^m(V) = \bigoplus_\mu L_\mu(V)$$

where the sum if over all partitions μ of degree $|\mu| = |\lambda| + m$ and height $\leq n$ whose Young *diagram is obtained from* $\mathrm{YD}(\lambda)$ *by adding m boxes, at most one to each column. Similarly, one gets*

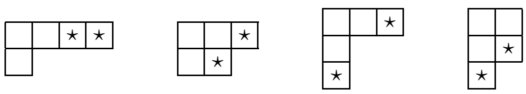
$$L_\lambda \otimes \bigwedge^m(V) = \bigoplus_\mu L_\mu(V)$$

where the sum if over all partitions μ of degree $|\mu| = |\lambda| + m$ and height $\leq n$ whose Young *diagram is obtained from* $\mathrm{YD}(\lambda)$ *by adding m boxes, at most one to each row.*

Example. We have the following decompositions

$$L_{(2,1)}(V) \otimes S^2(V) \cong L_{(4,1)} \oplus L_{(3,2)} \oplus L_{(3,1,1)} \oplus L_{(2,2,1)}$$

corresponding to the Young diagrams



and similarly $L_{(2,1)}(V) \otimes \bigwedge^2(V) \cong L_{(3,2)} \oplus L_{(3,1,1)} \oplus L_{(2,2,1)} \oplus L_{(2,1,1,1)}$.

> **Exercise 13.** Show that $\mathrm{ht}\,\lambda$ is the smallest integer r such that $\det^r L_\lambda(n)^*$ is a polynomial module and prove that $\det^{\mathrm{ht}\,\lambda} L_\lambda(n)^* \simeq L_{\lambda^c}(n)$ where λ^c is the complementary partition $(n - \lambda_r, \dots, n - \lambda_1)$, $r = \mathrm{ht}\,\lambda$.
> (Hint: Use 6.1 Proposition (c) and Exercises 2 and 4.)

6.7 Decomposition of tensor products. Let G be a linear group and assume that every rational representation of G is completely reducible. Given two irreducible representations V and W of G it is an important task to determine the decomposition of the tensor product $V \otimes W$ as a direct sum of irreducible representations. This means that we want to calculate the multiplicities

$$m_\lambda := \operatorname{mult}(V_\lambda, V \otimes W)$$

where λ parametrizes the irreducible representations of G. We have already seen two such examples in the previous paragraph, deduced from PERIE's formula, namely the decomposition of the GL(V)-modules $L \otimes \mathrm{S}^m(V)$ and $L \otimes \bigwedge^m(V)$ (6.6 Corollary 3). We will first show that there is an interesting relation between such multiplicities for the general linear group and those for the symmetric group.

For three partitions λ, μ, ρ of height $\leq n$ we define the following multiplicities:

$$N_\rho^{\lambda\mu} := \operatorname{mult}(L_\rho, L_\lambda \otimes L_\mu).$$

where all modules are considered as GL_n-modules. We will see in a moment that $N_\rho^{\lambda\mu}$ does not depend on n as long as $n \geq \operatorname{ht}\lambda, \operatorname{ht}\mu, \operatorname{ht}\rho$. Equivalently, we have

$$L_\lambda \otimes L_\mu \simeq \bigoplus_\rho N_\rho^{\lambda\mu} L_\rho.$$

Proposition. *Let λ be a partition of p, μ a partition of q and ρ a partition of $m := p + q$. Consider $\mathcal{S}_p \times \mathcal{S}_q$ as a subgroup of $\mathcal{S}_m$ in the usual way. Then we have*

$$N_\rho^{\lambda\mu} := \operatorname{mult}(L_\rho, L_\lambda \otimes L_\mu) = \operatorname{mult}(M_\lambda \otimes M_\mu, M_\rho|_{\mathcal{S}_p \times \mathcal{S}_q}),$$

and $N_\rho^{\lambda\mu} = 0$ if $|\rho| \neq |\lambda| + |\mu|$.

PROOF: This is an easy consequence of CAUCHY's formula (6.6 Corollary 1). We have

$$V^{\otimes p} \cong \bigoplus_{|\lambda|=p} M_\lambda \otimes L_\lambda \quad \text{and} \quad V^{\otimes q} \cong \bigoplus_{|\mu|=q} M_\mu \otimes L_\mu.$$

Thus we get for $V^{\otimes m} = V^{\otimes p} \otimes V^{\otimes q}$ as an $\mathcal{S}_p \times \mathcal{S}_q \times \mathrm{GL}_n$-module

$$\begin{aligned} V^{\otimes m} &\cong \bigoplus_{|\rho|=m} M_\rho \otimes L_\rho \\ &\cong \bigoplus_{|\lambda|=p,|\mu|=q} (M_\lambda \otimes M_\mu) \otimes (L_\lambda \otimes L_\mu) \\ &\cong \bigoplus_{|\lambda|=p,|\mu|=q,|\rho|=m} (M_\lambda \otimes M_\mu) \otimes (N_\rho^{\mu\lambda} L_\rho) \\ &\cong \bigoplus_{|\rho|=m} \left(\bigoplus_{|\lambda|=p,|\mu|=q} N_\rho^{\mu\lambda} (M_\lambda \otimes M_\mu) \right) \otimes L_\rho . \end{aligned}$$

This shows that $M_\rho|_{\mathcal{S}_p \times \mathcal{S}_q} \cong \bigoplus_{|\lambda|=p,|\mu|=q} N_\rho^{\mu\lambda}(M_\lambda \otimes M_\mu)$ and the claim follows. □

§ 7 Some Classical Invariant Theory

In the following paragraphs we give another approach to the First Fundamental Theorem for GL_n which will enable us to generalize it to all classical groups. It is based on the representation theory of the general linear group and on the CAPELLI-DERUYTS expansion. Along these lines we discuss two fundamental results due to H. WEYL concerning invariants of several copies of a representation V of an arbitrary group $G \subset \mathrm{GL}(V)$. In this context we introduce the algebra of differential operators generated by the polarization operators. They will play an important rôle in the theory of CAPELLI. Finally, we study multiplicity free algebras and show how they are related to the problems discussed so far and how they can be used to obtain different proofs of some of our previous results.
Throughout the end of the text we assume $\operatorname{char} K = 0$.

7.1 GL_p-action on invariants. We come back to the general situation from the very beginning of these notes (see §1.5). Let V be a finite dimensional K-vector space and $G \subset \mathrm{GL}(V)$ an arbitrary subgroup. We want to discuss the G-invariants of p copies of the representation V where p is any natural number. For this purpose we introduce the following linear action of GL_p on $V^p = V \oplus \cdots \oplus V$:

$$h(v_1, \ldots, v_p) := (v_1, \ldots, v_p) \cdot h^{-1} \quad \text{for } h \in \mathrm{GL}_p$$

where the right hand side has the obvious meaning of multiplication of matrices: If $A = (a_{ij})_{i,j=1}^p$ then

$$(v_1, \ldots, v_p) \cdot A := (\ldots, \sum_{i=1}^p a_{ij} v_i, \ldots).$$

Choosing a basis and identifying V^p with the $n \times p$ matrices $\mathrm{M}_{n \times p}$ this action is just right multiplication with the matrix h^{-1}.

Remark. There are two natural actions of GL_p on V^p depending on the following two identifications:

$$V^p = K^p \otimes V \quad \text{or} \quad V^p = \mathrm{Hom}(K^p, V).$$

The two actions are obtained from each other by the *outer automorphism* $h \mapsto (h^{-1})^t$. All results we are going to describe are true for both actions, perhaps with some minor changes in notation. We have chosen the latter since it has the advantage that the corresponding representation on the coordinate ring $K[V^p]$ is polynomial.

The action of GL_p on V^p clearly commutes with the natural diagonal action of $\mathrm{GL}(V)$. Hence, we get the following result about the G-invariants.

Lemma. *The ring of invariants $K[V^p]^G$ is stable under GL_p.*

For $r \le p$ we include $V^r \subset V^p$ using the first r copies of V:

$$V^r \hookrightarrow V^p \quad : \quad (v_1, \ldots, v_r) \mapsto (v_1, \ldots, v_r, 0, \ldots, 0).$$

This map is equivariant with respect to the inclusion of the corresponding groups:

$$\mathrm{GL}_r \hookrightarrow \mathrm{GL}_p \quad : \quad A \mapsto \begin{pmatrix} A & 0 \\ 0 & E_{p-r} \end{pmatrix}. \tag{1}$$

Moreover, we have an inclusion of the coordinate rings

$$K[V^r] \hookrightarrow K[V^p]$$

corresponding to the linear projection $\mathrm{pr}\colon V^p \to V^r$ onto the first r copies, i.e., we identify $K[V^r]$ with those functions $f \in K[V^p]$ which do not depend on the last $p - r$ copies of V. This inclusion is clearly $\mathrm{GL}(V)$-equivariant, but also equivariant with respect to the homomorphism (1) above. In particular, we obtain

$$K[V^r]^G = K[V^r] \cap K[V^p]^G \subset K[V^p]^G.$$

By the lemma above $K[V^p]^G$ even contains the GL_p-module generated by $K[V^r]^G$:

$$\langle K[V^r]^G \rangle_{\mathrm{GL}_p} \subset K[V^p]^G. \tag{2}$$

(As before, we use the notation $\langle S \rangle_{\mathrm{GL}_p}$ for the GL_p-module generated by a subset S of a representation of GL_p.)

A fundamental result which is due to Weyl says that we have equality in (2) as soon as $r \ge \dim V$ ([Wey46] II.5 Theorem 2.5A). This result is sometimes refered to by saying that *"one can see all invariants already on $n = \dim V$ copies"*.

Theorem A (H. Weyl)**.** *For every $p \ge n := \dim V$ we have $K[V^p]^G = \langle K[V^n]^G \rangle_{\mathrm{GL}_p}$. In particular, if $S \subset K[V^n]^G$ is a system of generators then $\langle S \rangle_{\mathrm{GL}_p}$ generates the invariant ring $K[V^p]^G$.*

The classical proof of this result is based on the Capelli-Deruyts expansion. We will discuss it in the next paragraph (see 8.2). Here we use the representation theory of GL_p to obtain a more direct proof based on our results from §5 about highest weights of GL_p-modules. (See Corollary 1 of the following section.)

Exercises 1. Verify the theorem for the standard (1-dimensional) representation of the finite cyclic group $\mu_m := \{\zeta \in K^* \mid \zeta^m = 1\} \subset \mathrm{GL}_1$ of order m.

2. Determine the invariants of several copies of the standard 2-dimensional

representation of $T := \left\{ \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \,\middle|\, t \in K^* \right\} \subset \mathrm{GL}_2$.
(Cf. 1.2 Exercise 6.)

3. Let V be an n-dimensional representation of G and assume that $K[V^n]^G$ is generated by the invariants of degree $\leq N$ for some $N > 0$. Then this holds for the invariants of any number of copies of V.

7.2 U_p-invariants and primary covariants. We use the notation from the previous section.

Proposition. *Assume $p \geq n := \dim V$ and let $M \subset K[V^p]$ be a GL_p-stable subspace. Then, as a GL_p-module, M is generated by its intersection with $K[V^n]$:*

$$M = \langle M \cap K[V^n] \rangle_{\mathrm{GL}_p}.$$

Before giving the proof of the proposition we want to draw some consequences. Part (a) of the following corollary is WEYL's Theorem A of the previous section 7.1.

Corollary 1. *Assume that the subspace $F \subset K[V^n]^G$ generates $K[V^n]^G$ as an algebra.*

(a) *For $p \geq n$ the invariants $K[V^p]^G$ are generated by $\langle F \rangle_{\mathrm{GL}_p}$.*
(b) *For $p \leq n$ the invariants $K[V^p]^G$ are generated by the restrictions $\operatorname{res} F := \{f|_{V^p} \mid f \in F\}$. Moreover, $\operatorname{res} F = F \cap K[V^p]$ in case F is GL_n-stable.*

PROOF: (a) Let $p \geq n$ and denote by $R \subset K[V^p]$ the subalgebra generated by $\langle F \rangle_{\mathrm{GL}_p}$. Then $R \subset K[V^p]^G$, and R is GL_p-stable. Since $R \supset K[F] = K[V^n]^G = K[V^p]^G \cap K[V^n]$ the proposition above implies that $R = K[V^p]^G$.
(b) For $p \leq n$ consider the restriction map $\operatorname{res}\colon K[V^n] \to K[V^p]$, $f \mapsto f|_{V^p}$ which is G-equivariant and induces the identity on $K[V^p] \subset K[V^n]$. Hence, the composition

$$K[V^p]^G \hookrightarrow K[V^n]^G \xrightarrow{\text{res}} K[V^p]^G$$

is the identity, too, and so $K[V^n]^G \to K[V^p]^G$ is surjective. This shows that $\operatorname{res}(F)$ generates $K[V^p]^G$. Finally, the equation $\operatorname{res} F = F \cap K[V^p]$ certainly holds in case F is a multihomogeneous subspace of $K[V^n]$. Since F is GL_n-stable this follows from the next lemma. □

Lemma. *A GL_p-stable subspace $F \subset K[V^p]$ is multihomogeneous.*

PROOF: We give the argument for $p = 1$; the general case follows easily by induction. Let $f \in F \subset K[V]$ and let $f = f_0 + f_1 + \cdots + f_s$ be the decomposition into homogeneous components. For $t \in K^* = \mathrm{GL}_1$ we get $tf = f_0 + t \cdot f_1 + t^2 \cdot f_2 + \cdots + t^s \cdot f_s \in F$. Choosing $s+1$ different values $t_0, t_1, \dots, t_s \in K^*$, the linear system

$$t_i f = f_0 + t_i \cdot f_1 + \cdots + t_i{}^s \cdot f_s, \quad i = 0, 1, \dots, s$$

has an invertible coefficient matrix: Its determinant is a VANDERMONDE determinant. Hence $f_i \in \langle t_i f \mid i = 0, 1, \dots, s \rangle \subset F$. □

Now we come to the proof of the proposition above. It is a consequence of the theory of highest weights developed in 5.7.

PROOF OF THE PROPOSITION: Since the representation of GL_p on $K[V^p]$ is completely reducible (5.3) it suffices to show that for every simple submodule $M \subset K[V^p]$ we have $M \cap K[V^n] \neq 0$. In fact, this implies $\langle M \cap K[V^n] \rangle_{\mathrm{GL}_p} = M$ for all simple submodules M, hence for any submodule. We have seen in Proposition 5.7 that $M^{U_p} \neq 0$ where

$$U_p := \{ \begin{pmatrix} 1 & * & * \\ & \ddots & * \\ & & 1 \end{pmatrix} \in \mathrm{GL}_p \}$$

is the subgroup of the upper triangular unipotent matrices. So it remains to prove that $K[V^p]^{U_p} \subseteq K[V^n]$. By induction it suffices to show that for $p > n$ a U_p-invariant function $f \in K[V^p]$ does not depend on the last variable v_p. Consider an element

$$v = (v_1, v_2, \dots, v_n, v_{n+1}, \dots, v_p) \in V^p$$

where $v_1, v_2, \dots, v_n$ are linearly independent. Then there exist $\alpha_1, \dots, \alpha_{p-1} \in K$ such that $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{p-1} v_{p-1} + v_p = 0$. Putting

$$u = \begin{pmatrix} 1 & 0 & \cdots & \alpha_1 \\ & 1 & \cdots & \alpha_2 \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix} \in U_p,$$

we get $u(v_1, \dots, v_{p-1}, 0) = (v_1, \dots, v_{p-1}, v_p)$. Since the elements v as above form a ZARISKI-dense subset of V^p this shows that a U_p-invariant function does not depend on the last variable v_p. □

From the proof we get the following corollary:

Corollary 2. *For any $p \geq n$ we have $K[V^p]^{U_p} \subset K[V^n]$.*

The subalgebra $\mathrm{PC} := K[V^n]^{U_n}$ appears in the classical literature as the *algebra of primary covariants*. We will discuss this in more details in 8.1 (see also 7.7 Example 2).

Remark. Another proof of the proposition above follows from the CAPELLI-DERUYTS expansion (Theorem 8.1, see 8.2). This expansion can be regarded as an explicit way to write a function $f \in K[V^p]$ in the form $f = \sum A_i f_i$ where the f_i belong to the primary covariants and the A_i are certain linear operators preserving GL_p-stable subspaces.

7.3 Polarization operators. We want to give another description of the GL_p-submodule $\langle S \rangle_{\mathrm{GL}_p}$ generated by a subset $S \subset K[V^p]$ using certain *differential operators*. In modern terms, we study the action of the *Lie algebra* of GL_p on the coordinate ring $K[V^p]$ by derivations and of its *envelopping algebra* by differential operators.

First we introduce the following linear operators Δ_{ij} on $K[V^p]$ where $1 \leq i, j \leq p$:

$$\Delta_{ij} f\,(v_1, \ldots, v_p) := \left. \frac{f(v_1, \ldots, v_j + tv_i, \ldots, v_p) - f(v_1, \ldots, v_p)}{t} \right|_{t=0}.$$

They are called *polarization operators*. Using coordinates in V and putting $v_i = (x_1^{(i)}, \ldots, x_n^{(i)})$ we see that

$$\Delta_{ij} = \sum_{\nu=1}^{n} x_\nu^{(i)} \cdot \frac{\partial}{\partial x_\nu^{(j)}}.$$

Clearly, Δ_{ij} is a *derivation* of the algebra $K[V^p]$, i.e. $\Delta_{ij}(fh) = f\,\Delta_{ij} h + h\,\Delta_{ij} f$ for $f, h \in K[V^p]$. Or, in more geometric terms, Δ_{ij} is the *vector field* on V^p given by

$$(\Delta_{ij})_{(v_1, \ldots, v_p)} = (0, \ldots, 0, \underset{j}{v_i}, 0, \ldots, 0).$$

The next lemma explains the meaning of these operators with respect to the TAYLOR-expansion.

Lemma. *For $f \in K[V^p]$ and $t \in K$ we get*

$$f(v_1, \ldots, v_j + tv_i, \ldots, v_p) = \sum_{\nu=0}^{\infty} \frac{t^\nu}{\nu!} \cdot \Delta_{ij}^\nu f\,(v_1, \ldots, v_p).$$

PROOF: The coefficients in the usual TAYLOR-expansion with respect to t are

$$f_\nu(x_1, \dots, x_p) := \left(\frac{\partial^\nu}{\partial t^\nu} f(v_1, \dots, v_j + tv_i, \dots, v_p) \right)_{t=0}.$$

But it is easy to see that

$$\left(\frac{\partial^{\nu+1}}{\partial t^{\nu+1}} f(\dots, v_j + tv_i, \dots) \right)_{t=0} = \left(\frac{\partial}{\partial t} f_\nu(\dots, v_j + tv_i, \dots) \right)_{t=0}.$$

Hence, by induction, $f_\nu = \Delta_{ij}^\nu f$. □

Examples. (a) If v_j does not occur in f then $\Delta_{ij} f = 0$.

(b) Assume that f is linear in v_j. Then

$$\Delta_{ij} f(v_1, \dots, v_j, \dots, v_p) = f(v_1, \dots, \underset{j}{v_i}, \dots, v_p),$$

i.e., v_j is replaced by v_i.

(c) The full polarization $\mathcal{P}f$ of a homogeneous polynomial $f \in K[V]$ of degree d which was defined in 4.4 has the following description:

$$\mathcal{P}f = \Delta_{d0} \cdots \Delta_{20} \Delta_{10} f.$$

Here we work in the coordinate ring $K[V^{d+1}]$ where $d = \deg f$ and use the notation $(v_0, v_1, v_2, \dots, v_d)$ for elements in V^{d+1}, assuming that $f = f(v_0)$ is a function depending only on the first copy of V^{d+1}.

7.4 Differential operators. Let W be a finite dimensional K-vector space. The algebra of linear operators on $K[W]$ generated by the *derivations*

$$\frac{\partial}{\partial w} : f(u) \mapsto \frac{f(u + tw) - f(u)}{t} |_{t=0}$$

and the multiplication with functions $f \in K[W]$ is called the *ring of differential operators* on W. We will denote it by $\mathcal{D}(W)$ and consider it as a subalgebra of $\mathrm{End}_K(K[W])$. Using coordinates in W it is easy to see that every $D \in \mathcal{D}(W)$ can be (uniquely) written in the form

$$D = \sum f_\alpha \frac{\partial^\alpha}{\partial x^\alpha} \quad \text{where } f_\alpha \in K[W] \text{ and}$$

$$\frac{\partial^\alpha}{\partial x^\alpha} := \frac{\partial^{\alpha_1 + \alpha_2 + \cdots + \alpha_m}}{\partial {x_1}^{\alpha_1} \partial {x_2}^{\alpha_2} \cdots \partial {x_m}^{\alpha_m}}.$$

The linear operators of the form $\sum f_i \frac{\partial}{\partial x_i}$ are the *polynomial vector fields* on W, i.e., the derivations of $K[W]$.

Definition. Let $\mathcal{U} = \mathcal{U}(p) \subset \mathcal{D}(V^p)$ be the subalgebra of differential operators on V^p generated by the polarization operators Δ_{ij}, $i, j = 1, \dots, p$.

Proposition. *Let $M \subset K[V^p]$ be a linear subspace. Then M is stable under* GL_p *if and only if it is stable under* $\mathcal{U}(p)$. *If M is finite dimensional and* GL_p*-stable then the subalgebra of* $\mathrm{End}(M)$ *generated by* GL_p *coincides with the image of* $\mathcal{U}(p)$ *in* $\mathrm{End}(M)$.

In other words, a subspace $M \subset K[V^p]$ is GL_p-stable if and only if it is *stable under polarization*, i.e., if $f \in M$ implies $\Delta_{ij} f \in M$ for all Δ_{ij}. Furthermore, the GL_p-span $\langle S\rangle_{\mathrm{GL}_p}$ of a subset $S \subset K[V^p]$ can be obtained by applying successively polarization operators in all possible ways and forming the linear span:

$$\langle S\rangle_{\mathrm{GL}_p} = \langle \mathcal{U}(p)\, S\rangle.$$

We shortly say $\langle S\rangle_{\mathrm{GL}_p}$ *is obtained from S by polarization.* In these terms we can rephrase Weyl's Theorem A (7.1).

Corollary. *Let $G \subset \mathrm{GL}(V)$ be a subgroup and let $S \subset K[V^n]^G$ be a system of generators of the ring of invariants, $n := \dim V$. For every $p \geq n$ we get a system of generators for $K[V^p]^G$ by polarizing S.*

Proof of the Proposition: For $t \in K$ and $1 \leq i, j \leq p$ define as in 5.7

$$u_{ij}(t) := E + tE_{ij} \in \mathrm{GL}_p \quad (t \neq -1 \text{ in case } i = j).$$

It is well known that these matrices generate GL_p (see §5, Exercise 18). Furthermore, for $f \in K[V^p]$ we find

$$\begin{aligned} (u_{ij}(t)f)(v_1, \dots, v_p) &= f(v_1, \dots, v_j + tv_i, \dots, v_p) \\ &= \sum_\nu \frac{t^\nu}{\nu!} \Delta_{ij}^\nu f(v_1, \dots, v_p), \end{aligned} \tag{3}$$

by Lemma 7.3 above. Clearly, the sum is finite and the same argument as in the proof of Lemma 7.2 implies that

$$\langle u_{ij}(t) f \mid t \in K\rangle = \langle \Delta_{ij}^\nu f \mid \nu = 0, 1, \dots\rangle.$$

Thus we get the first claim. In addition, formula (3) shows that

$$u_{ij}(t) = \exp(t\Delta_{ij}) := \sum_{\nu \geq 0} \frac{t^\nu}{\nu!} \Delta_{ij}^\nu$$

as an operator on $K[V^p]$. (Note that for a given f we have $\Delta_{ij}^\nu f = 0$ for large ν.) If we restrict this operator to a finite dimensional GL_p-stable subspace M the right hand side becomes a finite sum and is therefore an element of $\mathcal{U}(p)$. Arguing again as in the proof of Lemma 7.2 we see that the finite set $\{\Delta_{ij}^\nu|_M\} \subset \mathrm{End}(M)$ lies in the span of the linear operators $u_{ij}(t)$, $t \in K$. □

Exercise 4. Prove that the commutator subgroup of $\mathrm{GL}_p(K)$, i.e., the subgroup generated by all commutators $(g,h) := ghg^{-1}h^{-1}$ $(g,h \in \mathrm{GL}_p)$, is equal to $\mathrm{SL}_p(K)$.

7.5 Invariants of unimodular groups. In case of a unimodular subgroup $G \subset \mathrm{SL}(V)$ we can improve Weyl's Theorem A (7.1). Let $n := \dim V$ and fix a basis of V. Then the determinant $\det(v_1, \dots, v_n)$ is defined for every n-tuple of vectors $v_i \in V = K^n$ as the determinant of the $n \times n$ matrix consisting of the column vectors $v_1, \dots v_n$. This allows to define, for every sequence $1 \le i_1 < i_2 < \dots < i_n \le p$, an $\mathrm{SL}(V)$-invariant function

$$[i_1, \dots, i_n] \colon V^p \to K, \quad (v_1, \dots, v_p) \mapsto \det(v_{i_1}, \dots, v_{i_n}).$$

The following result is again due to Weyl (cf. [Wey46] II.5 Theorem 2.5A). The proof will be given in 8.2.

Theorem B. *Assume that G is a subgroup of $\mathrm{SL}(V)$. For $p \ge n = \dim V$ the invariant ring $K[V^p]^G$ is generated by $\langle K[V^{n-1}]^G \rangle_{\mathrm{GL}_p}$ together with all determinants $[i_1, \dots, i_n]$.*

As before, this can be rephrased in the following way:

Corollary 1. *Let $S \subset K[V^{n-1}]^G$ be a system of generators. Then we get a generating system for $K[V^p]^G$ by polarizing S and adding all possible determinants $[i_1, \dots, i_n]$.*

Exercise 5. Show that a polarization of a determinant $[i_1, \dots, i_n]$ is either 0 or again a determinant.

As a consequence we obtain a preliminary version of the FFT for the special linear group. The general case will be given in the next paragraph (see Theorem 8.4).

Corollary 2 (Preliminary FFT for SL_n). *For $\mathrm{SL}(V)$ acting on several copies of V the invariant ring $K[V^p]^{\mathrm{SL}(V)}$ is generated by the determinants $[i_1, \dots, i_n]$. In particular, $K[V^p]^{\mathrm{SL}(V)} = K$ for $p < n$.*

Proof: The first statement follows from the second by the corollary above, and the second claim is clear since $\mathrm{SL}(V)$ has a Zariski-dense orbit in V^{n-1}. □

Exercises

6. Consider the subgroup $N := \left\{ \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix}, \begin{pmatrix} 0 & -t \\ t^{-1} & 0 \end{pmatrix} \middle| \, t \in K^* \right\}$ of SL_2 with its natural representation on $V = K^2$. Show that the invariant ring $K[V^p]^N$ is generated by the invariants $x_i^2 y_i^2$ and $x_i y_j - x_j y_i$ $(1 \le i < j \le p)$.

7. The invariants of any number of 2×2 matrices $(A_1, \dots, A_r) \in \mathrm{M}_2(K)^r$ under simultaneous conjugation are generated by the following functions: $\operatorname{Tr} A_i$, $\operatorname{Tr} A_i A_j$ and $\operatorname{Tr} A_i A_j A_k$.
(Hint: Since $\mathrm{M}_2 = K \oplus \mathrm{M}_2{}'$ where $\mathrm{M}_2{}'$ are the traceless 2×2 one can apply Theorem B for $n = 3$. Now use Example 2.4 and show that $[A, B, C] = \operatorname{Tr} ABC$ for $A = \left(\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right)$ corresponding to the column vector $(a, b, c)^t$).

7.6 The First Fundamental Theorem for $\mathcal{S}_n$. Consider the standard representation of the symmetric group $\mathcal{S}_n$ on $V = K^n$ by permutation:

$$\sigma e_i = e_{\sigma(i)} \quad \text{or} \quad \sigma(x_1, \dots, x_n) = (x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)}).$$

The invariant ring $K[x_1, \dots, x_n]^{\mathcal{S}_n}$ is the algebra of symmetric polynomials which is generated by the elementary symmetric functions $\sigma_1, \dots, \sigma_n$ (see 1.2 Example 3). It is natural to ask for a description of the invariants of several copies of the standard representation of $\mathcal{S}_n$. Again this question has been answered by H. Weyl (see [Wey46] II.A.3).

Theorem. *The invariants $K[V^p]^{\mathcal{S}_n}$ are generated by the polarizations of the elementary symmetric functions.*

Proof: In order not to overload notation let us denote the variables on the different copies of V by $x, y, \dots, z$. As in the proof of Proposition 1.2 we use induction on n and denote by $\sigma'_1, \dots, \sigma'_{n-1}$ the elementary symmetric functions of $n-1$ variables. Let $A \subset K[V^p]^{\mathcal{S}_n} = K[x_1, \dots, x_n, y_1, \dots, z_n]^{\mathcal{S}_n}$ be the subalgebra generated by the polarizations of the elementary symmetric functions σ_i and by $A' \subset A$ the subalgebra generated by the polarizations of the elementary symmetric functions σ'_i. Polarizing the relations

$$\begin{aligned} \sigma_1 &= \sigma'_1 + x_n, \\ \sigma_2 &= \sigma'_2 + x_n \sigma'_1, \\ &\vdots \\ \sigma_{n-1} &= \sigma'_{n-1} + x_n \sigma'_{n-2}, \\ \sigma_n &= x_n \sigma'_{n-1}, \end{aligned}$$

we find that the $A \subset A'[x_n, y_n, \dots, z_n]$.

Now let $f = f(x, y, \dots, z)$ be an invariant and write

$$f = \sum_{\alpha, \beta, \dots, \gamma} f_{\alpha, \beta, \dots, \gamma} x_n^\alpha y_n^\beta \cdots z_n^\gamma$$

where the $f_{\alpha, \beta, \dots, \gamma}$ do not depend on the last variables $x_n, y_n, \dots, z_n$. Clearly, the coefficients $f_{\alpha, \beta, \dots, \gamma}$ are invariants under $\mathcal{S}_{n-1}$ and therefore elements from

A', by induction. It follows that $f \in A'[x_n, y_n, \dots, z_n] \subset A[x_n, y_n, \dots, z_n]$. Since f is an invariant under $\mathcal{S}_n$ we have $f = \frac{1}{n!}\sum_{\sigma\in\mathcal{S}_n} \sigma f$. Therefore, it suffices to show that $\sum_i x_i^\alpha y_i^\beta \cdots z_i^\gamma \in A$ for all $\alpha, \beta, \dots, \gamma \geq 0$. But clearly, this sum is a polarization of the invariant $\sum_i x_i^{\alpha+\beta+\cdots+\gamma}$ which is contained in $K[\sigma_1, \dots, \sigma_n]$ and the claim follows. □

Remark. It was asked by several people whether this result holds for every finite reflection group, i.e., for every finite subgroup $G \subset \mathrm{GL}_n(\mathbb{C})$ which is generated by pseudo-reflections. (An element of finite order in $\mathrm{GL}_n(\mathbb{C})$ is called a *pseudo-reflection* if it fixes a hyperplane pointwise.) It is not difficult to see that the result holds for the Weyl-groups of type B_n and also for the dihedral groups (see Exercise 8 below). But WALLACH showed that it is false for the Weyl group of type D_4. We refer to the paper [Hun96] of HUNZIKER for a detailed discussion of this problem.

Exercise 8. Let $D_{2n} \subset \mathrm{GL}_2(\mathbb{C})$ be the dihedral group of order $2n$:

$$D_{2n} := \left\{ \begin{pmatrix} \zeta & 0 \\ 0 & \zeta^{-1} \end{pmatrix} \,\middle|\, \zeta^n = 1 \right\} \cup \left\{ \begin{pmatrix} 0 & \zeta \\ \zeta^{-1} & 0 \end{pmatrix} \,\middle|\, \zeta^n = 1 \right\},$$

i.e., $D_{2n} = C_n \cup \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \cdot C_n$. Then $K[x,y]^{D_{2n}} = K[x^n y^n]$ and the invariants of any number of copies of K^2 are generated by the polarizations of $x^n y^n$.

7.7 Multiplicity free representations. Let G be a product of general linear and special linear groups or, more generally, a linear group as in 5.4. Assume that G acts on a (commutative) K-algebra A by means of algebra automorphisms

$$\rho\colon G \to \mathrm{Aut}_{\mathrm{alg}}\, A \subset \mathrm{GL}(A)$$

and that this representation is locally finite and rational (5.2). Such an algebra will be called a *G-algebra*. Typical examples are the coordinate rings of rational representations of G (Lemma 5.4).

Exercise 9. Let H be an arbitrary group acting on an algebra A by means of algebra automorphisms. Assume that every finite dimensional subrepresentation of A is completely reducible. Then $A = \bigoplus_\lambda A_\lambda$ where A_λ is the isotypic component of type λ, i.e., the sum of all simple submodules of isomorphism class λ. Moreover, $A_0 = A^H$ is a subalgebra and each A_λ is an A^H-submodule.

Definition. A G-algebra is called *multiplicity free* if A is a direct sum of simple G-modules which are pairwise non-isomorphic. A representation W of G is called multiplicity free if the coordinate ring $K[W]$ is multiplicity free.

Multiplicity free algebras appear in the literature at several places, see e.g. [Had66], [ViP72], [Kra85]. Recently, Howe used it to deduce many results from classical invariant theory in a systematic way [How95].

Examples. (a) The symmetric algebra $\mathrm{S}(V) = \bigoplus_i \mathrm{S}^i V$ is multiplicity free with respect to $\mathrm{GL}(V)$ and $\mathrm{SL}(V)$. The same holds for the coordinate ring $K[V] = S(V^*)$.

(b) The Cauchy formula (6.6 Corollary 2) shows that $\mathrm{S}(V \otimes W)$ and $K[V \otimes W]$ are multiplicity free with respect to $\mathrm{GL}(V) \times \mathrm{GL}(W)$. For the group $\mathrm{SL}(V) \times \mathrm{SL}(W)$ it is multiplicity free only in case $\dim V \neq \dim W$. In fact, for $V = W$ there are non-constant invariants.

(c) The exterior algebra $\bigwedge V$ is multiplicity free for $\mathrm{GL}(V)$.

From now on assume that G is a product $\mathrm{GL}_{p_1} \times \mathrm{GL}_{p_2} \times \cdots \times \mathrm{SL}_{q_1} \times \mathrm{SL}_{q_2} \times \cdots$ and denote by Λ_G the monoid of highest weights. Any G-algebra A has an isotypic decomposition

$$A = \bigoplus_{\lambda \in \Lambda_G} A_\lambda,$$

see Exercise 9 above. Define

$$\Omega_A := \{\lambda \in \wedge_G \mid A_\lambda \neq 0\}.$$

Lemma. *Assume that A is a domain.*

(a) *Ω_A is a submonoid, i.e., $\Omega_A + \Omega_A = \Omega_A$.*

(b) *Suppose that A is multiplicity free. If Ω_A is generated by $\lambda_1, \ldots, \lambda_s$ then A is generated as an algebra by $A_{\lambda_1} + \cdots + A_{\lambda_s}$. In particular, A is finitely generated.*

Proof: (a) For any $\lambda \in \Omega_A$ denote by $a_\lambda \in A_\lambda$ a highest weight vector. If $\lambda, \mu \in \Omega_A$ then $a_\lambda a_\mu$ is a vector $\neq 0$ of maximal weight in $A_\lambda A_\mu \subset A$ and so its weight $\lambda + \mu$ occurs in Ω_A.

(b) Let $a_i \in A_{\lambda_i}$ be highest weight vectors, $i = 1, \ldots, s$. If $\lambda \in \Omega_A$, $\lambda = \sum_{i=1}^s m_i \lambda_i$ then $a := a_1^{m_1} a_2^{m_2} \cdots a_s^{m_s}$ is a highest weight vector of weight λ and so $A_\lambda = \langle a \rangle_{\mathrm{GL}_n} \subset A_{\lambda_1}^{m_1} A_{\lambda_2}^{m_2} \cdots A_{\lambda_s}^{m_s}$. □

7.8 Applications and Examples. We regard V^p as a $\mathrm{GL}(V) \times \mathrm{GL}_p$-module as in 7.1. For $k \leq m := \min(p, \dim V)$ we have a $\mathrm{GL}(V)$-equivariant multilinear map

$$\varphi_k \colon V^p \to \bigwedge^k V, \quad (v_1, v_2, \ldots, v_p) \mapsto v_1 \wedge v_2 \wedge \cdots \wedge v_k.$$

Thus the simple module $(\bigwedge^k V)^* = \bigwedge^k V^*$ occurs in $K[V^p]$ in degree k. Moreover, φ_k is also equivariant with respect to the action of $U_p \subset \mathrm{GL}_p$ which is trivial on $\bigwedge^k V$. Thus we find $(\bigwedge^k V)^* = \bigwedge^k V^* \subset K[V^p]_k^{U_p}$.

We claim that the algebra $K[V^p]^{U_p}$ is generated by $V^*, \bigwedge^2 V^*, \ldots, \bigwedge^m V^*$ where $m := \min(p, \dim V)$.

Proof: The Cauchy formula (6.6 Corollary 2) tells us that

$$K[V^p] = \mathrm{S}(V^* \otimes K^p) = \bigoplus_\lambda L_\lambda(V^*) \otimes L_\lambda(p).$$

where λ runs through the dominant weights of height $\leq m$. Therefore, the algebra

$$A := K[V^p]^{U_p} = \bigoplus_\lambda L_\lambda(V^*) \otimes L_\lambda(p)^{U_p} \simeq \bigoplus_\lambda L_\lambda(V^*)$$

is multiplicity free, and $\Omega_A = \{\lambda = (p_1, \ldots, p_m)\}$ is the monoid generated by $\omega_1, \ldots, \omega_m$. Clearly, $L_{\omega_k}(V^*)$ corresponds to the submodule $\bigwedge^k V^*$ of A constructed above, and the claim follows by Lemma 1 above. □

Choosing a basis $(e_1, \ldots, e_n)$ in V we can identify the elements of V^p with $n \times p$ matrices

$$X = \begin{pmatrix} x_1^{(1)} & x_1^{(2)} & \cdots & x_1^{(p)} \\ x_2^{(1)} & x_2^{(2)} & \cdots & x_2^{(p)} \\ \vdots & \vdots & & \vdots \\ x_n^{(1)} & x_n^{(2)} & \cdots & x_n^{(p)} \end{pmatrix}.$$

Then the map $\varphi_k \colon V^p \to \bigwedge^k V$ is given by

$$\varphi_k(X) = \sum_{r_1 < \cdots < r_k} \det \begin{pmatrix} x_{r_1}^{(1)} & \cdots & x_{r_1}^{(k)} \\ \vdots & & \vdots \\ x_{r_k}^{(1)} & \cdots & x_{r_k}^{(k)} \end{pmatrix} e_{r_1} \wedge e_{r_2} \wedge \cdots \wedge e_{r_k}.$$

Thus we have proved the following result.

Proposition. *The invariant ring $K[V^p]^{U_p}$ is generated by the subspaces V^*, $\bigwedge^2 V^*, \ldots, \bigwedge^m V^*$, $m := \min(p, \dim V)$. Identifying V^p with the $n \times p$ matrices as above an explicit system of generators is given by the $k \times k$ minors extracted from the first k columns of the matrix X for $k = 1, \ldots, m$. In particular, we have $K[V^p]^{U_p} = K[V^n]^{U_n}$ for $p \geq n = \dim V$.*

The algebra $K[V^n]^{U_n}$ is classically called the algebra of *primary covariants* (see 7.2. We will discuss it in more details in the next paragraph (see 8.1).

Next we want to give an easy criterion for a representation to be multiplicity free which is due to Servedio [Ser73]. We denote by T and U the subgroups of G of diagonal and upper triangular unipotent matrices:

$$\begin{aligned} T &= T_{p_1} \times T_{p_2} \times \cdots \times T'_{q_1} \times T'_{q_2} \times \cdots \\ U &= U_{p_1} \times U_{p_2} \times \cdots \times U_{q_1} \times U_{q_2} \times \cdots . \end{aligned}$$

T normalizes U, and $B := T \cdot U = U \cdot T$ is the subgroup of upper triangular matrices, usually called Borel *subgroup* of G. There is a canonical projection $p\colon B \to T$ with kernel U which induces an inclusion $p^*\colon K[T] \hookrightarrow K[B]$ of regular functions in the usual way. It follows that the restriction $\chi \mapsto \chi|_T$ defines an isomorphism of character groups which allows us to indentify $\mathcal{X}(B)$ with $\mathcal{X}(T)$. Moreover, p^* induces an isomorphism

$$p^*\colon K[T] \xrightarrow{\sim} K[B]^U \subset K[B].$$

The proofs of these statements are easy (see the following exercise 10).

Exercises

10.

11. A G-algebra is multiplicity free if and only if all weight spaces of the invariant algebra A^U are one-dimensional.

Lemma. *Let W be a representation of G and assume that there is a vector $w_0 \in W$ such that the B-orbit $Bw_0 = \{bw_0 \mid b \in B\}$ is* Zariski*-dense in W. Then*

(a) *$K[W]$ is multiplicity free.*

(b) *If $K[W]$ contains a simple G-module of highest weight λ then the character λ vanishes on the stabilizer $B_{w_0} := \{b \in B \mid bw_0 = w_0\}$.*

Proof: The orbit map $B \to W$, $b \mapsto bw_0$, induces an inclusion $K[W] \hookrightarrow K[B]$, $f \mapsto \tilde{f}$ where $\tilde{f}(b) := f(bw_0)$. This is a B-homomorphim where B acts on B by left multiplication. Thus

$$K[W]^U \hookrightarrow K[B]^U \simeq K[T].$$

and this map is T-equivariant. It follows that all weight spaces of $K[W]^U$ are one-dimensional and so $K[W]$ is multiplicity free (Exercise 11).

If $f \in K[W]^U$ is a highest weight vector of weight λ and $b \in B_{w_0}$ then $f(w_0) \neq 0$ and $f(w_0) = f(bw_0) = \lambda(b^{-1})f(w_0)$ and so $\lambda(b) = 1$ □

Examples. (a) The standard representations of GL_n and SL_n on $V = K^n$ admit a dense B_n-orbit $B_n e_n = V \setminus \{0\}$ and $(B_n)_{e_n} = B_{n-1}$ embedded in the obvious way. Thus we see again that $K[V]$ is multiplicity free and that every highest weight λ of $K[V]$ must be a multiple of $-\varepsilon_n$: $K[V] = \mathrm{S}(V^*) = \bigoplus_j \mathrm{S}^j V^* \simeq \bigoplus_j L_{-j\varepsilon_n}(V)$. For SL_n we have $-\varepsilon_n = \varepsilon_1 + \cdots \varepsilon_{n-1}$ which means that $V^* = \bigwedge^{n-1} V$ (see 5.2 Exercises 6 and 7).

(b) Consider M_n as a $\mathrm{GL}_n \times \mathrm{GL}_n$-module where $(g,h)A = gAh^t$, i.e., we identify M_n with $V \otimes V$. The open cell $B_n B_n^t$ is the $B = B_n \times B_n$-orbit of E and the stabilizer is given by $B_E = \{(t, t^{-1}) \mid t \in T_n\}$. It follows that $K[\mathrm{M}_n]$ is multiplicity free and if (λ, μ) occurs as a highest weight then $\lambda(t) + \mu(t^{-1}) = 0$ for all $t \in T_n$. i.e., $\lambda = \mu$. Hence

$$K[V \otimes V] = \mathrm{S}(V^* \otimes V^*) \simeq \bigoplus_\lambda L_\lambda(V^*) \otimes L_\lambda(V^*).$$

Similarly on finds $K[\mathrm{End}(V)] \simeq \bigoplus_\lambda L_\lambda(V^*) \otimes L_\lambda(V)$.

(c) With the same considerations we obtain a short proof of the Cauchy formula (6.6 Corollary 2). Consider $V \otimes W$ as a representation $\mathrm{GL}(V) \times \mathrm{GL}(W)$-module. As above, one sees that there is a dense B-orbit, namely the orbit of the element $u := e_1 \otimes f_1 + \cdots e_m \otimes f_m$ where $(e_i)_i$ and $(f_j)_j$ are bases of V and W and $m := \min(\dim V, \dim W)$. The stabilizer is $\{(t, t^{-1}) \mid t \in T_m\}$. As above we see that the possible irreducible components of $A := \mathrm{S}(V \otimes W)$ are given by $L_\lambda(V) \otimes L_\lambda(W)$ where $\operatorname{ht} \lambda \le m$. They all occur since the generators of the corresponding monoid Ω_A occur. Thus we have

$$\mathrm{S}(V \otimes W) = \bigoplus_{\lambda \text{ dominant of height } \le m} L_\lambda(V) \otimes L_\lambda(W).$$

We remark that this completes Example 2, independent of earlier considerations.

(d) Consider the representation $\mathrm{S}^2 V$ of $\mathrm{GL}(V)$. Choosing a basis we can identify $\mathrm{S}^2 V$ with the symmetric 2×2 matrices with the GL_n-action $A \mapsto gAg^t$. Again the B_n-orbit of E is Zariski-dense and its stabilizer is $\left\{ \begin{pmatrix} \pm 1 & & \\ & \ddots & \\ & & \pm 1 \end{pmatrix} \right\}$. Thus $\mathrm{S}(\mathrm{S}^2 V)$ is multiplicity free and if $\lambda = \sum_{i=1}^n p_i \varepsilon_i$ occurs as a highest weight then λ is *even*, i.e., all p_i are even. The monoid of even dominant weights of height $\le n = \dim V$ is generated by $2\omega_1, \ldots, 2\omega_n$. It is easy to see that the representations $L_{2\omega_i}(V)$ all occur in $\mathrm{S}(\mathrm{S}^2 V)$. Hence

$$\mathrm{S}(\mathrm{S}^2 V) = \bigoplus_{\lambda \text{ even ht } \lambda \leq n} L_\lambda(V).$$

In particular we have the following *plethysm*:

$$\mathrm{S}^m(\mathrm{S}^2 V) = \bigoplus_{\substack{\lambda \text{ even} \\ |\lambda| = m, \ \mathrm{ht}\, \lambda \leq n}} L_\lambda(V)$$

§ 8 Capelli-Deruyts Expansion

8.1 Primary covariants. We fix a bases of V and identify V^p with the $n \times p$-matrices:

$$V^p = \mathrm{M}_{n\times p}(K) \; : \; (v_1, \dots , v_p) \longleftrightarrow \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix}.$$

Hence $K[V^p] = K[x_{ij} \mid i = 1, \dots , n, \; j = 1, \dots , p]$.

Definition. Let $\mathrm{PC} = \mathrm{PC}_n \subset K[V^n]$ be the subalgebra generated by all $k \times k$-minors extracted from the first k columns of the matrix $(x_{ij})_{i,j=1,\dots,n}$ (7.8). This subalgebra PC_n is classically called the algebra of *primary covariants*.

For $n = 2$ we have

$$\mathrm{PC} = K[x_{11}, x_{21}, x_{11}x_{22} - x_{21}x_{12}]$$

and in general

$$\mathrm{PC} = K[x_{11}, \dots , x_{n1}, x_{11}x_{22} - x_{21}x_{12}, \dots , x_{n-11}x_{n2} - x_{n1}x_{n-12}, \dots \\ \dots , \det(x_{ij})^n_{i,j=1}].$$

Remark. It follows from the definition that for any $p \geq n$ the primary covariants $\mathrm{PC} \subset K[V^p]$ are invariant under the subgroup

$$U_p := \{ \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in \mathrm{GL}_p \}$$

of GL_p which acts by right multiplication on V^p as in 7.1. In fact, we have already seen in Proposition 7.8 that $\mathrm{PC}_n = K[V^p]^{U_p}$. We will discuss this again in §9.

The fundamental result here is the following Capelli-Deruyts-*expansion* which is also called Gordan-Capelli-*expansion*.

Theorem. *For every multihomogeneous $f \in K[V^p]$ there are linear operators A_i, $B_i \in \mathcal{U}(p)$ such that*

$$f = \sum_i A_i B_i f \quad \text{and} \quad B_i f \in \mathrm{PC} \qquad \text{for all } i.$$

The operators A_i and B_i only depend on the multidegree of f.

(Since f is multihomogeneous and PC_n a homogeneous subalgebra one can always assume if necessary that the elements $B_i f$ are multihomogeneous, too.)

This theorem will follow from CAPELLI*'s identity* 9.3, see 9.5. Here we first draw some consequences.

8.2 Proof of WEYL's Theorems. We will show now that the two Theorems of WEYL (7.1 and 7.5) follow easily from the expansion-formula above. Recall that the operators A_i, $B_i \in \mathcal{U}(p)$ stabilize every GL_p-stable subset $U \subset K[V^p]$ (Proposition 7.4). Hence for every $f \in U$ we get

$$B_i f \in U \cap \mathrm{PC}_n \subset U \cap K[V^n]$$

and therefore

$$f = \sum A_i B_i f \in \langle U \cap K[V^n] \rangle_{\mathrm{GL}_p}$$

which proves Theorem A of 7.1.

For the proof of Theorem B (7.5) we can assume that $B_i f$ is multihomogeneous. Since the determinant $[1, \ldots, n]$ is the only generator of PC_n which contains the variables x_{in} of the last copy of V in V^n we have $B_i f = h \cdot [1, \ldots, n]^d$ with $h \in \mathrm{PC}_n \cap K[V^{n-1}]$ and $d \in \mathbb{N}$. Now

$$\Delta_{ij}(h \cdot [1, \ldots, n]^d) = \\ = \Delta_{ij} h \cdot [1, \ldots, n]^d + dh \cdot [1, \ldots, n]^{d-1} \cdot \Delta_{ij}[1, \ldots, n].$$

Since $\Delta_{ij}[1, \ldots, n]$ is again a determinant we obtain

$$A_i B_i f \in K[\langle K[V^{n-1}]^G \rangle_{\mathrm{GL}_p}, [i_1, \ldots, i_n]],$$

and the claim follows.

Remark. In these proofs we have not used the full strength of the CAPELLI-DERUYTS-expansion (cf. the remarks in 9.4).

8.3 A generalization of WEYL's Theorems. Let $V_1, \ldots, V_r$ be (finite dimensional) representations of a group G and consider the direct sum $V^{\mathbf{p}} := {V_1}^{p_1} \oplus {V_2}^{p_2} \oplus \cdots \oplus {V_r}^{p_r}$. There is an obvious action of

$$\mathrm{GL}_{\mathbf{p}} := \mathrm{GL}_{p_1} \times \mathrm{GL}_{p_2} \times \cdots \times \mathrm{GL}_{p_r}$$

on $V^{\mathbf{p}}$ commuting with G. Furthermore

$$K[V^{\mathbf{p}}] := K[{V_1}^{p_1} \oplus \cdots \oplus {V_r}^{p_r}] = K[{V_1}^{p_1}] \otimes K[{V_2}^{p_2}] \otimes \cdots \otimes K[{V_r}^{p_r}].$$

Again we have a G-equivariant inclusion

$$K[{V_1}^{p'_1} \oplus \cdots \oplus {V_r}^{p'_r}] \hookrightarrow K[{V_1}^{p_1} \oplus \cdots \oplus {V_r}^{p_r}]$$

in case $p_i' \le p_i$ for all i which is also equivariant with respect to the inclusion

$$\mathrm{GL}_{p_1'} \times \cdots \mathrm{GL}_{p_r'} \hookrightarrow \mathrm{GL}_{p_1} \times \cdots \times \mathrm{GL}_{p_r}.$$

Theorem A. *Assume $p_i \ge n_i := \dim V_i$ for all i and let $U \subset K[{V_1}^{p_1} \oplus \cdots \oplus {V_r}^{p_r}]$ be a $\mathrm{GL}_{p_1} \times \cdots \times \mathrm{GL}_{p_r}$-stable subset. Then*

$$U = \langle U \cap K[{V_1}^{n_1} \oplus \cdots \oplus {V_r}^{n_r}] \rangle_{\mathrm{GL}_{p_1} \times \cdots \times \mathrm{GL}_{p_r}}.$$

Theorem B. *Assume that all representations V_i of G are unimodular (i.e. the image of G is contained in $\mathrm{SL}(V_i)$ for all i). Then the ring of invariants $K[{V_1}^{p_1} \oplus \cdots \oplus {V_r}^{p_r}]^G$ is generated by $\langle K[{V_1}^{n_1} \oplus \cdots \oplus {V_r}^{n_r}]^G \rangle_{\mathrm{GL}_{p_1} \times \cdots \times \mathrm{GL}_{p_r}}$ and all determinants $[i_1, \ldots, i_{n_j}]_j$, $j = 1, \ldots, r$.*

(Here $[\ldots]_j$ denotes the determinant extracted from the p_j copies of V_j in $V^{\mathbf{p}}$.)

The proofs are the same as before, once we have established the following result, generalizing Theorem 8.1:

Theorem 3. *For every multihomogeneous $f \in K[{V_1}^{p_1} \oplus \cdots \oplus {V_r}^{p_r}]$ there are operators A_i, $B_i \in \mathcal{U}(p_1, \ldots, p_r)$ such that*

$$f = \sum_i A_i B_i f \quad \textit{and} \quad B_i f \in \ \mathrm{PC} \subset K[{V_1}^{n_1} \oplus \cdots \oplus {V_r}^{n_r}].$$

(Here $\mathcal{U}(p_1, \ldots, p_r)$ is the subalgebra of $\mathrm{End}(K[{V_1}^{p_1} \oplus \cdots \oplus {V_r}^{p_r}])$ generated by the polarization operators $\Delta_{ij}^{(\nu)} \in \mathrm{End}(K[V_\nu)^{p_\nu}])$ and PC is the tensor product of the $\mathrm{PC}_\nu \subset K[{V_\nu}^{n_\nu}]$, i.e., PC is generated by all $k \times k$-minors extracted from the first k rows in each block ${V_\nu}^{n_\nu}$ of $V^{\mathbf{n}}$.)

PROOF: By Theorem 8.1 we have for each $\nu = 1, \ldots, r$ operators $A_i^{(\nu)}$, $B_i^{(\nu)} \in \mathcal{U}(p_\nu)$ such that

$$f = \sum_i A_i^{(\nu)} B_i^{(\nu)} f \quad \text{and} \quad B_i^{(\nu)} f \in K[{V_1}^{p_1}] \otimes \cdots \overset{\nu}{\mathrm{PC}_\nu} \cdots \otimes K[{V_r}^{p_r}].$$

(The operators only affect the variables in ${V_\nu}^{p_\nu}$.) Clearly the operators ${A_i}^\nu$, ${B_i}^\nu$ commute with the operators ${A_i}^{\nu'}$, ${B_i}^{\nu'}$ for $\nu' \ne \nu$. Hence we find

$$f = \sum_i A_i B_i f$$

where each A_i is of the form $A_{j_1}^{(1)} A_{j_2}^{(2)} \cdots A_{j_r}^{(r)}$ and similarly for B_i. But then

$$B_i f \in \mathrm{PC}_1 \otimes \cdots \otimes \mathrm{PC}_r = \mathrm{PC}$$

and the claim follows. □

As in the previous paragraph 7 we can draw the following consequences from the theorems above:

Corollary 1. *Let $V_1, V_2, \dots, V_r$ be representations of a group G, $n_i := \dim V_i$, and let $S \subset K[{v_1}^{n_1} \oplus \cdots \oplus {V_r}^{n_r}]^G$ be a generating set. Then for every representation V of the form ${V_1}^{p_1} \oplus \cdots \oplus {V_r}^{p_r}$ we get a generating set for the ring of invariants $K[V]^G$ by polarizing S.*

Corollary 2. *If $K[{V_1}^{n_1} \oplus \cdots \oplus {V_r}^{n_r}]^G$ is finitely generated resp. generated by invariants of degree $\leq N$, then this holds for the invariant ring of any representation of the form ${V_1}^{p_1} \oplus \cdots \oplus {V_r}^{p_r}$.*

(This is clear since polarizing a homogeneous function does not change the degree.)

Corollary 3. *Let $V_1, \dots, V_r$ be unimodular representations of a group G, $\dim V_i = n_i$, and let $S \subset K[{V_1}^{n_1-1} \oplus \cdots \oplus {V_r}^{n_r-1}]^G$ be a generating set. Then the ring of invariants of any representation of the form ${V_1}^{p_1} \oplus \cdots \oplus {V_r}^{p_r}$ is generated by the polarization of S and all possible determinats $[i_1, \dots, i_{n_\nu}]_\nu$, $\nu = 1, \dots, r$.*

8.4 The First Fundamental Theorem for SL_n. As an application of the methods developed so far we give another proof of the First Fundamental Theorem for SL_n and obtain a new proof of the FFT for GL_n, which is more direct than the first one.

Theorem (FFT for SL_n). *The ring of invariants $K[V^p \oplus V^{*q}]^{\mathrm{SL}(V)}$ is generated by the scalar products $\langle j \mid i \rangle$ and the determinants $[i_1, \dots, i_n]$ and $[j_1, \dots, j_n]_*$.*

(In order to define the determinants we assume that some bases of V has been fixed and we choose in V^* the dual bases. Then $[i_1, \dots, i_n](v, \varphi) := \det(v_{i_1}, \dots, v_{i_n})$ and $[j_1, \dots, j_n]_*(v, \varphi) := \det(\varphi_{j_1}, \dots, \varphi_{j_n})$ as usual.)

PROOF: In view of 8.3 Corollary 3 we have to show that the invariant ring $K[V^{n-1} \oplus V^{*n-1}]^{\mathrm{SL}(V)}$ is generated by the scalar products $\langle j \mid i \rangle = \langle \varphi_j \mid v_i \rangle$. The method of "cross-sections" which we are going to use now will be important also in the next chapter.

Fix a bases $\varepsilon_1, \dots, \varepsilon_n$ of V^* and define the following subset U of $V^{n-1} \oplus V^{*n-1}$:

$$U := \{(v_1, \dots, \varphi_1, \dots, \varphi_{n-1}) \mid \varphi_1, \dots, \varphi_{n-1} \text{ linearly independent}\}.$$

Put

$$S := \{(v_1, \dots, v_{n-1}, \varepsilon_1, \dots, \varepsilon_{n-1}) \mid v_i \in V\} \subset U.$$

(This set S is a cross-section.) Restricting a function to S defines a homomorphism

$$\alpha\colon K[V^{n-1} \oplus V^{*n-1}] \to K[V^{n-1}] = K[x_{ij} \,|\, 1 \le i \le n,\ 1 \le j \le n-1]$$

by $\alpha(f)(v_1, \dots, v_{n-1}) = f(v_1, \dots, v_{n-1}, \varepsilon_1, \dots, \varepsilon_{n-1})$. (Here the coordinate function $x_{ij} \in (V^{n-1})^*$ is the composition $V^{n-1} \xrightarrow{\mathrm{pr}_j} V \xrightarrow{\varepsilon_i} K$ as usual.) Since U is ZARISKI-dense in $V^{n-1} \oplus V^{*n-1}$ and since every $\mathrm{SL}(V)$-orbit in U meets S (i.e. $\mathrm{SL}(V) \cdot S = U$) the homomorphism α restricted to the invariants $J := K[V^{n-1} \oplus V^{*n-1}]^{\mathrm{SL}(V)}$ is injective:

$$\alpha|_J\colon J \hookrightarrow K[V^{n-1}].$$

In order to prove the theorem we have to show that $J = K[\langle j \mid i \rangle \mid i, j = 1, \dots, n-1] =: A$. Now

$$\alpha(\langle i \mid j \rangle)(v_1, \dots, v_{n-1}) = \langle v_j \mid \varepsilon_i \rangle = x_{ij}(v_1, \dots, v_{n-1}),$$

hence $\alpha(A) = K[x_{ij} \mid i, j = 1, \dots, n-1]$. Since $\alpha|_J$ is injective it remains to see that $\alpha(J) \subset K[x_{ij} \mid i, j = 1, \dots, n-1]$. Consider the subgroup

$$\begin{aligned} H &:= \{g \in \mathrm{SL}(V) \mid g\varepsilon_i = \varepsilon_i \quad \text{for } i = 1, 2, \dots, n-1\} \\ &= \left\{ \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} & 1 \end{pmatrix} \mid \alpha_i \in K \right\}. \end{aligned}$$

This group H clearly stabilizes S, and every $\mathrm{SL}(V)$-invariant function restricts to an H-invariant function on $S \simeq V^{n-1}$, i.e.

$$\alpha(J) \subseteq K[V^{n-1}]^H.$$

We claim that $K[V^{n-1}]^H \subseteq K[x_{ij} \mid i, j = 1, \dots, n-1] = \alpha(A)$ which means that an H-invariant function $f \in K[x_{ij} \mid i = 1, \dots, n,\ j = 1, \dots, n-1]$ does not depend of $x_{n1}, \dots, x_{n\,n-1}$. (This clearly implies the theorem by what we have said above.)

Write the matrix $(x_{ij}) = (v_1, \dots, v_{n-1}) \in \mathrm{M}_{n \times n-1}$ as a matrix of row vectors $w_i \in K^{n-1}$:

$$(v_1, \dots, v_{n-1}) = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}, \quad w_i \in K^{n-1}.$$

Then we get for $g = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ \alpha_1 & \cdots & \alpha_{n-1} & 1 \end{pmatrix}$:

$$g \begin{pmatrix} w_1 \\ \vdots \\ w_{n-1} \\ 0 \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_{n-1} \\ \sum \alpha_i w_i \end{pmatrix}.$$

It follows that the subset $H \cdot \{ \begin{pmatrix} w_1 \\ \vdots \\ w_{n-1} \\ 0 \end{pmatrix} \mid w_i \in K^{n-1} \}$ is Zariski-dense in $\mathrm{M}_{n\times n-1}$. Hence an H-invariant function on $\mathrm{M}_{n\times n-1}$ does not depend on the last row of the matrix (x_{ij}). □

8.5 Remark. The theorem above implies the FFT for GL_n. In fact, let $f \in K[V^p \oplus V^{*q}]$ be a $\mathrm{GL}(V)$-invariant. Then it is also an $\mathrm{SL}(V)$-invariant and can therefore be written in the form

$$f = \sum_\nu \alpha_\nu p_\nu d_\nu d_\nu^* \tag{$*$}$$

where $\alpha_\nu \in K$, p_ν is a product of factors $\langle j \mid i \rangle$, d_ν a product of n_ν factors $[i_1, \ldots, i_n]$ and d_ν^* a product of m_ν factors $[j_1, \ldots, j_n]_*$. Now for every scalar $\lambda \in K^* \subset \mathrm{GL}(V)$ we have

$$f = {}^\lambda f = \sum_\nu \alpha_\nu s_\nu d_\nu d_\nu^* \lambda^{n(m_\nu - n_\nu)}.$$

Hence we may assume that $n_\nu = m_\nu$ in $(*)$. But

$$[i_1, \ldots, i_n] \cdot [j_1, \ldots, j_n]_* = \det(\langle j_\rho \mid i_\mu \rangle)_{\rho,\mu=1,\ldots,n},$$

and the claim follows.

Example. Let G be a finite group and $V_{\mathrm{reg}} = K[G]$ the *regular representation*. It's well known that every irreducible representation W of G occurs exactly $\dim W$ times in V_{reg}. (We assume K to be algebraically closed.) Hence a generating system for the ring of invariants of any representation of G can be obtained from a generating system of $K[V_{\mathrm{reg}}]^G$ by polarization.

If the group G is simple then every representation is unimodular. Hence we can replace V_{reg} by the smaller representation containing each non-trivial representation W only $\dim W - 1$ times, but we have to add all possible determinants as generators of the ring of invariants.

§ 9 The Theory of Capelli

In this paragraph we prove a fundamental relation between the polarization operators Δ_{ij} which is due to Capelli. As a consequence, we obtain a proof of the Capelli-Deruyts expansion formula from the previous paragraph (Theorem 8.1) which was the basic ingredient for the two theorems of Weyl (Theorem 7.1 and 7.5).
We start with a special case, the famous Clebsch-Gordan formula.

9.1 Clebsch-Gordan **formula.** One of the first occurrences of polarization operators and the formal relations among them was in the Clebsch-Gordan *formula.*

Consider two copies of the vector space $V = K^2$ and denote the variables by $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. On $K[V^2] = K[x_1, x_2, y_1, y_2]$ define the following linear differential operators:

$$\begin{aligned}
\Omega &:= \det \begin{pmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial y_1} \\ \frac{\partial}{\partial x_2} & \frac{\partial}{\partial y_2} \end{pmatrix} = \frac{\partial}{\partial x_1}\frac{\partial}{\partial y_2} - \frac{\partial}{\partial x_2}\frac{\partial}{\partial y_1}, \\
\Delta_{xx} &:= x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2}, \\
\Delta_{yy} &:= y_1 \frac{\partial}{\partial y_1} + y_2 \frac{\partial}{\partial y_2}, \\
\Delta_{xy} &:= x_1 \frac{\partial}{\partial y_1} + x_2 \frac{\partial}{\partial y_2}, \\
\Delta_{yx} &:= y_1 \frac{\partial}{\partial x_1} + y_2 \frac{\partial}{\partial x_2}, \\
[x, y] &:= \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}.
\end{aligned}$$

These operators belong to the ring $\mathcal{D} = \mathcal{D}(V^2) \subset \operatorname{End} K[V^2]$ of differential operators, i.e., the subalgebra generated by the derivations $\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial y_1}, \frac{\partial}{\partial y_2}$ and the multiplications with x_1, x_2, y_1, y_2 (see 7.4). Every $\delta \in \mathcal{D}$ has a unique expression in the form

$$\delta = \sum_{\alpha_1, \alpha_2, \beta_1, \beta_2} f_{\alpha_1\alpha_2\beta_1\beta_2} \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} \left(\frac{\partial}{\partial y_1}\right)^{\beta_1} \left(\frac{\partial}{\partial y_2}\right)^{\beta_2}$$

where $f_{\alpha_1,\alpha_2,\beta_1,\beta_2} \in K[V^2]$. The basic relations in $\mathcal{D}$ are

$$\frac{\partial}{\partial t} s = s \frac{\partial}{\partial t} \text{ for } s \neq t \quad \text{and} \quad \frac{\partial}{\partial t} t = t \frac{\partial}{\partial t} + 1.$$

We have also introduced the subalgebra $\mathcal{U} = \mathcal{U}(2) \subset \mathcal{D}(V^2)$ generated by the polarization operators $\Delta_{xx}, \Delta_{xy}, \Delta_{yx}, \Delta_{yy}$. The natural representations of $\mathrm{GL}(V)$ and GL_2 on V^2 define linear actions of these groups (and even of the group $\mathrm{GL}(V^2)$) on $\operatorname{End} K[V^2]$ which normalize the subalgebra $\mathcal{D}(V^2)$. We have already remarked and used the fact that the polarization operators (and hence all operators from $\mathcal{U}$) are $\mathrm{GL}(V)$-invariant, i.e., they commute with $\mathrm{GL}(V)$ (see 7.3).

Exercises

1. Prove the following identities:

(a) $\Delta_{yx}[x,y] = [x,y]\Delta_{yx}$ and $\Delta_{xx}[x,y] = [x,y](\Delta_{xx}+1)$.

(b) $\Delta_{xy}\Omega = \Omega\Delta_{xy}$ and $\Omega\Delta_{xx} = (\Delta_{xx}+1)\Omega$.

2. Show that $[x,y]\mathcal{U} = \mathcal{U}[x,y]$ and $\Omega\mathcal{U} = \mathcal{U}\Omega$. In other words, the left ideals generated by $[x,y]$ or Ω are two-sided.

3. The subalgebra of $\mathcal{D}$ consisting of constant coefficient differential operators $\delta = \sum c_{\alpha_1\alpha_2\beta_1\beta_2} \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} \left(\frac{\partial}{\partial y_1}\right)^{\beta_1} \left(\frac{\partial}{\partial y_2}\right)^{\beta_2}$, $c_{\alpha_1\alpha_2\beta_1\beta_2} \in K$, is canonically isomorphic to the symmetric algebra $\mathrm{S}(V^2)$. The multiplication induces an isomorphism

$$K[V^2] \otimes \mathrm{S}(V^2) \xrightarrow{\sim} \mathcal{D}(V^2)$$

which is equivariant with respect to $\mathrm{GL}(V)$ and $\mathrm{GL}_2(K)$ (and even $\mathrm{GL}(V^2)$).

4. Show that $\mathcal{D}(V^2)^{\mathrm{GL}(V)} = \mathcal{U}(2)$ and that $\mathcal{D}(V^2)^{\mathrm{SL}(V)}$ is generated by the polarization operators together with $[x,y]$ and Ω.

An explicit computation (which we leave to the reader) shows that we have the following identity:

$$[x,y] \cdot \Omega = \det \begin{pmatrix} \Delta_{xx}+1 & \Delta_{xy} \\ \Delta_{yx} & \Delta_{yy} \end{pmatrix} := (\Delta_{xx}+1)\Delta_{yy} - \Delta_{yx}\Delta_{xy} \qquad (2)$$

Remark. One has to be careful when expanding the determinant on the right because of the non-commutativity of the operators. We agree that a determinant will be read conventionally by columns from left to right.

Let $f \in K[V^2]$ be homogeneous of degree (m,n). From (2) we first get

$$(\Delta_{xx}+1)\Delta_{yy} = \Delta_{yx}\Delta_{xy} + [x,y]\Omega,$$

hence

$$(m+1)nf = \Delta_{yx}\Delta_{xy}f + [x,y]\Omega f$$

or

$$f = \frac{1}{(m+1)n}(\Delta_{yx}\Delta_{xy}f + [x,y]\Omega f). \tag{3}$$

Theorem (CLEBSCH-GORDAN formula). *For every degree (m,n) there are rational coefficients $a_i = a_i(m,n) \in \mathbb{Q}$ such that*

$$f = \sum_{i=0}^{\min(m,n)} a_i [x,y]^i \Delta_{yx}^{n-i} \Delta_{xy}^{n-i} \Omega^i f$$

for all homogeneous polynomials $f \in K[x_1,x_2,y_1,y_2]$ of degree (m,n).

PROOF: We use induction on $n = \deg_y f$. If $n = 0$ there is nothing to prove. So let us assume that $n > 0$. By (3) we have

$$f = \frac{1}{(m+1)n}(\Delta_{yx}\Delta_{xy}f + [x,y]\Omega f),$$

and we can apply induction on $\Delta_{xy}f$ and Ωf: There are rational coefficients b_i and c_i such that

$$\Delta_{xy}f = \sum_{i=0}^{\min(m+1,n-1)} b_i [x,y]^i \Delta_{yx}^{n-1-i} \Delta_{xy}^{n-1-i} \Omega^i \Delta_{xy} f$$

and

$$\Omega f = \sum_{j=0}^{\min(m-1,n-1)} c_j [x,y]^j \Delta_{yx}^{n-1-j} \Delta_{xy}^{n-1-j} \Omega^j \Omega f.$$

Thus

$$\begin{aligned} f &= \frac{1}{(m+1)n}(\Delta_{yx} \sum_{i=0}^{\min(m+1,n-1)} b_i [x,y]^i \Delta_{yx}^{n-1-i} \Delta_{xy}^{n-1-i} \Omega^i \Delta_{xy} f \\ &+ [x,y] \sum_{j=0}^{\min(m-1,n-1)} c_j [x,y]^j \Delta_{yx}^{n-1-j} \Delta_{xy}^{n-1-j} \Omega^j \Omega f). \end{aligned} \tag{4}$$

Now we use the fact that $\Delta_{y,x}$ commutes with $[x,y]$ (see Exercise 1) and obtain

$$f = \frac{1}{(m+1)n}(\sum_{i \geq 0} b_i [x,y]^i \Delta_{yx}^{n-i} \Delta_{xy}^{n-i} \Omega^i f + \sum_{j \geq 1} c_{j-1} [x,y]^j \Delta_{yx}^{n-j} \Delta_{xy}^{n-j} \Omega^j f),$$

hence the claim. □

9.2 CLEBSCH-GORDAN decomposition. What is the meaning of the formula of CLEBSCH-GORDAN? We first show that it has the properties of the CAPELLI-DERUYTS expansion (Theorem 8.1). Remember that in the present situation the primary covariants are given by $\mathrm{PC} = K[x_1, x_2, [x,y]]$ (see 8.1). We claim that the operator

$$B_i := [x,y]^i \Delta_{xy}^{n-i} \Omega^i$$

belongs to the algebra $\mathcal{U} = \mathcal{U}(2)$ generated by the polarization operators and that $B_i f \in \mathrm{PC}$ for any f of degree (m,n). The second claim is clear since $\Delta_{xy}^{n-i}\Omega^i f$ has degree $(m+n-2i, 0)$. For the first we remark that $B_i = \Delta_{xy}^{n-i}[x,y]^i\Omega^i$ (see Lemma 9.1) and that $[x,y]^i\Omega^i \in \mathcal{U}$. In fact, $[x,y]\Omega \in \mathcal{U}$ by formula 2 in 9.1. By induction, we can assume that $[x,y]^i\Omega^i =: P \in \mathcal{U}$. Since $\Omega\mathcal{U} = \mathcal{U}\Omega$ (Exercise 2) we obtain $[x,y]^{i+1}\Omega^{i+1} = [x,y]P\Omega = [x,y]\Omega P' \in \mathcal{U}$.

So we have an expansion

$$f = \sum A_i B_i f \quad \text{where} \quad A_i := a_i \Delta_{yx}^{n-i}$$

which has the required properties of the CAPELLI-DERUYTS expansion 8.1.

Next we remark that the polarization operators Δ commute with the action of $\mathrm{GL}(V)$ and that $[x,y]$ commutes with $\mathrm{SL}(V)$. Hence the operator $\Delta_{xy}^{n-i}\Omega^i$ defines a $\mathrm{SL}(V)$-*equivariant linear map*

$$p_i \colon R_{m,n} \to R_{m+n-2i,0}$$

where $R_{r,s} := \{f \in K[V \oplus V] \mid f \text{ homogeneous of degree } (r,s)\}$.

Proposition. *There is an* $\mathrm{SL}(V)$*-equivariant isomorphism*

$$R_{m,n} \xrightarrow{\sim} \bigoplus_{i=0}^{\min(m,n)} R_{m+n-2i,0}$$

given by $f \mapsto (\dots, \Delta_{xy}^{n-i}\Omega^i f, \dots)$. *The inverse map is*

$$(f_0, f_1, \dots) \mapsto \sum_{i=0}^{\min(m,n)} a_i [x,y]^i \Delta_{yx}^{n-i} f_i.$$

PROOF: By the CLEBSCH-GORDAN formula the composition of the two maps is the identity on $R_{m,n}$. It therefore suffices to show that the two spaces have the save dimension, i.e.

$$\sum_{i=0}^{\min(m,n)} (m+n-2i+1) = (m+1)(n+1).$$

We leave this as an exercise. □

Remark. Denote by $R_i := K[V]_i = K[x_1, x_2]_i$ the *binary forms of degree i.* Then the proposition above says that we have the following decomposition of the tensor product of the two SL_2-modules R_m and R_n:

$$R_m \otimes R_n \simeq \bigoplus_{i=0}^{\min(m,n)} R_{m+n-2i}.$$

Exercises

5. The representations R_i of SL_2 are *selfdual*, i.e. equivalent to the dual representation R_i^*. In particular $R_i \simeq \mathrm{S}^i V$.

6. The R_i, $i = 0, 1, 2, \dots$ form a complete set of representatives of the irreducible SL_2-modules.

7. We have the following decomposition formulas:

$$\begin{aligned} \mathrm{S}^2 R_i &= R_{2i} \oplus R_{2i-4} \oplus R_{2i-8} \oplus \cdots \\ \textstyle\bigwedge^2 R_i &= R_{2i-2} \oplus R_{2i-6} \oplus \cdots \end{aligned}$$

9.3 CAPELLI**'s identity.** CAPELLI was able to generalize the formula (1) of 9.1 to any number of variables. We write the vectors $x_1, x_2, \dots, x_p \in V = K^n$ as column vectors:

$$x_i = \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix} \qquad i = 1, \dots, p.$$

Recall that the polarization operators Δ_{ij} are the differential operators on $K[V^p]$ given by

$$\Delta_{ij} = \Delta_{x_i x_j} = \sum_{\nu=1}^{n} x_{\nu i} \frac{\partial}{\partial x_{\nu j}}$$

(see 7.3). Now we define the CAPELLI *operator* C by

$$C := \det \begin{pmatrix} \Delta_{11} + (p-1) & \Delta_{12} & \Delta_{13} & \cdots & \Delta_{1p} \\ \Delta_{21} & \Delta_{22} + (p-2) & \Delta_{23} & \cdots & \Delta_{2p} \\ \vdots & & & & \\ \Delta_{p1} & \Delta_{p2} & & \cdots & \Delta_{pp} \end{pmatrix}$$

(Remember that the expansion of the determinant is by columns from left to right.) The main result which is due to CAPELLI states that we have the formal analogs of the usual rules for determinants.

Theorem (CAPELLI). *The differential operator C satisfies the following relations:*

(a) *If $n < p$, then $C = 0$.*

(b) *If $n = p$ then $C = [x_1, \dots, x_n]\Omega$ where $\Omega := \det\left(\frac{\partial}{\partial x_{ij}}\right)_{i,j=1}^{n}$.*

(c) *If $n > p$ then*

$$C = \sum_{j_1 < j_2 < \dots < j_p} \det \begin{pmatrix} x_{j_1 1} & \cdots & x_{j_1 p} \\ x_{j_2 1} & \cdots & x_{j_2 p} \\ \vdots & & \\ x_{j_p 1} & \cdots & x_{j_p p} \end{pmatrix} \det \begin{pmatrix} \frac{\partial}{\partial x_{j_1 1}} & \cdots & \frac{\partial}{\partial x_{j_1 p}} \\ \vdots & & \\ \frac{\partial}{\partial x_{j_p 1}} & \cdots & \frac{\partial}{\partial x_{j_p p}} \end{pmatrix}.$$

PROOF: The idea of the proof is quite simple. The identities are similar to the usual determinant identities except for the additional summands $p - i$ in the diagonal of C which would hold if we were dealing with commuting operators. We want to use this and introduce a second set of variables $\xi_1, \dots, \xi_p$,

$$\xi_i = \begin{pmatrix} \xi_{1i} \\ \vdots \\ \xi_{ni} \end{pmatrix}$$

with corresponding operators

$$\Delta_{\xi_i x_j} := \sum_{\nu=1}^{n} \xi_{\nu i} \frac{\partial}{\partial x_{\nu j}} \quad \text{and} \quad C_{\xi,x} := \det(\Delta_{\xi_i x_j}).$$

Since we are working with distinct variables ξ_i and x_j we can use the usual rules to calculate the determinant $C_{\xi,x}$ in the three cases $n < p$, $n = p$, $n > p$ and obtain

(a) $C_{\xi,x} = 0$ for $n < p$,

(b) $C_{\xi,x} = [\xi_1, \dots, \xi_n] \cdot \det\left(\frac{\partial}{\partial x_{ij}}\right)$ for $n = p$,

(c) $C_{\xi,x} = \sum_{i_1,\dots,i_p} |\xi|_{i_1,\dots,i_p} \cdot \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p}$ for $n > p$,

where $|\xi|_{i_1,\dots,i_p}$ and $\left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p}$ are the minors of the matrices $(\xi_1, \dots, \xi_n)$ and $\left(\frac{\partial}{\partial x_{ij}}\right)$ corresponding to the rows $i_1, \dots, i_p$. Now we apply on both sides of these equations the operator $\Delta_{x_p \xi_p} \Delta_{x_{p-1}\xi_{p-1}} \cdots \Delta_{x_1 \xi_1}$ which eliminates the variables ξ_i. What we need to prove is the following:

(a) The operator $\Delta_{x_p\xi_p} \cdots \Delta_{x_1\xi_1} C_{\xi,x}$ coincides with C on functions which do not depend on the variables ξ_i.

(b) The operator $\Delta_{x_p\xi_p} \cdots \Delta_{x_1\xi_1} |\xi|_{i_1,\dots,i_p} \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p}$ coincides with the operators $|x|_{i_1,\dots,i_p} \cdot \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p}$ on functions not depending on the variables ξ_i.

The second statement is clear since

$$\Delta_{x_p\xi_p} \cdots \Delta_{x_1\xi_1} |\xi|_{i_1,\dots,i_p} = |x|_{i_1,\dots,i_p}.$$

For the proof of (a) we need the following commutation relations: Let $a, b, c, d \in \{x_i, \xi_i\}$ be distinct. Then

(a) Δ_{ab} commutes with Δ_{cd},

(b) $\Delta_{ab}\Delta_{bc} = \Delta_{bc}\Delta_{ab} + \Delta_{ac}$,

(c) $\Delta_{ab}\Delta_{ca} = \Delta_{ca}\Delta_{ab} - \Delta_{cb}$,

(d) $\Delta_{ab}\Delta_{ba} = \Delta_{ba}\Delta_{ab} + \Delta_{aa} - \Delta_{bb}$.

The verification is straightforward; we leave it as an exercise.

We want to prove by induction that $\Delta_{x_k\xi_k} \cdots \Delta_{x_1\xi_1} C_{\xi,x}$ equals the following operator

$$C_k := \begin{vmatrix} \Delta_{x_1x_1} + (p-1) & \cdots & \Delta_{x_1x_k} & \Delta_{x_1x_{k+1}} & \cdots & \Delta_{x_1x_p} \\ \Delta_{x_2x_1} & \cdots & \Delta_{x_2x_k} & \Delta_{x_2x_{k+1}} & \cdots & \Delta_{x_2x_p} \\ \vdots & & \vdots & \vdots & & \vdots \\ \Delta_{x_kx_1} & \cdots & \Delta_{x_kx_k} + (p-k) & \Delta_{x_kx_{k+1}} & \cdots & \Delta_{x_kx_p} \\ \Delta_{\xi_{k+1}x_1} & \cdots & \Delta_{\xi_{k+1}x_k} & \Delta_{\xi_{k+1}x_{k+1}} & \cdots & \Delta_{\xi_{k+1}x_p} \\ \vdots & & \vdots & \vdots & & \vdots \\ \Delta_{\xi_px_1} & \cdots & \Delta_{\xi_px_k} & \Delta_{\xi_px_{k+1}} & \cdots & \Delta_{\xi_px_p} \end{vmatrix}$$

i.e., C_k it is the determinant of the matrix whose first k rows are those of C and the last $p-k$ rows are those of $C_{\xi,x}$. (We write $|A|$ for the determinant of the matrix A.) Clearly, it is enough to show that $\Delta_{x_k\xi_k} C_{k-1} = C_k$. Using the commutation rules we obtain (for $k > 1$)

$$\Delta_{x_k\xi_k} C_{k-1} = \begin{vmatrix} (\Delta_{x_1x_1} + p - 1)\Delta_{x_k\xi_k} & \\ \vdots & \\ \Delta_{x_{k-1}x_1}\Delta_{x_k\xi_k} & \\ \Delta_{\xi_kx_1}\Delta_{x_k\xi_k} + \Delta_{x_kx_1} & \\ \Delta_{\xi_{k+1}x_1}\Delta_{x_k\xi_k} & \\ \vdots & \end{vmatrix}$$

$$= \begin{vmatrix} 0 & \\ \vdots & \\ 0 & \\ \Delta_{x_kx_1} & \\ 0 & \\ \vdots & \\ 0 & \end{vmatrix} + \begin{vmatrix} (\Delta_{x_1x_1} + p - 1)\Delta_{x_k\xi_k} & \\ \vdots & \\ \Delta_{x_{k-1}x_1}\Delta_{x_k\xi_k} & \\ \Delta_{\xi_kx_1}\Delta_{x_k\xi_k} & \\ \vdots & \\ \Delta_{\xi_px_1}\Delta_{x_k\xi_k} & \end{vmatrix}$$

We agree that all entries of our matrices where there is no indication are those of the original determinant C_{k-1}. For example, in the matrices above the last $p-1$ columns are those of C_{k-1}.

If we apply the second summand to a function $f(x_1, .., x_p)$ the term $\Delta_{\xi_kx_1}\Delta_{x_k\xi_k}$ gives contribution zero: Look at the cofactor D of this term (i.e., delete the first column and the k-th row and take the determinant); it does not contain a term of the form $\Delta_{\xi_kx_i}$, hence Df does not depend on the variable ξ_k. Therefore we get for the sum above

$$\begin{vmatrix} 0 & & & \\ \vdots & & & \\ 0 & & & \\ \Delta_{x_kx_1} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{vmatrix} + \begin{vmatrix} (\Delta_{x_1x_1} + p - 1)\Delta_{x_k\xi_k} & \\ \vdots & \\ \Delta_{x_{k-1}x_1}\Delta_{x_k\xi_k} & \\ 0 & \\ \Delta_{\xi_{k+1}x_1}\Delta_{x_k\xi_k} & \\ \vdots & \\ \Delta_{\xi_px_1}\Delta_{x_k\xi_k} & \end{vmatrix}$$

In the second summand—call it S—we bring the term $\Delta_{x_k\xi_k}$ over to the second column:

$$S = \begin{vmatrix} \Delta_{x_1x_1} + p - 1 & \Delta_{x_k\xi_k}\Delta_{x_1x_2} & \\ \Delta_{x_2x_1} & \Delta_{x_k\xi_k}(\Delta_{x_2x_2} + p - 2) & \\ \vdots & \vdots & \\ \Delta_{x_{k-1}x_1} & \Delta_{x_k\xi_k}\Delta_{x_{k-1}x_k} & \\ 0 & \Delta_{x_k\xi_k}\Delta_{\xi_kx_2} & \\ \Delta_{\xi_{k+1}x_1} & \Delta_{x_k\xi_k}\Delta_{\xi_{k+1}x_2} & \\ \vdots & \vdots & \end{vmatrix}$$

Applying again the commutation rules and dropping the factor $\Delta_{\xi_kx_2}\Delta_{x_k\xi_k}$ using the same reasoning as above, we get (in case $k > 2$)

$$S = \begin{vmatrix} \Delta_{x_1x_1} + p - 1 & 0 & & & \\ \Delta_{x_2x_1} & 0 & & & \\ \vdots & \vdots & & & \\ \Delta_{x_{k-1}x_1} & 0 & & & \\ 0 & \Delta_{x_kx_2} & 0 & \cdots & 0 \\ \Delta_{\xi_{k+1}x_1} & 0 & & & \\ \vdots & \vdots & & & \\ \Delta_{\xi_px_1} & 0 & & & \end{vmatrix}$$

$$+ \begin{vmatrix} \Delta_{x_1x_1} + p - 1 & \Delta_{x_1x_2} & \Delta_{x_k\xi_k}\Delta_{x_1x_3} & \\ \Delta_{x_2x_1} & \Delta_{x_2x_2} + p - 2 & \Delta_{x_k\xi_k}\Delta_{x_2x_3} & \\ \vdots & \vdots & \vdots & \\ \Delta_{x_{k-1}x_1} & \Delta_{x_{k-1}x_2} & \Delta_{x_k\xi_k}\Delta_{x_{k-1}x_3} & \\ 0 & 0 & \Delta_{x_k\xi_k}\Delta_{\xi_kx_3} & \\ \Delta_{\xi_{k+1}x_1} & \Delta_{\xi_{k+1}x_2} & \Delta_{x_k\xi_k}\Delta_{\xi_{k+1}x_3} & \\ \vdots & \vdots & \vdots & \\ \Delta_{\xi_px_1} & \Delta_{\xi_px_2} & \Delta_{x_k\xi_k}\Delta_{\xi_px_3} & \end{vmatrix}$$

We continue this until we bring $\Delta_{x_k\xi_k}$ over to the k-th column and obtain:

$$\Delta_{x_k\xi_k}C_{k-1} =$$

$$\sum_{i=1}^{k-1} \begin{vmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & \Delta_{x_k x_i} & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{vmatrix} + \begin{vmatrix} & & & \Delta_{x_k\xi_k}\Delta_{x_1x_k} & & \\ & & & \vdots & & \\ & & & \Delta_{x_k\xi_k}\Delta_{x_{k-1}x_k} & & \\ 0 & \cdots & 0 & \Delta_{x_k\xi_k}\Delta_{\xi_kx_k} & & \\ & & & \Delta_{x_k\xi_k}\Delta_{\xi_{k+1}x_k} & & \\ & & & \vdots & & \\ & & & \Delta_{x_k\xi_k}\Delta_{\xi_px_k} & & \end{vmatrix}$$

(Remember that in the empty spaces we have to put the entries of C_{k-1}.) Applying the commutation rules to the last term—call it T—and dropping the factor $\Delta_{\xi_kx_k}\Delta_{x_k\xi_k}$ as before, we obtain three summands:

$$T = U + V + W :$$

$$\begin{aligned} U &= \begin{vmatrix} & & & \Delta_{x_1x_k}\Delta_{x_k\xi_k} & & \\ & & & \vdots & & \\ & & & \Delta_{x_{k-1}x_k}\Delta_{x_k\xi_k} & & \\ 0 & \cdots & 0 & 0 & & \\ & & & \Delta_{\xi_{k+1}x_k}\Delta_{x_k\xi_k} & & \\ & & & \vdots & & \\ & & & \Delta_{\xi_px_k}\Delta_{x_k\xi_k} & & \end{vmatrix} \\ &= \begin{vmatrix} & & & \Delta_{x_1x_k} & & & \\ & & & \vdots & & & \\ & & & \Delta_{x_{k-1}x_k} & & & \\ 0 & \cdots & 0 & 0 & \Delta_{x_k\xi_k}\Delta_{\xi_kx_{k+1}} & \cdots & \Delta_{x_k\xi_k}\Delta_{\xi_kx_p} \\ & & & \Delta_{\xi_{k+1}x_k} & & & \\ & & & \vdots & & & \\ & & & \Delta_{\xi_px_k} & & & \end{vmatrix} \end{aligned}$$

$$V = \begin{vmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & \Delta_{x_kx_k} & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{vmatrix}$$

$$W = \begin{vmatrix} & & & -\Delta_{x_1\xi_k} & & & \\ & & & \vdots & & & \\ & & & -\Delta_{x_{k-1}\xi_k} & & & \\ 0 & \cdots & 0 & 0 & & & \\ & & & -\Delta_{\xi_{k+1}x_k} & & & \\ & & & \vdots & & & \\ & & & -\Delta_{\xi_p x_k} & & & \end{vmatrix}.$$

For U we have used the fact that $\Delta_{x_k\xi_k}$ commutes with all operators in the last $p-k$ columns except those in the k-th row. Applying again commutation relations and dropping factors as above, we can replace the k-th row of U by $(0,\dots,0,\Delta_{x_kx_{k+1}},\dots,\Delta_{x_kx_p})$. Here we get for $C' := \Delta_{x_k\xi_k}C_{k-1} - W$:

$$C' := \sum_{i=1}^{k-1} \begin{vmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & \Delta_{x_kx_i} & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{vmatrix} + U + V =$$

$$\begin{vmatrix} \Delta_{x_1x_1}+p-1 & \cdots & \Delta_{x_1x_{k-1}} & \Delta_{x_1x_k} & \cdots & \Delta_{x_1x_p} \\ \Delta_{x_2x_1} & \cdots & \Delta_{x_2x_{k-1}} & \Delta_{x_2x_k} & \cdots & \Delta_{x_2x_p} \\ \vdots & & \vdots & \vdots & & \vdots \\ \Delta_{x_{k-1}x_1} & \cdots & \Delta_{x_{k-1}x_{k-1}}+p-k+1 & \Delta_{x_{k-1}x_k} & \cdots & \Delta_{x_{k-1}x_p} \\ \Delta_{x_kx_1} & \cdots & \Delta_{x_kx_{k-1}} & \Delta_{x_kx_k} & \cdots & \Delta_{x_kx_p} \\ \Delta_{\xi_{k+1}x_1} & \cdots & \Delta_{\xi_{k+1}x_{k-1}} & \Delta_{\xi_{k+1}x_k} & \cdots & \Delta_{\xi_{k+1}x_p} \\ \vdots & & \vdots & \vdots & & \vdots \\ \Delta_{\xi_px_1} & \cdots & \Delta_{\xi_px_{k-1}} & \Delta_{\xi_px_k} & \cdots & \Delta_{\xi_px_p} \end{vmatrix}$$

This is almost C_k except the (k,k)-entry which should be $\Delta_{x_kx_k}+p-k$. Therefore, we are done if we can show that W equals $(p-k)C''$ with

$$C'' := \begin{vmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{vmatrix}$$

where the integer 1 is in position (k,k) and the empty spaces have to be filled with the entries of C_{k-1}. In fact, we have

$$(p-k)C'' = \begin{vmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ \Delta_{x_k x_1} & \cdots & \Delta_{x_k x_{k-1}} & p-k & \Delta_{x_k x_{k+1}} & \cdots & \Delta_{x_k x_p} \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{vmatrix}$$

which has the same columns as C' except the k-th and therefore $C'+(p-k)C'' = C_k$. Let us expand W with respect to the k-th column:

$$W = -\sum_{i=1}^{k-1} \begin{vmatrix} & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & \Delta_{x_i \xi_k} & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ 0 & \cdots & 0 & 0 & \Delta_{\xi_k x_{k+1}} & \cdots & \Delta_{\xi_k x_p} \\ & & & \vdots & & & \end{vmatrix} -$$

$$\sum_{j=k+1}^{p} \begin{vmatrix} & & & \vdots & & & \\ 0 & \cdots & 0 & 0 & \Delta_{\xi_k x_{k+1}} & \cdots & \Delta_{\xi_k x_p} \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & \Delta_{\xi_j \xi_k} & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \end{vmatrix} .$$

In the first $k-1$ summands W_i' we can bring the factor $\Delta_{x_i \xi_k}$ to the k-th row (the only place where ξ_k appears). Using $\Delta_{x_i \xi_k} \Delta_{\xi_k x_j} = \Delta_{\xi_k x_j} \Delta_{x_i \xi_k} + \Delta_{x_i x_j}$ and dropping the first summand we obtain

$$W_i' = \begin{vmatrix} & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & \Delta_{x_i x_{k+1}} & \cdots & \Delta_{x_i x_p} \\ & & & \vdots & & & \end{vmatrix}$$

$$= - \begin{vmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 0 & \Delta_{x_i x_{k+1}} & \cdots & \Delta_{x_i x_p} \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{vmatrix} .$$

Similarly we find for the second $p-k$ summands W''_j:

$$W''_j = -\begin{vmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & \Delta_{\xi_j x_{k+1}} & \cdots & \Delta_{\xi_j x_p} \\ & & & \vdots & & & \\ & & & 0 & & & \end{vmatrix}.$$

Now we see that these matrices are obtained from C'' by replacing the first k entries in the i-th row of W'_i, $i = 1, \dots, k-1$ and in the j-th row of W''_j, $j = k+1, \dots, p$ by zero. Let us call D the matrix obtained from C'' by removing the k-th row and the k-th column, and denote by D_t, $t = 1, \dots, p-1$ the matrix obtained from D by replacing the first $k-1$ elements in the t-th row by zero. Then we clearly have

$$\begin{aligned} C'' &= \det D, \\ W'_i &= \det D_i \quad i = 1, \dots, k-1 \\ W''_j &= \det D_{j-1} \quad j = k+1, \dots, p. \end{aligned}$$

Now the claim follows from the next lemma. □

Lemma. *Let D be an $m \times m$ matrix with non necessarily commuting entries. Denote by $D_t^{(l)}$ the matrix obtained from D by replacing the first l entries in the t-th row by zero. Then for every $l = 0, 1, \dots, m$ we have*

$$\sum_{t=1}^{m} \det D_t^{(l)} = (m-l) \det D.$$

Proof: We use induction on l. Remark that

$$\det D_t^{(l-1)} = \det D_t^{(l)} + \begin{vmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & a_{tl} & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{vmatrix}.$$

Hence

$$\sum_{t=1}^{m} \det D_t^{(l-1)} = \sum_{t=1}^{m} \det D_t^{(l)} + \det D.$$

□

9.4 Some applications. Let $f = f(x_1, \dots, x_p) \in K[V^p]$ be multihomogeneous of degree $(d_1, d_2, \dots, d_p)$. As before, we denote by C_p the Capelli operator

$$C_p = \det \begin{pmatrix} \Delta_{11} + p - 1 & \Delta_{12} & \cdots & \Delta_{1p} \\ \Delta_{21} & \Delta_{22} + p - 2 & & \\ \vdots & & & \\ \Delta_{p1} & \cdots & & \Delta_{pp} \end{pmatrix}.$$

Lemma. *$C_p f = c_p \cdot f + \sum_{i<j} C_{ij} \Delta_{ij} f$, where $c_p = (d_1 + p - 1) \cdot (d_2 + p - 2) \cdots d_p$ and $C_{ij} \in \mathcal{U}(p)$, the algebra generated by the Δ_{ij}.*

Proof: By definition we have

$$C_p = \sum_{\sigma \in \S_p} \operatorname{sgn} \sigma \, \tilde{\Delta}_{\sigma(1)1} \tilde{\Delta}_{\sigma(2)2} \cdots \tilde{\Delta}_{\sigma(p)p}$$

where

$$\tilde{\Delta}_{ij} = \begin{cases} \Delta_{ij} & \text{for } i \neq j \\ \Delta_{ii} + p - i & \text{for } i = j \end{cases}$$

Now $\tilde{\Delta}_{11} \tilde{\Delta}_{22} \cdots \tilde{\Delta}_{pp} f = (d_1 + p - 1)(d_2 + p - 2) \cdots d_p \cdot f$ and every other monomial in the sum above has the form

$$\tilde{\Delta}_{i_1 1} \cdots \tilde{\Delta}_{i_j j} \tilde{\Delta}_{j+1\, j+1} \tilde{\Delta}_{j+2\, j+2} \cdots \tilde{\Delta}_{pp}$$

with $i_j < j$. This term applied to f can be written as $C_{ij} \Delta_{ij} f$ with $i = i_j < j$.

□

Application 1. *There are operators A_i, $B_i \in \mathcal{U}(p)$ such that $f = \sum_i A_i B_i f$ and $B_i f$ depends only on $x_1, \dots, x_n$.*

Proof: We may assume that $p > n$ and $d_p \neq 0$; otherwise there is nothing to prove. Thus $C_p = 0$ by Capelli's identity, and we obtain from the lemma above:

$$c \cdot f = \sum_{i<j} C_{ij} \Delta_{ij} f \qquad \text{with some } c \neq 0.$$

Now we use induction on the multidegree of f, ordered in the antilexicographic way. Then $\deg \Delta_{ij} f < \deg f$ for $i < j$, hence

$$\Delta_{ij} f = \sum_l A_l^{ij} B_l^{ij} (\Delta_{ij} f)$$

for some operators A_l^{ij}, $B_l^{ij} \in \mathcal{U}(p)$. It follows that

$$f = \frac{1}{c} \sum_{i<j} C_{ij} \sum_l A_l^{ij} B_l^{ij} \Delta_{ij} f = \sum_{i<j} (\frac{1}{c} C_{ij} A_l^{ij})(B_l^{ij} \Delta_{ij}) f$$

which has the required form. □

Remarks. (a) The proof shows that the operators A_i, B_i depend only on the multidegree of f.

(b) Application 1 is sufficient to prove the first theorem of Weyl (Theorem 7.1; see 8.2).

Application 2. *There are operators A_i, $B_i \in \mathcal{U}(p)$ such that $f = \sum A_i B_i f$ where $B_i f$ is of the form $B_i f = [x_1, \dots, x_n]^{s_i} h_i(x_1, \dots, x_{n-1})$.*

Proof: Using Application 1 it is easy to reduce to the case where f depends only on $x_1, \dots, x_n$. In this situation ($p = n$) we have the identity

$$C_n = [x_1, \dots, x_n] \cdot \Omega.$$

Hence, with the lemma above

$$c \cdot f + \sum_{i<j} C_{ij} \Delta_{ij} f = [x_1, \dots, x_n] \Omega f. \tag{1}$$

We can assume that f depends on x_n, i.e. $c = c_n \neq 0$. By induction on the multidegree (in antilexicographic order) we get

$$\Delta_{ij} f = \sum A_l^{ij} B_l^{ij} \Delta_{ij} f$$

with $B_l^{ij}\Delta_{ij}f$ of the form $[x_1,\dots,x_n]^s h(x_1,\dots,x_{n-1})$, which implies that the sum $\sum_{i<j} C_{ij}\Delta_{ij}f$ has the required form.

It remains to handle the right hand side of (1). The induction hypothesis also applies to Ωf

$$\Omega f = \sum A_k B_k \Omega f \tag{2}$$

with $B_k\Omega f = [x_1,\dots,x_n]^{s_k}\cdot h_k(x_1,\dots,x_{n-1})$. From the two relations

$$\begin{aligned} \Delta_{ij}[x_1,\dots,x_n] &= [x_1,\dots,x_n]\Delta_{ij} \qquad \text{if } i\neq j, \text{ and} \\ \Delta_{ii}[x_1,\dots,x_n] &= [x_1,\dots,x_n](\Delta_{ii}+1) \end{aligned}$$

whose proofs are straightforward we see that for every polarization operator A we have an equation $[x_1,\dots,x_n]A = A'[x_1,\dots,x_n]$ with some other operator A' (cf. Exercise 1). Hence we obtain from (2)

$$[x_1,\dots,x_n]\Omega f = \sum A'_k([x_1,\dots,x_n]B_k\Omega)f.$$

Now $[x_1,\dots,x_n]B_k\Omega f = [x_1,\dots,x_n]^{s_k+1}h_k(x_1,\dots,x_{n-1})$ and has therefore the required form. On the other hand

$$[x_1,\dots,x_n]B_k\Omega = B'_k[x_1,\dots,x_n]\Omega = B'_kC_n \in \mathcal{U}(n)$$

which shows that the right hand side of (1) has the required form, too. □

Remarks. (a) Again we see that the operators A_i, B_i only depend on the multidegree of f.

(b) As above Application 2 is sufficient to prove the second theorem of WEYL (Theorem 7.5; see 8.2).

(c) We have seen in the proof that for every $D\in\mathcal{U}(n)$ we have

$$[x_1,\dots,x_n]D = D'[x_1,\dots,x_n]$$

for some $D'\in\mathcal{U}(n)$, i.e., *conjugation with* $[x_1,\dots,x_n]$ *induces an automorphism of* $\mathcal{U}(n)$. We will denote it by $D\mapsto D'$.
(Remember that all this takes place in $\mathrm{End}(K[V^n])$.)

9.5 Proof of the CAPELLI-DERUYTS **expansion.** Let us recall the notation. The vectors x_i are written as column vectors

$$x_i = \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix} \in V = K^n.$$

In case $p \le n$ we denoted by $|x|_{i_1,\dots,i_p}$ the minor of $(x_1,\dots,x_n)$ extracted from the rows $i_1,\dots,i_p$. Similarly, $\left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p}$ is defined with respect to the matrix $\left(\frac{\partial}{\partial x_{ij}}\right)_{i,j=1,\dots,n}$.

CAPELLI-DERUYTS **expansion.** *For every multihomogeneous polynomial $f = f(x_1,\dots,x_p)$ there are operators A_i, $B_i \in \mathcal{U}(p)$ depending only on the multidegree of f such that*

$$f = \sum_i A_i B_i f \quad \textit{and} \quad B_i f \in \mathrm{PC},$$

where $\mathrm{PC} \subset K[V^n]$ is the subalgebra of primary covariants, generated by all $k \times k$-minors extracted from the first k columns of (x_{ij}), $k = 1, 2, \dots, n$.

Using 9.4 Application 1 one reduces to the case $p = n$. The claim now follows from the next proposition which is slightly more general.

Proposition. *Let $p \le n$ and $f(x_1,\dots,x_p)$ multihomogeneous. Then there are operators A_i, $B_i \in \mathcal{U}(p)$ depending only on the multidegree of f such that*

$$f = \sum A_i B_i f$$

and $B_i f$ is a sum of monomials in the $k \times k$-minors extracted from the first k columns of (x_{ij}), where $k = 1,\dots,p$.

PROOF: The case $p = 1$ is obvious. Let us assume $p > 1$ and that f depends on x_p. Now the CAPELLI-identity says

$$C_p f = \sum_{i_1,\dots,i_p} |x|_{i_1,\dots,i_p} \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p} f. \tag{1}$$

By Lemma 9.4 we have

$$C_p f = c \cdot f + \sum_{i<j} C_{ij} \Delta_{ij} f \quad \text{with } c = c_p \neq 0$$

Now the induction hypothesis applies to $\Delta_{ij} f$ and we obtain operators A_l^{ij}, $B_l^{ij} \in \mathcal{U}(p)$ such that

$$\Delta_{ij} f = \sum_l A_l^{ij} B_l^{ij} f \quad \text{with } B_l^{ij} \text{ of the required form.}$$

To handle the right hand side of (1) we can also use induction:

$$\left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p} f = \sum A_l B_l \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p}$$

where A_l,$B_l \in \mathcal{U}(p)$ and $B_l \left|\frac{\partial}{\partial x}\right|_{i_1\dots i_p} f$ are of the required form.
(Since $\deg \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p} f = \deg f - (1,1,\dots,1)$, the operators A_l, B_l do not depend on $i_1,\dots,i_p$.) Now we have

$$|x|_{i_1,\dots,i_p} \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p} f = \sum A_l' |x|_{i_1,\dots,i_p} B_l \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p} f$$

and $|x|_{i_1,\dots,i_p} B_l \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p} f$ has the required form. On the other hand we get

$$\begin{aligned}\sum_{i_1,\dots,i_p} |x|_{i_1,\dots,i_p} \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p} f &= \sum_l A_l' B_l' \sum_{i_1,\dots,i_p} |x|_{i_1,\dots,i_p} \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p} f \\ &= \sum_l A_l' B_l' C_p f.\end{aligned}$$

Hence, $B_l' C_p = \sum_{i_1,\dots,i_p} |x|_{i_1,\dots,i_p} B_l \left|\frac{\partial}{\partial x}\right|_{i_1,\dots,i_p} \in \mathcal{U}(p)$ and we are done. □

§ 10 The First Fundamental Theorem for Orthogonal and Symplectic Groups

As another application of the Theorems of WEYL (7.1, 7.5) we prove the First Fundamental Theorem for the orthogonal group O_n, the special orthogonal group SO_n and the symplectic group Sp_n, i.e., we describe a system of generators for the invariants on several copies of the natural representation of these groups.
Throughout the whole paragraph we assume that $\operatorname{char} K = 0$.

10.1 Orthogonal and symplectic groups. Let $V = K^n$ and denote by

$$(v \mid w) := \sum_{i=1}^{n} x_i y_i \quad \text{for } v = (x_1, \ldots, x_n),\ w = (y_1, \ldots, y_n) \in V$$

the standard symmetric form on V. As usual the *orthogonal group* is defined by

$$O_n := O_n(K) := \{g \in GL_n(K) \mid (gv \mid gw) = (v \mid w) \text{ for all } v, w \in V\}$$

and the *special orthogonal group* by

$$SO_n := SO_n(K) := \{g \in O_n(K) \mid \det g = 1\} = O_n(K) \cap SL_n(K).$$

It is easy to see that SO_n is a subgroup of O_n of index 2. In terms of matrices A we have $A \in O_n$ if and only if $A^t A = E$ where A^t denotes the transposed matrix.

Exercises

1. Show that $SO_2(\mathbb{C})$ is isomorphic to $\mathbb{C}^*$ and that the natural 2-dimensional representation of $SO_2(\mathbb{C})$ corresponds to the representation of $\mathbb{C}^*$ with weights ± 1.
(Hint: An explicit isomorpism is given by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto a + ib$.)

2. Show that $O_2(\mathbb{C})$ is a semidirect product of $\mathbb{C}^*$ with $\mathbb{Z}/2$. More generally, O_n is a direct product of SO_n with $\mathbb{Z}/2$ for odd n and a semidirect product for even n.

$$\langle v \mid w \rangle := \sum_{j=1}^{m} (x_{2j-1} y_{2j} - x_{2j} y_{2j-1})$$

be the standard skew symmetric form on V and define the *symplectic group* by

$$Sp_n := Sp_n(K) := \{g \in GL_n(K) \mid \langle gv \mid gw \rangle = \langle v \mid w \rangle \text{ for } v, w \in V\}.$$

To get a more explicit description of Sp_n we introduce the following skew-symmetric $n \times n$-matrix

$$J := \begin{pmatrix} I & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{pmatrix} \quad \text{where } I := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

By definition we have $\langle v \mid w \rangle = vJw^t$. Therefore, a $n \times n$-matrix A belongs to $\mathrm{Sp_n}$ if and only if $A^t JA = J$.

Exercises

3. Show that $\mathrm{Sp}_2(K) = \mathrm{SL}_2(K)$.

4. Show that all non-degenerate skew forms on K^{2m} are equivalent, i.e., for any skew-symmetric $2m \times 2m$-matrix S there is a $g \in \mathrm{GL_n}(K)$ such that $g^t Sg = J$. This holds for any field K in any characteristic. (In characteristic 2 one has to assume that the form is alternating, i.e., $\langle v \mid v \rangle = 0$ for all $v \in V$.)

5. Let $n = 2m$ or $n = 2m + 1$ and consider the subgroup

$$T := \underbrace{\mathrm{SO}_2 \times \mathrm{SO}_2 \times \cdots \times \mathrm{SO}_2}_{m \text{ times}} \subset \mathrm{SO}_n$$

embedded in the usual way.

(a) The matrices $A \in \mathrm{SO}_n(\mathbb{C})$ which are diagonalizable (in $\mathrm{GL}_n(\mathbb{C})$) form a Zariski-dense subset.

(b) Every diagonalizable matrix $A \in \mathrm{SO}_n(\mathbb{C})$ is conjugate (in $\mathrm{SO}_n(\mathbb{C})$) to a matrix of T.

(c) For any field extension L/K the subgroup $\mathrm{SO}_n(K)$ is Zariski-dense in $\mathrm{SO}_n(L)$.

Lemma. $\mathrm{Sp_{2m}}$ *is a subgroup of* SL_{2m}.

Proof: Recall that the *Pfaffian* $\mathrm{Pf}(S)$ of a skew-symmetric $2m \times 2m$-matrix $S \in \mathrm{M}_{2m}(K)$ is a homogeneous polynomial of degree m in the entries of S which is determined by the following two conditions:

$$\det S = (\mathrm{Pf}\, S)^2 \quad \text{and} \quad \mathrm{Pf}(J) = 1.$$

We claim that $\mathrm{Pf}(g^t Sg) = \det g \cdot \mathrm{Pf}\, S$ for all $g \in \mathrm{GL}_{2m}(K)$. In fact, consider the function $f(g) := \mathrm{Pf}(g^t Sg) \cdot (\det g \cdot \mathrm{Pf}\, S)^{-1}$ which is defined for $g \in \mathrm{GL}_{2m}(K)$ and S an invertible skew-symmetric $2m \times 2m$-matrix. The first condition above implies that $f(g)^2 = \det(g^t Sg) \cdot ((\det g)^2 \cdot \det S)^{-1} = 1$. Since $f(e) = 1$ the claim follows.

Now let $g \in \mathrm{Sp_{2m}}$. Then $g^t Jg = J$ and so $1 = \mathrm{Pf}\, J = \mathrm{Pf}(g^t Jg) = \det g \cdot \mathrm{Pf}\, J = \det g$. □

Exercises

6. Let $v_1, \dots, v_{2m} \in K^{2m}$. Then $\mathrm{Pf}((\langle v_i \mid v_j \rangle)_{i,j} = [v_1, \dots, v_{2m}]$.
(Hint: If A is the matrix with rows $v_1, \dots, v_{2m}$, then $AJA^t = (\langle v_i \mid v_j \rangle)_{i,j}$.)

7. For any field extension L/K the subgroup $\mathrm{Sp_n}(K)$ is ZARISKI-dense in $\mathrm{Sp_n}(L)$.
(Hint: As in Exercise 5 we can define a subgroup

$$T := \underbrace{\mathrm{Sp}_2 \times \cdots \times \mathrm{Sp}_2}_{m \text{ times}} \subset \mathrm{Sp_{2m}}$$

with similar properties.)

p copies of the natural representation $V = K^n$: For $v = (v_1, \dots, v_p) \in V^p$ and $g \in \mathrm{O_n}$, $\mathrm{SO_n}$ or $\mathrm{Sp_{2m}}$ set $g \cdot v := (gv_1, \dots, gv_p)$. If we apply the symmetric form to the ith and jth factor of V^p we obtain for every pair $1 \le i, j \le p$ an $\mathrm{O_n}$-invariant function $v = (v_1, \dots, v_p) \mapsto (v_i \mid v_j)$ which we denote by $(i \mid j)$:

$$(i \mid j)(v_1, \dots, v_p) := (v_i \mid v_j).$$

Now the First Fundamental Theorem (shortly FFT) states that these functions generate the ring of invariants:

10.2 First Fundamental Theorem for $\mathrm{O_n}$ and $\mathrm{SO_n}$.

(a) *The invariant algebra $K[V^p]^{\mathrm{O_n}}$ is generated by the invariants $(i \mid j)$, $1 \le i \le j \le p$.*
(b) *The invariant algebra $K[V^p]^{\mathrm{SO_n}}$ is generated by the invariants $(i \mid j)$, $1 \le i \le j \le p$ together with the determinants $[i_1, \dots, i_n]$, $1 \le i_1 < i_2 < \ldots < i_n \le p$.*

every pair $1 \le i, j \le p$ the following invariant functions on V^p:

$$\langle i \mid j \rangle(v_1, \dots, v_p) := \langle v_i \mid v_j \rangle.$$

10.3 First Fundamental Theorem for $\mathrm{Sp_{2m}}$. *The algebra $K[V^p]^{\mathrm{Sp_{2m}}}$ is generated by the invariants $\langle i \mid j \rangle$, $1 \le i < j \le p$.*

PROOF OF THE FFT FOR $\mathrm{O_n}$ AND $\mathrm{SO_n}$: We first remark that (a) follows from (b). In fact, since a determinant $[i_1, \dots, i_n]$ is mapped to $-[i_1, \dots, i_n]$ under any $g \in \mathrm{O_n} \setminus \mathrm{SO_n}$ we see from (b) that the $\mathrm{O_n}$-invariants are generated by the $(i \mid j)$ and all products $[i_1, \dots, i_n][j_1, \dots, j_n]$ of two determinants. But

$$[i_1, \dots, i_n][j_1, \dots, j_n] = \det((i_k \mid i_l)_{k,l=1}^n)$$

and the claim follows.

Secondly, we can assume that K is algebraically closed. In fact, we have $\overline{K}[V^p]^{\mathrm{SO_n}(K)} = \overline{K} \otimes_K (K[V^p]^{\mathrm{SO_n}(K)})$ where $\overline{K}$ denotes the algebraic closure of K (cf. 1.5 Exercise 29), and $\overline{K}[V^p]^{\mathrm{SO_n}(K)} = \overline{K}[V^p]^{\mathrm{SO_n}(\overline{K})}$ because $\mathrm{SO_n}(K)$ is ZARISKI-dense in $\mathrm{SO_n}(\overline{K})$ (see Exercise 5). This will allow us to use geometric arguments. Moreover, by WEYL's Theorem B (7.5) it suffices to prove (b) for $p = n-1$. We will procede by induction on n.

For $n = 1$ the group $\mathrm{SO_n}$ is trivial and there is nothing to prove. Assume now that $n > 1$. We identify $V' := K^{n-1}$ with the subspace $\{v = (x_1, \ldots, x_{n-1}, 0)\} \subset V$. The normalizer N in $\mathrm{SO_n}$ of the orthogonal decomposition $V = V' \oplus Ke_n$ is the intersection $(\mathrm{GL}_{n-1} \times \mathrm{GL}_1) \cap \mathrm{SO_n}$ which has the form

$$N = \{(g, \lambda) \mid g \in \mathrm{O}_{n-1},\ \lambda = \det g\}.$$

It follows that the restriction homomorphism $K[V^{n-1}] \mapsto K[V'^{n-1}]$ induces a map

$$\rho \colon K[V^{n-1}]^{\mathrm{SO_n}} \to K[V'^{n-1}]^{\mathrm{O}_{n-1}}.$$

Clearly, the invariant $(i \mid j)$ is mapped onto the corresponding invariant on V'^{n-1} which we denote by $(i \mid j)'$. By induction (and by what we said above), the $(i \mid j)'$ generate $K[V'^{n-1}]^{\mathrm{O}_{n-1}}$. Thus it suffices to show that the homomorphism ρ is injective. This follows if we can prove that the set

$$\mathrm{O_n} \cdot V'^{n-1} = \{g \cdot v \mid g \in \mathrm{O_n},\ v \in V'^{n-1}\}$$

is ZARISKI-dense in V^{n-1}. Define

$$Z := \{v = (v_1, \ldots, v_{n-1}) \in V^{n-1} \mid \det((v_i \mid v_j)) \neq 0\}.$$

This is the ZARISKI-open subset of V^{n-1} where the function $\det((v_i \mid v_j))$ does not vanish (see 1.3). Given $v = (v_1, \ldots, v_n) \in Z$, the subspace $W(v)$ spanned by $v_1, \ldots, v_n$ has dimension $n-1$ and the symmetric form $(.\mid.)$ restricted to $W(v)$ is non-degenerate. Hence, there is a $g \in \mathrm{O_n}$ such that $g(W(v)) = V'$. (in fact, $W(v)$ and V' are isomorphic as orthogonal spaces, because K is algebraically closed, and $V = W(v) \oplus W(v)^\perp = V' \oplus Ke_n$ are both orthogonal decompositions.) It follows that $g \cdot v = (gv_1, \ldots, gv_{n-1}) \in V'^{n-1}$. As a consequence, $Z \subset \mathrm{O_n} \cdot V'^{n-1}$ which shows that $\mathrm{O_n} \cdot V'^{n-1}$ is ZARISKI-dense in V^{n-1}. □

PROOF OF THE FFT FOR $\mathrm{Sp_{2m}}$: Again we can use WEYL's Theorem B (7.5) since $\mathrm{Sp_{2m}}$ is a subgroup of SL_{2m} by Lemma 10.1 above and the determinants

$[v_1, \dots, v_{2m}]$ are contained in $K[\langle i \mid j\rangle]$ (see Exercise 6). Thus it suffices to consider the case $p = 2m-1$. As before, we assume that K is algebraically closed and we proceed by induction on m.

For $m = 1$ we have $\mathrm{Sp}_2 = \mathrm{SL}_2$ (Exercise 3) and the claim is obvious since there are no non-constant invariants. Assume $m > 1$ and consider the two subspaces

$$\begin{aligned} V' &:= \{v = (x_1, \dots, x_{2m-2}, 0, \dots, 0)\} \quad \text{and} \\ V'' &:= \{v = (0, \dots, 0, x_{2m-1}, \dots, x_{2m})\} \end{aligned}$$

of V. The restriction of the skew form $\langle . \mid . \rangle$ to both is non-degenerated and $V = V' \oplus V''$ is an orthogonal decomposition. The embeddings

$$V'^{2m-1} \hookrightarrow V'^{2m-2} \oplus (V' \oplus V'') \hookrightarrow V^{2m-1}$$

induce homomorphisms (by restriction of functions)

$$\begin{array}{ccc} K[V^{2m-1}]^{\mathrm{Sp_{2m}}} & \xrightarrow{\ \rho\ } & K[V'^{2m-1}]^{\mathrm{Sp}_{2m-2}} \otimes K[V'']^{\mathrm{Sp}_2} \\ & & \downarrow \simeq \\ & & K[V'^{2m-1}]^{\mathrm{Sp}_{2m-2}} \end{array}$$

where the second map is an isomorphism since $K[V'']^{\mathrm{Sp}_2} = K$. By induction (and WEYL's Theorem) we can assume that the images of the functions $\langle i \mid j\rangle$ generate the invariants $K[V'^{2m-1}]^{\mathrm{Sp}_{2m-2}}$. Hence, it suffices to prove that ρ is injective. Again this will follow if we show that the set

$$\mathrm{Sp_{2m}} \cdot (V'^{2m-1} \oplus V) = \\ \{(gv_1, \dots, gv_{2m-1}) \mid g \in \mathrm{Sp_{2m}},\ v_1, \dots, v_{2m-2} \in V',\ v_{2m} \in V\}$$

is ZARISKI-dense in V^{2m-1}. For this consider the subset

$$Z := \{v = (v_1, \dots, v_{2m-1}) \in V^{2m-1} \mid \det(\langle v_i \mid v_j\rangle)_{i,j=1}^{2m-2} \neq 0\}$$

which is ZARISKI-open in V^{2m-1}. For any $v \in Z$ the subspace $W(v)$ spanned by $v_1, \dots, v_{2m-2}$ is of dimension $2m-2$ and the restriction of the skew form $\langle . \mid . \rangle$ to $W(v)$ is non-degenerate. Thus, there is a $g \in \mathrm{Sp_{2m}}$ such that $gW(v) = V'$ and therefore $gv = (gv_1, \dots, gv_{2m-2}, gv_{2m-1}) \in V'^{2m-2} \oplus V$. Or, $Z \subset \mathrm{Sp_{2m}} \cdot (V'^{2m-2} \oplus V)$ and the claim follows. □

10.4 Simultaneous conjugation of 2×2-matrices. Let K be algebraically closed, e.g. $K = \mathbb{C}$. The adjoint representation of SL_2 is isomorphic to the standard representation of SO_3. In fact, the representation of SL_2 on $\mathfrak{sl}_2 := \{A \in \mathrm{M}_2 \mid \mathrm{Tr}\, A = 0\}$ by conjugation leaves the non-degenerate quadratic form

$q(A) := \det A - \frac{1}{2}\operatorname{Tr} A^2$ invariant. Using the basis $E_1 := \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$, $E_2 := \left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right)$, $E_3 := \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ we find $q(x_1E_1 + x_2E_2 + x_3E_3) = x_1^2 + x_2^2 + x_3^2$. It follows that the image of SL_2 in $\mathrm{GL}(\mathfrak{sl}_2) = \mathrm{GL}_3$ is contained in SO_3. One verifies that they are in fact equal (see Exercise 8 and 9 below). As a consequence from the FFT for SO_3 we find that the invariants of any number of copies of $\mathfrak{sl}_2$ under simultaneous conjugation are generated by the "quadratic" traces $\mathrm{Tr}_{ij}\colon (A_1, \ldots, A_p) \mapsto \mathrm{Tr}(A_iA_j)$ and the "cubic" traces $\mathrm{Tr}_{ijk}\colon (A_1, \ldots, A_p) \mapsto \mathrm{Tr}(A_iA_jA_k)$:

$$K[\mathfrak{sl}_2{}^p]^{\mathrm{SL}_2} = K[\mathrm{Tr}_{ij}, \mathrm{Tr}_{ijk} \mid 1 \le i \le j \le k \le p].$$

In fact, we have $(A_i, A_j) = \frac{1}{2}\operatorname{Tr} A_iA_j$ and $[A_iA_jA_k] = \frac{1}{2}\operatorname{Tr} A_iA_jA_K$. Similarly, one finds that the invariants of several copies of 2×2-matrices under simultaneous conjugation are generated by the traces Tr_{ij}, Tr_{ijk} together with the usual traces $\mathrm{Tr}_i\colon (A_1, \ldots, A_p) \mapsto \operatorname{Tr} A_i$:

Proposition. $K[\mathrm{M}_2{}^p]^{\mathrm{GL}_2} = K[\mathrm{Tr}_i, \mathrm{Tr}_{ij}, \mathrm{Tr}_{ijk} \mid 1 \le i \le j \le k \le p]$.

multi-traces $\mathrm{Tr}_{i_1,\ldots,i_k}$ are generators, but it did not give any bounds for the degree k. As a consequence of the proposition above, every trace function $\mathrm{Tr}_{i_1,\ldots,i_k}$ for $k > 2$ can be expressed as a polynomial in the traces Tr_i and Tr_{ij}.

For further reading about this interesting example and related questions we refer to the article [Pro84] of Procesi.

Exercises

8. Show that a special linear automorphism φ of $\mathfrak{sl}_2(\mathbb{C})$ which leaves the determinant invariant is given by conjugation with an invertible matrix. In particular, the image of $\mathrm{SL}_2(\mathbb{C})$ in $\mathrm{GL}(\mathfrak{sl}_2) \simeq \mathrm{GL}_3(\mathbb{C})$ under the adjoint representation is equal to $\mathrm{SO}_3(\mathbb{C})$.
(Hint: Since $\varphi(\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right))$ has determinant -1 (and trace 0) it is conjugate to $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, hence we can assume that $\varphi(\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Now $\mathbb{C}\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)^{\perp} = \mathbb{C}\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \oplus \mathbb{C}\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ and φ acts on this space by an element of SO_2. It is easy to see that every such element is given by conjugation with some $\left(\begin{smallmatrix} t & 0 \\ 0 & t^{-1} \end{smallmatrix}\right)$.)

9. Give another proof of the equality $\mathrm{SL}_2(\mathbb{C})/\{\pm 1\} \simeq \mathrm{SO}_3(\mathbb{C})$ by using the following fact:
If $H \subset G \subset \mathrm{GL}_n$ are subgroups such that H contains an open Zariski-densesubset of G then $H = G$.

10.5 Remark. The First Fundamental Theorem for all classical groups has been shown to be valid over the integers $\mathbb{Z}$ by DeConcini and Procesi in [DeP76]. This fundamental result has influenced a lot the further development in this area, in particular, the so-called "Standard Monomial Theory" of Lakshmibai, Musili, Seshadri et al. which is an important tool in the geometry of flag varieties [LaS78] ([LMS79], [LaS86], [Lak]).

References

[Art91] Artin, Michael: *Algebra*, Prentice Hall, Englewood Cliffs, N.J., 1991.

[BiH63] Białynicki-Birula, A.; Hochschild, G.; Mostow, G.: *Extensions of representations of algebraic linear groups.* Amer. J. Math. **85** (1963) 131–144.

[Bor91] Borel, A.: *Linear Algebraic Groups*, Graduate Texts in Math., vol. 126, Springer-Verlag, New York–Heidelberg–Berlin, 1991. (2nd enl. ed.)

[Bou75] Bourbaki, N.: *Groupes et Algèbres de Lie, Chap. IV–VI*, Hermann, Paris, 1975.

[BtD85] Bröcker, Th.; tom Dieck, T.: *Representations of compact Lie groups*, Graduate Texts in Math., vol. 28, Springer Verlag, Berlin–Heidelberg–New York, 1985.

[Cap02] Capelli, A.: *Lezioni sulla teoria delle forme algebriche*, Napoli: Pellerano, 1902.

[Cle72] Clebsch, A.: *Theorie der binären algebraischen Formen*, Leipzig, 1872.

[DeP76] DeConcini, C.; Procesi, C.: *A characteristic free approach to invariant theory.* Adv. in Math. **21** (1976) 169–193.

[Der91] Deruyts, J.: *Essai d'une théorie générale des fomres algébriques*, Bruxelles, 1891.

[Dic14] Dickson, L. E.: *Algebraic invariants*, New York, 1914.

[DiC70] Dieudonné, J.; Carrell, J. B.: *Invariant theory, old and new.* Adv. in Math. **4** (1970) 1–80. (Als Buch bei Academic Press, New York, 1971)

[FdB76] Faa di Bruno: *Théorie de formes binaires.* 1876.

[Fog69] Fogarty, J.: *Invariant theory*, W.A. Benjamin, Inc., New York–Amsterdam, 1969.

[For84] Formanek, E.: *Invariants and the ring of generic matrices.* J. Algebra **89** (1984) 178–223.

[For87] Formanek, E.: *The invariants of $n \times n$ matrices.* In: Invariant Theory (S. S. Koh, eds.)., Lecture Notes in Math. vol. 1287, Springer Verlag, Berlin–Heidelberg–New York, 1987, pp. 18–43.

[FuH91] Fulton, W.; Harris, J.: *Representation Theory*, Graduate Texts in Math., vol. 129, Springer-Verlag, New York–Heidelberg–Berlin, 1991.

[GeS92] Geroldinger, A.; Schneider, R.: *On Davenport's Constant.* J. Comb. Theory **A61** (1992) 147–152.

[Gor68] Gordan, P.: *Beweis, dass jede Covariante und Invariante einer binären Form eine ganze Funktion mit numerischen Koeffizienten einer endlichen Anzahl solcher Formen ist.* J. Reine Angew. Math. **69** (1868) 323–354.

[Gor87] Gordan, P.: *Vorlesungen über Invariantentheorie. Zweiter Band: Binäre Formen*, Teubner Verlag, Leipzig, 1887. (Herausgegeben von G. Kerschensteiner)

[GrY03] Grace, J. H.; Young, A.: *The algebra of invariants*, Chelsea Publishing Company, New York, 1903. (Reprint)

[Gur64] Gurevič, G. B.: *Foundations of the theory of algebraic invariants*, Noordhof, Groningen, 1964.

[Had66] Hadžiev, D.: *Some problems in the theory of vector invariants.* Soviet Math. Dokl. **7** (1966) 1608–1610.

[Har77] Hartshorne, R.: *Algebraic Geometry*, Graduate Texts in Math., vol. 52, Springer-Verlag, New York–Heidelberg–Berlin, 1977.

[Hil90] Hilbert, D.: *Über die Theorie der algebraischen Formen.* Math. Ann. **36** (1890) 473–534.

[Hil93] Hilbert, D.: *Über die vollen Invariantensysteme.* Math. Ann. **42** (1893) 313–373.

[Hoc81] Hochschild, G.: *Basic theory of algebraic groups and Lie algebras*, Graduate Texts in Math., vol. 75, Springer-Verlag, New York–Heidelberg–Berlin, 1981.

[HoM73] Hochschild, G.; Mostow, G. D.: *Unipotent groups in invariant theory.* Proc. Nat. Acad. Sci. USA **70** (1973) 646–648.

[How95] Howe, R.: *Perspectives on invariant theory: Schur duality, multiplicity-free actions and beyond.* In: Israel Math. Conference Proceedings., 1995

[How95] Howe, R.: *Perspectives on invariant theory: Schur duality, multiplicity-free actions and beyond.* In: Israel Math. Conference Proceedings., 1995, pp. 1–182.

[Hum72] Humphreys, J.E.: *Introduction to Lie Algebras and Representation Theory*, Graduate Texts in Math., vol. 9, Springer-Verlag, New York–Heidelberg–Berlin, 1972.

[Hum75] Humphreys, J.E.: *Linear Algebraic Groups*, Graduate Texts in Math., vol. 21, Springer-Verlag, New York–Heidelberg–Berlin, 1975.

[Hun96] Hunziker, M.: *Classical invariant theory for finite reflection groups.*

[KKL89] Knop, F.; Kraft, H.; Luna, D.; Vust, Th.: *Local properties of algebraic group actions.* In: Algebraische Transformationsgruppen und Invariantentheorie (H. Kraft, P. Slodowy, T. A. Springer, eds.)., DMV-Seminar vol. 13, Birkhäuser Verlag, Basel–Boston, 1989, pp. 63–75.

[KKV89] Knop, F.; Kraft, H.; Vust, Th.: *The Picard group of a G-variety.* In: Algebraische Transformationsgruppen und Invariantentheorie (H. Kraft, P. Slodowy, T. A. Springer, eds.)., DMV-Seminar vol. 13, Birkhäuser Verlag, Basel–Boston, 1989, pp. 77–87.

[Kra82] Kraft, H.: *Geometric methods in representation theory.* In: Representations of Algebras, Workshop Proceedings (Puebla, Mexico 1980)., Lecture Notes in Math. vol. 944, Springer Verlag, Berlin–Heidelberg–New York, 1982

[Kra85] Kraft, H.: *Geometrische Methoden in der Invariantentheorie*, Aspekte der Mathematik, vol. D1, Vieweg Verlag, Braunschweig/Wiesbaden, 1985. (2., durchgesehene Auflage)

[KrR86] Kraft, H.; Riedtmann, Ch.: *Geometry of representations of quivers.* In: Representations of Algebras, Proceedings of the Durham symposium 1985., London Math. Soc. Lecture Notes vol. 116, 1986, pp. 109–145.

[KrS92] Kraft, H.; Schwarz, G.-W.: *Reductive group actions with one-dimensional quotient.* Publ. Math. IHES **76** (1992) 1–97.

[KSS89] Kraft, H.; Slodowy, P.; Springer, A. T.: *Algebraische Transformationsgruppen und Invariantentheorie*, DMV-Seminar, vol. 13, Birkhäuser Verlag, Basel–Boston, 1989.

[KrS94] Kraft, H.; Small, L.W.: *Invariant algebras and completely reducible representations.* Math. Research Letters **1** (1994) 297–307.

[Lak] Lakshmibai, V.: *Geometry of G/P, VI.* J. Algebra.

[Lak] Lakshmibai, V.: *Geometry of G/P, VII.* J. Algebra.

[Lak] Lakshmibai, V.: *Geometry of G/P, VIII.* J. Algebra.

[LMS79] Lakshmibai, V.; Musili, C.; Seshadri, C. S.: *Geometry of G/P, IV.* Proc. Indian Acad. Sci. Math. Sci. **87A** (1979) 279–362.

[LaS78] Lakshmibai, V.; Seshadri, C. S.: *Geometry of G/P-II (The work of DeConcini and Procesi and the basic conjectures).* Proc. Indian Acad. Sci. Math. Sci. **87** (1978) 1–54.

[LaS86] Lakshmibai, V.; Seshadri, C. S.: *Geometry of G/P, V.* J. Algebra **100** (1986) 462–557.

[Mat89] Matsumura, H.: *Commutative algebra*, Cambridge Univ. Press, Cambridge, 1989.

[Mey99] Meyer, F.: *Invariantentheorie.* In: Encyklopädie der mathematischen Wissenschaften. vol. IB2, 1899, pp. 320–403.

[Mey92] Meyer, W. Fr.: *Bericht über den gegenwärtigen Stand der Invariantentheorie.* Jahresber. Deutsch. Math.-Vereinigung **1** (1892fborel).

[Mil87] Miller, W.: *The maximum order of an element of a finite symmetric group.* Math. Monthly **94** (1987) 497–506.

[Mum88] Mumford, D.: *The Red Book of varieties and schemes*, Lecture Notes in Math., vol. 1358, Springer Verlag, Berlin–Heidelberg–New York, 1988.

[MFK94] Mumford, D.; Fogarty, J., Kirwan, F.: *Geometric Invariant Theory*, Ergebnisse der Mathematik und Grenzgebiete, vol. 34, Springer-Verlag, New York–Heidelberg–Berlin, 1994. (3rd edition)

[Noe16] Noether, E.: *Der Endlichkeitssatz der Invarianten endlicher Gruppen.* Math. Ann. **77** (1916) 89–92.

[OnV90] Onishchik, A.L.; Vinberg, E.B.: *Seminar on Algebraic Groups and Lie Groups*, Springer-Verlag, New York–Heidelberg–Berlin, 1990.

[Pop86] Popov, V. L.: *Modern developments in invariant theory.* In: Proc. International Congress of Mathematicians in Berkeley, USA, 1986.

[Pop82] Popov, V. L.: *The constructive theory of invariants.* Math. USSR-Izv. **19** (1982) 359–376. (erste Fassung: Constructive invariant theory, in: Young tableaux and Schur functors in algebra and geometry, Astérisque 87/88 (1981), 303–334)

[Pop84] Popov, V. L.: *Syzygies in the theory of invariants.* Math. USSR-Izv. **22** (1984) 507–585.

[Pop92] Popov, V. L.: *Groups, Generators, Syzygies, and Orbits in Invariant Theory*, Transl. Math. Monographs, vol. 100, Amer. Math. Soc., Providence, 1992.

[Pro82] Procesi, C.: *A Primer of Invariant Theory*, Brandeis Lecture Notes, vol. 1, Brandeis University, Waltham MA, 1982.

[Pro84] Procesi, C.: *Computing with* 2×2 *matrices.* J. Algebra **87** (1984) 342–359.

[Sch89] Schmid, B.: *Generating invariants of finite groups.* C. R. Acad. Sci. Paris **308** (1989) 1–6.

[Sch91] Schmid, B.: *Finite groups and invariant theory.* In: Séminaire d'Algèbre Paul Dubreil et M.-P. Malliavin., Lecture Notes in Math. vol. 1478, Springer Verlag, Berlin–Heidelberg–New York, 1991, pp. 35–66.

[Sch68] Schur, I.: *Vorlesungen über Invariantentheorie*, Grundlehren Math. Wiss., vol. 143, Springer-Verlag, New York–Heidelberg–Berlin, 1968. (bearbeitet und herausgegeben von H. Grunsky)

[Ser73] Servedio, F. J.: *Prehomogeneous vector spaces and varieties.* Trans. Amer. Math. Soc. **176** (1973) 421–444.

[Spr77] Springer, T. A.: *Invariant theory*, Lecture Notes in Mathematics, vol. 585, Springer-Verlag, New York–Heidelberg–Berlin, 1977.

[Spr89] Springer, T. A.: *Aktionen reduktiver Gruppen auf Varietäten.* In: Algebraische Transformationsgruppen und Invariantentheorie (H. Kraft, P. Slodowy, T. A. Springer, eds.)., DMV-Seminar vol. 13, Birkhäuser Verlag, Basel–Boston, 1989, pp. 3–39.

[Stu89] Study, E.: *Methoden zur Theorie der ternären Formen*, Leipzig, 1889.

[Tit83] Tits, J.: *Liesche Gruppen und Algebren*, Hochschultext, Springer-Verlag, New York–Heidelberg–Berlin, 1983.

[ViP72] Vinberg, E. B.; Popov, V. L.: *On a class of quasihomogeneous affine varieties.* Math. USSR-Izv. **6** (1972) 743–758.

[Wei22] Weitzenböck, R.: *Neuere Arbeiten der algebraischen Invariantentheorie. Differentialinvarianten.* In: Encyklopädie der mathematischen Wissenschaften. vol. III E1, 1922

[Wei23] Weitzenböck, R.: *Invariantentheorie*, Noordhoff, Groningen, 1923.

[Wei32] Weitzenböck, R.: *Über die Invarianten von linearen Gruppen.* Acta Math. **58** (1932) 230–250.

[Wey46] Weyl, H.: *The Classical Groups, their Invariants and Representations*, Princeton Mathematical Series, vol. 1, Princeton Univ. Press, Princeton, 1946.

Index