Invariant Theory with Applications

Jan Draisma and Dion Gijswijt

October 8 2009

Contents

Chapter 1

Lecture 1. Introducing invariant theory

The first lecture gives some flavor of the theory of invariants. Basic notions such as (linear) group representation, the ring of regular functions on a vector space and the ring of invariant functions are defined, and some instructive examples are given.

1.1 Polynomial functions

Let V be a complex vector space. We denote by $V^* := \{f : V \to \mathbb{C} \text{ linear map}\}$ the dual vector space. Viewing the elements of V^* as functions on V, and taking the usual pointwise product of functions, we can consider the algebra of all $\mathbb{C}$-linear combinations of products of elements from V^*.

Definition 1.1.1. The *coordinate ring* $\mathcal{O}(V)$ of the vectorspace V is the algebra of functions $F : V \to \mathbb{C}$ generated by the elements of V^*. The elements of $\mathcal{O}(V)$ are called *polynomial* or *regular* functions on V.

If we fix a basis $e_1, \ldots, e_n$ of V, then a dual basis of V^* is given by the coordinate functions $x_1, \ldots, x_n$ defined by $x_i(c_1e_1 + \cdots + c_ne_n) := c_i$. For the coordinate ring we obtain $\mathcal{O}(V) = \mathbb{C}[x_1, \ldots, x_n]$. This is a polynomial ring in the x_i, since our base field $\mathbb{C}$ is infinite.

Exercise 1.1.2. Show that indeed $\mathbb{C}[x_1, \ldots, x_n]$ is a polynomial ring. In other words, show that the x_i are algebraically independent over $\mathbb{C}$: there is no nonzero polynomial $p \in \mathbb{C}[X_1, \ldots, X_n]$ in n variables $X_1, \ldots, X_n$, such that $p(x_1, \ldots, x_n) = 0$. Hint: this is easy for the case $n = 1$. Now use induction on n.

We call a regular function $f \in \mathcal{O}(V)$ *homogeneous of degree d* if $f(tv) = t^d f(v)$ for all $v \in V$ and $t \in \mathbb{C}$. Clearly, the elements of V^* are regular of degree

1, and the product of polynomials f, g homogeneous of degrees d, d' yields a homogeneous polynomial of degree $d + d'$. It follows that every regular function f can be written as a sum $f = c_0 + c_1 f_1 + \cdots + c_k f_k$ of regular functions f_i homogeneous of degree i. This decomposition is unique (disregarding the terms with zero coefficient). Hence we have a direct sum decomposition $\mathcal{O}(V) = \bigoplus_{d \in \mathbb{N}} \mathcal{O}(V)_d$, where $\mathcal{O}(V)_d := \{f \in \mathcal{O}(V) \mid f \text{ homogeneous of degree } d\}$, making $\mathcal{O}(V)$ into a *graded algebra*.

Exercise 1.1.3. Show that indeed the decomposition of a regular function f into its homogeneous parts is unique.

In terms of the basis $x_1, \ldots, x_n$, we have $\mathcal{O}(V)_d = \mathbb{C}[x_1, \ldots, x_n]_d$, where $\mathbb{C}[x_1, \ldots, x_n]_d$ consists of all polynomials of total degree d and has as basis the monomials $x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n}$ for $d_1 + d_2 + \cdots + d_n = d$.

1.2 Representations

Central objects in this course are linear representations of groups. For any vector space V we write $\mathrm{GL}(V)$ for the group of all invertible linear maps from V to itself. When we have a fixed basis of V, we may identify V with $\mathbb{C}^n$ and $\mathrm{GL}(V)$ with the set of invertible matrices $n \times n$ matrices $\mathrm{GL}(\mathbb{C}^n) \subset \mathrm{Mat}_n(\mathbb{C})$.

Definition 1.2.1. Let G be a group and let X be a set. An *action of G on X* is a map $\alpha : G \times X \to X$ such that $\alpha(1, x) = x$ and $\alpha(g, \alpha(h, x)) = \alpha(gh, x)$ for all $g, h \in G$ and $x \in X$.

If α is clear from the context, we will usually write gx instead of $\alpha(g, x)$. What we have just defined is sometimes called a *left action* of G on X; *right actions* are defined similarly.

Definition 1.2.2. If G acts on two sets X and Y, then a map $\phi : X \to Y$ is called *G-equivariant* if $\phi(gx) = g\phi(x)$ for all $x \in X$ and $g \in G$. As a particular case of this, if X is a subset of Y satisfying $gx \in X$ for all $x \in X$ and $g \in G$, then X is called *G-stable*, and the inclusion map is G-equivariant.

Example 1.2.3. The symmetric group S_4 acts on the set $\binom{[4]}{2}$ of unordered pairs of distinct numbers in $[4] := \{1, 2, 3, 4\}$ by $g\{i, j\} = \{g(i), g(j)\}$. Think of the edges in a tetrahedron to visualise this action. The group S_4 also acts on the set $X := \{(i, j) \mid i, j \in [4] \text{ distinct}\}$ of all ordered pairs by $g(i, j) = (g(i), g(j))$—think of directed edges—and the map $X \to \binom{[4]}{2}$ sending (i, j) to $\{i, j\}$ is S_4-equivariant.

Definition 1.2.4. Let G be a group and let V be a vector space. A *(linear) representation of G on V* is a group homomorphism $\rho : G \to \mathrm{GL}(V)$.

If ρ is a representation of G, then the map $(g, v) \mapsto \rho(g)v$ is an action of G on V. Conversely, if we have an action α of G on V such that $\alpha(g, .) : V \to V$ is a linear map for all $g \in G$, then the map $g \mapsto \alpha(g, .)$ is a linear representation.

As with actions, instead of $\rho(g)v$ we will often write gv. A vector space with an action of G by linear maps is also called a *G-module*.

Given a linear representation $\rho : G \to \mathrm{GL}(V)$, we obtain a linear representation $\rho^* : G \to \mathrm{GL}(V^*)$ on the dual space V^*, called the *dual representation* or *contragredient representation* and defined by

$$(\rho^*(g)x)(v) := x(\rho(g)^{-1}v) \text{ for all } g \in G,\ x \in V^* \text{ and } v \in V. \tag{1.1}$$

Exercise 1.2.5. Let $\rho : G \to \mathrm{GL}_n(\mathbb{C})$ be a representation of G on $\mathbb{C}^n$. Show that with respect to the dual basis, ρ^* is given by $\rho^*(g) = (\rho(g)^{-1})^\mathsf{T}$, where A^T denotes the transpose of the matrix A.

1.3 Invariant functions

Definition 1.3.1. Given a representation of a group G on a vector space V, a regular function $f \in \mathcal{O}(V)$ is called *G-invariant* or simply *invariant* if $f(v) = f(gv)$ for all $g \in G, v \in V$. We denote by $\mathcal{O}(V)^G \subseteq \mathcal{O}(V)$ the subalgebra of invariant functions. The actual representation of G is assumed to be clear from the context.

Observe that $f \in \mathcal{O}(V)$ is invariant, precisely when it is constant on the orbits of V under the action of G. In particular, the constant functions are invariant.

The representation of G on V induces an action on the (regular) functions on V by defining $(gf)(v) := f(g^{-1}v)$ for all $g \in G, v \in V$. This way the invariant ring can be discribed as the set of regular functions fixed by the action of G: $\mathcal{O}(V)^G = \{f \in \mathcal{O}(V) \mid gf = f \text{ for all } g \in G\}$. Observe that when restricted to $V^* \subset \mathcal{O}(V)$, this action coincides with the action corresponding to the dual representation. In terms of a basis $x_1, \ldots, x_n$ of V^*, the regular functions are polynomials in the x_i and the action of G is given by $gp(x_1, \ldots, x_n) = p(gx_1, \ldots, gx_n)$ for any polynomial p. Since for every d, G maps the set of polynomials homogeneous of degree d to itself, it follows that the homogeneous parts of an invariant are invariant as well. This shows that $\mathcal{O}(V)^G = \bigoplus_d \mathcal{O}(V)^G_d$, where $\mathcal{O}(V)^G_d := \mathcal{O}(V)_d \cap \mathcal{O}(V)^G$.

Example 1.3.2. Consider the representation $\rho : \mathbb{Z}/3\mathbb{Z} \to \mathrm{GL}_2(\mathbb{C})$ defined by mapping 1 to the matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right)$ (and mapping 2 to $\left(\begin{smallmatrix} -1 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and 0 to the identity matrix). With respect to the dual basis x_1, x_2, the dual representation is given by:

$$\rho^*(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \rho^*(1) = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \rho^*(2) = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}. \tag{1.2}$$

The polynomial $f = x_1^2 - x_1x_2 + x_2^2$ is an invariant:

$$\rho^*(1)f = (-x_1 + x_2)^2 - (-x_1 + x_2)(-x_1) + (-x_1)^2 = x_1^2 - x_1x_2 + x_2^2 = f, \tag{1.3}$$

and since 1 is a generator of the group, f is invariant under all elements of the group. Other invariants are $x_1^2x_2 - x_1x_2^2$ and $x_1^3 - 3x_1x_2^2 + x_2^3$. These three invariants generate the ring of invariants, althought it requires some work to show that.

A simpler example in which the complete ring of invariants can be computed is the following.

Example 1.3.3. Let D_4 be the symmetry group of the square, generated by a rotation r, a reflection s and the relations $r^4 = e, s^2 = e$ and $srs = r^3$, where e is the identity. The representation ρ of D_4 on $\mathbb{C}^2$ is given by

$$\rho(r) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \rho(s) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{1.4}$$

the dual representation is given by the same matrices:

$$\rho^*(r) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \rho^*(s) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.5}$$

It is easy to check that $x_1^2+x_2^2$ and $x_1^2x_2^2$ are invariants, and so are all polynomial expressions in these two invariants. We will show that in fact $\mathcal{O}(\mathbb{C}^2)^{D_4} = \mathbb{C}[x_1^2 + x_2^2, x_1^2x_2^2] =: R$. It suffices to show that all homogeneous invariants belong to R.

Let $p \in \mathbb{C}[x_1, x_2]$ be a homogeneous invariant. Since $sp = p$, only monomials having even exponents for x_1 can occur in p. Since r^2s exchanges x_1 and x_2, for every monomial $x_1^a x_2^b$ in p, the monomial $x_1^b x_2^a$ must occur with the same exponent. This proves the claim since every polynomial of the form $x_1^{2n}x_2^{2m} + x_1^{2m}x_2^{2n}$ is an element of R. Indeed, we may assume that $n \leq m$ and proceed by induction on $n+m$, the case $n+m = 0$ being trivial. If $n > 0$ we have $q = (x_1^2x_2^2)^n(x_2^{2m-2n} + x_1^{2m-2n})$ and we are done. If $n = 0$ we have $2q = 2(x_1^{2m} + x_2^{2m}) = 2(x_1^2+x_2^2)^m - \sum_{i=1}^{m-1} \binom{m}{i}(x_1^{2i}x_2^{2m-2i})$ and we are done by induction again.

1.4 Conjugacy classes of matrices

In this section we discuss the polynomial functions on the square matrices, invariant under conjugation of the matrix variable by elements of $\mathrm{GL}_n(\mathbb{C})$. This example shows some tricks that are useful when proving that certain invariants are equal. Denote by $M_n(\mathbb{C})$ the vectorspace of complex $n \times n$ matrices. We consider the action of $G = \mathrm{GL}_n(\mathbb{C})$ on $M_n(\mathbb{C})$ by conjugation: $(g, A) \mapsto gAg^{-1}$ for $g \in \mathrm{GL}_n(\mathbb{C})$ and $A \in M_n(\mathbb{C})$. We are interested in finding all polynomials in the entries of $n \times n$ matrices that are invariant under G. Two invariants are given by the functions $A \mapsto \det A$ and $A \mapsto \mathrm{tr} A$.

Let

$$\chi_A(t) := \det(tI - A) = t^n - s_1(A)t^{n-1} + s_2(A)t^{n-2} - \cdots + (-1)^n s_n(A) \tag{1.6}$$

be the characteristic polynomial of A. Here the s_i are polynomials in the entries of A. Clearly,

$$\chi_{gAg^{-1}}(t) = \det(g(tI - A)g^{-1}) = \det(tI - A) = \chi_A(t) \tag{1.7}$$

holds for all $t \in \mathbb{C}$. It follows that the functions $s_1, \ldots, s_n$ are G-invariant. Observe that $s_1(A) = \mathrm{tr} A$ and $s_n(A) = \det A$.

Proposition 1.4.1. *The functions $s_1, \ldots, s_n$ generate $\mathcal{O}(\mathrm{Mat}_n(\mathbb{C}))^{\mathrm{GL}_n(\mathbb{C})}$ and are algebraically independent.*

Proof. To each $c = (c_1, \ldots, c_n \in \mathbb{C}^n$ we associate the so-called *companion matrix*

$$A_c := \begin{pmatrix} 0 & \cdots & \cdots & 0 & -c_n \\ 1 & \ddots & & \vdots & -c_{n-1} \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & c_2 \\ 0 & \cdots & 0 & 1 & c_1 \end{pmatrix} \in M_n(\mathbb{C}). \tag{1.8}$$

A simple calculation shows that $\chi_{A_c}(t) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$.

Exercise 1.4.2. Verify that $\chi_{A_c}(t) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$.

This implies that $s_i(A_c) = (-1)^i c_i$ and therefore

$$\{(s_1(A_c), s_2(A_c), \ldots, s_n(A_c) \mid A \in M_n(\mathbb{C})\} = \mathbb{C}^n. \tag{1.9}$$

It follows that the s_i are algebraically independent over $\mathbb{C}$. Indeed, suppose that $p(s_1, \ldots, s_n) = 0$ for some polynomial p in n variables. Then

$$0 = p(s_1, \ldots, s_n)(A) = p(s_1(A), \ldots, s_n(A)) \tag{1.10}$$

for all A and hence $p(c_1, \ldots, c_n) = 0$ for all $c \in \mathbb{C}^n$. But this implies that p itself is the zero polynomial.

Now let $f \in \mathcal{O}(\mathrm{Mat}_n(\mathbb{C}))^G$ be an invariant function. Define the polynomial p in n variables by $p(c_1, \ldots, c_n) := f(A_c)$, and $P \in \mathcal{O}(\mathrm{Mat}_n(\mathbb{C}))^G$ by $P(A) := p(-s_1(A), s_2(A), \ldots, (-1)^n s_n(A))$. By definition, P and f agree on all companion matrices, and since they are both G-invariant they agree on $W := \{gA_cg^{-1} \mid g \in G, c \in \mathbb{C}^n\}$. To finish the proof, it suffices to show that W is dense in $\mathrm{Mat}_n(\mathbb{C})$ since $f - P$ is continuous and zero on W. To show that W is dense in $\mathcal{O}(\mathrm{Mat}_n(\mathbb{C}))$, it suffices to show that the set of matrices with n distinct nonzero eigenvalues is a subset of W and is itself dense in $\mathcal{O}(\mathrm{Mat}_n(\mathbb{C}))$. This we leave as an exercise.

Exercise 1.4.3. Let $A \in \mathrm{Mat}_n(\mathbb{C})$ have n distinct nonzero eigenvalues. Show that A is conjugate to A_c for some $c \in \mathbb{C}^n$. Hint: find $v \in \mathbb{C}^n$ such that

$v, Av, A^2v, \ldots, A^{n-1}v$ is a basis for $\mathbb{C}^n$. You might want to use the fact that the Vandermonde determinant

$$\det \begin{pmatrix} 1 & \ldots & 1 \\ c_1 & \ldots & c_n \\ c_1^2 & \ldots & c_n^2 \\ \vdots & \ddots & \vdots \\ c_1^{n-1} & \ldots & c_n^{n-1} \end{pmatrix} \tag{1.11}$$

is nonzero if $c_1, \ldots, c_n$ are distinct and nonzero.

Exercise 1.4.4. Show that the set of matrices with n distinct nonzero eigenvalues is dense in the set of all complex $n \times n$ matrices. Hint: every matrix is conjugate to an upper triangular matrix.

□

1.5 Exercises

Exercise 1.5.1. Let G be a finite group acting on $V = \mathbb{C}^n$, $n \geq 1$. Show that $\mathcal{O}(V)^G$ contains a nontrivial invariant. That is, $\mathcal{O}(V)^G \neq \mathbb{C}$. Give an example of an action of an infinite group G on V with the property that only the constant functions are invariant.

Exercise 1.5.2. Let $\rho : \mathbb{Z}/2\mathbb{Z} \to \mathrm{GL}_2(\mathbb{C})$ be the representation given by $\rho(1) := \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Compute the invariant ring. That is, give a minimal set of generators for $\mathcal{O}(\mathbb{C}^2)^{\mathbb{Z}/2\mathbb{Z}}$.

Exercise 1.5.3. Let $U := \{\left(\begin{smallmatrix} 1 & a \\ 0 & 1 \end{smallmatrix}\right) \mid a \in \mathbb{C}\}$ act on $\mathbb{C}^2$ in the obvious way. Denote the coordinate functions by x_1, x_2. Show that $\mathcal{O}(\mathbb{C}^2)^U = \mathbb{C}[x_2]$.

Exercise 1.5.4. Let $\rho : \mathbb{C}^* \to \mathrm{GL}_3(\mathbb{C})$ be the representation given by $\rho(t) = \left(\begin{smallmatrix} t^{-2} & 0 & 0 \\ 0 & t^{-3} & 0 \\ 0 & 0 & t^4 \end{smallmatrix}\right)$. Find a minimal system of generators for the invariant ring.

Exercise 1.5.5. Let $\pi : \mathrm{Mat}_n(\mathbb{C}) \to \mathbb{C}^n$ be given by $\pi(A) := (s_1(A), \ldots, s_n(A))$. Show that for every $c \in \mathbb{C}^n$ the fiber $\{A \mid \pi(A) = c\}$ contains a unique conjugacy class $\{gAg^{-1} \mid g \in \mathrm{GL}_n(\mathbb{C})\}$ of a diagonalizable (semisimple) matrix A.

Chapter 2

Lecture 2. Symmetric polynomials

In this chapter, we consider the natural action of the symmetric group S_n on the ring of polynomials in the variables $x_1, \ldots, x_n$. The fundamental theorem of symmetric polynomials states that the elementary symmetric polynomials generate the ring of invariants. As an application we prove a theorem of Sylvester that characterizes when a univariate polynomial with real coefficients has only real roots.

2.1 Symmetric polynomials

Let the group S_n act on the polynomial ring $\mathbb{C}[x_1, \ldots, x_n]$ by permuting the variables:

$$\sigma p(x_1, \ldots, x_n) := p(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \text{ for all } \sigma \in S_n. \tag{2.1}$$

The polynomials invariant under this action of S_n are called *symmetric polynomials*. As an example, for $n = 3$ the polynomial $x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2 + 7x_1 + 7x_2 + 7x_3$ is symmetric, but $x_1^2x_2 + x_1x_3^2 + x_2^2x_3$ is not symmetric (although it is invariant under the alternating group).

In terms of linear representations of a group, we have a linear representation $\rho : S_n \to \mathrm{GL}_n(\mathbb{C})$ given by $\rho(\sigma)e_i := e_{\sigma(i)}$, where $e_1, \ldots, e_n$ is the standard basis of $\mathbb{C}^n$. On the dual basis $x_1, \ldots, x_n$ the dual representation is given by $\rho^*(\sigma)x_i = x_{\sigma(i)}$, as can be easily checked. The invariant polynomial functions on $\mathbb{C}^n$ are precisely the symmetric polynomials.

Some obvious examples of symmetric polynomials are

$$\begin{aligned} s_1 &:= x_1 + x_2 + \cdots + x_n \text{ and} && (2.2)\\ s_2 &:= x_1x_2 + x_1x_3 + \cdots + x_1x_n + \cdots + x_{n-1}x_n && (2.3) \end{aligned}$$

More generally, for every $k = 1, \ldots, n$, the *k-th elementary symmetric polyno-*

mial

$$s_k := \sum_{i_1 < \ldots < i_k} x_{i_1} \cdots x_{i_k} \tag{2.4}$$

is invariant. Recall that these polynomials express the coefficients of a univariate polynomial in terms of its roots:

$$\prod_{i=1}^{n}(t - x_i) = x^n + \sum_{k=1}^{n}(-1)^k s_k t^{n-k}. \tag{2.5}$$

Moreover, if g is any polynomial in n variables $y_1, \ldots, y_n$, then $g(s_1, \ldots, s_n)$ is again a polynomial in the x_i which is invariant under all coordinate permutations. A natural question is: which symmetric polynomials are expressible as a polynomial in the elementary symmetric polynomials. For example $x_1^2 + \cdots + x_n^2$ is clearly symmetric and it can be expressed in terms of the s_i:

$$x_1^2 + \cdots + x_n^2 = s_1^2 - 2s_2. \tag{2.6}$$

It is a beautiful fact that the elementary symmetric polynomials generate *all* symmetric polynomials.

Theorem 2.1.1 (Fundamental theorem of symmetric polynomials)**.** *Every S_n-invariant polynomial $f(x_1, \ldots, x_n)$ in the x_i can be written as $g(s_1, \ldots, s_n)$, where $g = g(y_1, \ldots, y_n)$ is a polynomial in n variables. Moreover, given f, the polynomial g is unique.*

The proof of this result uses the *lexicographic order* on monomials in the variables $\underline{x} = (x_1, \ldots, x_n)$. We say that $\underline{x}^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ is (lexicographically) larger than $\underline{x}^\beta$ if there is a k such that $\alpha_k > \beta_k$ and $\alpha_i = \beta_i$ for all $i < k$. So for instance $x_1^2 > x_1 x_2^4 > x_1 x_2^3 > x_1 x_2 x_3^5$, etc. The *leading monomial* $\mathrm{lm}(f)$ of a non-zero polynomial f in the x_i is the largest monomial (with respect to this ordering) that has non-zero coefficient in f.

Exercise 2.1.2. Check that $\mathrm{lm}(fg) = \mathrm{lm}(f)\mathrm{lm}(g)$ and that $\mathrm{lm}(s_k) = x_1 \cdots x_k$.

Exercise 2.1.3. Show that there are no infinite lexicographically strictly decreasing chains of monomials.

Since every decreasing chain of monomials is finite, we can use this order to do induction on monomials, as we do in the following proof.

Proof of Theorem 2.1.1. Let f be any S_n-invariant polynomial in the x_i. Let $\underline{x}^\alpha$ be the leading monomial of f. Then $\alpha_1 \geq \ldots \geq \alpha_n$ because otherwise a suitable permutation applied to $\underline{x}^\alpha$ would yield a lexicographically larger monomial, which has the same non-zero coefficient in f as $\underline{x}^\alpha$ by invariance of f. Now consider the expression

$$s_n^{\alpha_n} s_{n-1}^{\alpha_{n-1}-\alpha_n} \cdots s_1^{\alpha_1-\alpha_2}. \tag{2.7}$$

The leading monomial of this polynomial equals

$$(x_1 \cdots x_n)^{\alpha_n} (x_1 \cdots x_{n-1})^{\alpha_{n-1}-\alpha_n} \cdots x_1^{\alpha_1-\alpha_2}, \tag{2.8}$$

which is just $\underline{x}^\alpha$. Subtracting a scalar multiple of the expression from f therefore cancels the term with monomial $\underline{x}^\alpha$, and leaves an S_n-invariant polynomial with a strictly smaller leading monomial. After repeating this step finitely many times, we have expressed f as a polynomial in the s_k.

This shows existence of g in the theorem. For uniqueness, let $g \in \mathbb{C}[y_1, \ldots, y_n]$ be a nonzero polynomial in n variables. It suffices to show that $g(s_1, \ldots, s_n) \in \mathbb{C}[x_1, \ldots, x_n]$ is not the zero polynomial. Observe that

$$\mathrm{lm}(s_1^{\alpha_1} \cdots s_n^{\alpha_n}) = x_1^{\alpha_1+\cdots+\alpha_n} x_2^{\alpha_2+\cdots+\alpha_n} \cdots x_n^{\alpha_n}. \tag{2.9}$$

It follows that the leading monomials of the terms $s_1^{\alpha_1} \cdots s_n^{\alpha_n}$, corresponding to the monomials occuring with nonzero coefficient in $g = \sum_\alpha \underline{y}^\alpha$, are pairwise distinct. In particular, the largest such leading monomial will not be cancelled in the sum and is the leading monomial of $g(s_1, \ldots, s_n)$. $\square$

Remark 2.1.4. The proof shows that in fact the coefficients of the polynomial g lie in the ring generated by the coefficients of f. In particular, when f has real coefficients, also g has real coefficients.

Exercise 2.1.5. Let $\pi : \mathbb{C}^n \to \mathbb{C}^n$ be given by

$$\pi(x_1, \ldots, x_n) = (s_1(x_1, \ldots, x_n), \ldots, s_n(x_1, \ldots, x_n)). \tag{2.10}$$

Use the fact that every univariate polynomial over the complex numbers can be factorised into linear factors to show that π is surjective. Use this to show that $s_1, \ldots, s_n$ are algebraically independent over $\mathbb{C}$. Describe for $b \in \mathbb{C}^n$ the fiber $\pi^{-1}(b)$.

Remark 2.1.6. The above proof of the fundamental theorem of symmetric polynomials gives an algorithm to write a given symmetric polynomial as a polynomial in the elementary symmetric polynomials. In each step the initial monomial of the residual symmetric polynomial is decreased, ending with the zero polynomial after a finite number of steps. Instead of using the described lexicographic order on the monomials, other linear orders can be used. An example would be the *degree lexicographic order*, where we set $\underline{x}^\alpha > \underline{x}^\beta$ if either $\alpha_1 + \cdots + \alpha_n > \beta_1 + \cdots + \beta_n$ or equality holds and there is a k such that $\alpha_k > \beta_k$ and $\alpha_i = \beta_i$ for all $i < k$.

Example 2.1.7. We write $x_1^3 + x_2^3 + x_3^3$ as a polynomial in the s_i. Since the leading monomial is $x_1^3 x_2^0 x_3^0$ we subtract $s_3^0 s_2^0 s_1^3$ and are left with $-3(x_1^2 x_2 + x_1^2 x_3 + x_1 x_2^2 + x_1 x_3^2 + x_2^2 x_3 + x_2 x_3^2) - 6(x_1 x_2 x_3)$. The leading monomial is now $x_1^2 x_2$, so we add $3 s_3^0 s_2^1 s_1^{2-1}$. This leaves $3 x_1 x_2 x_3 = 3 s_3^1 s_2^{1-1} s_1^{1-1}$, which is reduced to zero in the next step.

This way we obtain $x_1^3 + x_2^3 + x_3^3 = s_1^3 - 3 s_1 s_2 + 3 s_3$.

Exercise 2.1.8. Give an upper bound on the number of steps of the algorithm in terms of the number of variables n and the (total) degree of the input polynomial f.

2.2 Counting real roots

Given a (monic) polynomial $f(t) = t^n - a_1 t^{n-1} + \cdots + (-1)^n a_n$, the coefficients are elementary symmetric functions in the roots of f. Therefore, any property that can be expressed as a symmetric polynomial in the roots of f, can also be expressed as a polynomial in the coefficients of f. This way we can determine properties of the roots by just looking at the coefficients of f. For example: when are all roots of f distinct?

Definition 2.2.1. For a (monic) polynomial $f = (t - x_1) \cdots (t - x_n)$, define the *discriminant* $\Delta(f)$ *of* f by $\Delta(f) := \prod_{1 \le i < j \le n} (x_i - x_j)^2$.

Clearly, $\Delta(f) = 0$ if and only if all roots of f are distinct. It is not hard to see that $\Delta(f)$ is a symmetric polynomial in the roots of f. We will see later how f can be expressed in terms of the coefficients of f.

Exercise 2.2.2. Let $f(t) = t^2 - at + b$. Write $\Delta(f)$ as a polynomial in a and b.

Definition 2.2.3. Given n complex numbers $x_1, \ldots, x_n$, the *Vandermonde matrix* A for these numbers is given by

$$A := \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix}. \tag{2.11}$$

Lemma 2.2.4. *Given numbers $x_1, \ldots, x_n$, the Vandermonde matrix A has nonzero determinant if and only if the $x_1, \ldots, x_n$ are distinct.*

Proof. View the determinant of the Vandermonde matrix (called the *Vandermonde determinant*) as a polynomial p in the variables $x_1, \ldots, x_n$. For any $i < j$, $p(x_1, \ldots, x_n) = 0$ when $x_i = x_j$ and hence p is divisible by $(x_j - x_i)$. Expanding the determinant, we see that p is homogeneous of degree $\binom{n}{2}$, with lowest monomial $x_1^0 x_2^1 \cdots x_n^{n-1}$ having coefficient 1. It follows that

$$p = \prod_{1 \le i < j \le n} (x_j - x_i), \tag{2.12}$$

since the right-hand side divides p, and the two polynomials have the same degree and the same nonzero coefficient for $x_1^0 x_2^1 \cdots x_n^{n-1}$. $\square$

Exercise 2.2.5. Show that the Vandermonde matrix A of numbers $x_1, \ldots, x_n$ satisfies $\det A = \prod_{1 \le i < j \le n} (x_j - x_i)$ by doing row- and column-operations on A and applying induction on n.

Definition 2.2.6. Let $f = (t - \alpha_1)(t - \alpha_2) \cdots (t - \alpha_n) \in \mathbb{C}[t]$ be a monic polynomial of degree n in the variable t. We define the *Bezoutiant matrix* $\mathrm{Bez}(f)$ *of* f by

$$\mathrm{Bez}(f) = (p_{i+j-2}(\alpha_1, \ldots, \alpha_n))_{i,j=1}^n, \tag{2.13}$$

where $p_k(x_1, \ldots, x_n) := x_1^k + \cdots + x_n^k$ for $k = 0, 1, \ldots$ is the k-th *Newton polynomial.*

Since the entries of $\mathrm{Bez}(f)$ are symmetric polynomials in the roots of f, it follows by the fundamental theorem of symmetric polynomials that the entries are polynomials (with integer coefficients) in the elementary symmetric functions and hence in the coefficients of f. In particular, when f has real coefficients, $\mathrm{Bez}(f)$ is a real matrix. Another useful fact is that $\mathrm{Bez}(f) = A^{\mathsf{T}}A$, where A is the Vandermonde matrix for the roots $\alpha_1, \ldots, \alpha_n$ of f.

Exercise 2.2.7. Show that the discriminant of f satisfies: $\Delta(f) = \det \mathrm{Bez}(f)$.

Example 2.2.8. Let $f = t^2 - at + b$ have roots α and β. So $a = \alpha + \beta$ and $b = \alpha\beta$. We compute $\mathrm{Bez}(f)$. We have $p_0 = 2$, $p_1 = a$, $p_2 = a^2 - 2b$ so $\mathrm{Bez}(f) = \left(\begin{smallmatrix} 2 & a \\ a & a^2-2b \end{smallmatrix}\right)$. The determinant equals $a^2 - 4b$ and the trace equals $a^2 - 2b + 2$. There are three cases for the eigenvalues $\lambda_1 \geq \lambda_2$ of $\mathrm{Bez}(f)$:

- If $a^2 - 4b > 0$, we have $\lambda_1, \lambda_2 > 0$ and α, β are distinct real roots.
- If $a^2 - 4b = 0$, we have $\lambda_1 > 0, \lambda_2 = 0$ and $\alpha = \beta$.
- If $a^2 - 4b < 0$, we have $\lambda_1 > 0, \lambda_2 < 0$ and α and β are complex conjugate (nonreal) roots.

The determinant of $\mathrm{Bez}(f)$ determines whether f has double roots. The matrix $\mathrm{Bez}(f)$ can give us much more information about the roots of f. In particular, it describes when a polynomial with real coefficients has only real roots!

Theorem 2.2.9 (Sylverster)**.** *Let $f \in \mathbb{R}[t]$ be a polynomial in the variable t with real coefficients. Let r be the number of distinct roots in $\mathbb{R}$ and $2k$ the number of distinct roots in $\mathbb{C} \setminus \mathbb{R}$. Then the Bezoutiant matrix $\mathrm{Bez}(f)$ has rank $r + 2k$, with $r + k$ positive eigenvalues and k negative eigenvalues.*

proof of Theorem 2.2.9. Number the roots $\alpha_1, \ldots, \alpha_n$ of f in such a way that $\alpha_1, \ldots, \alpha_{2k+r}$ are distinct. We write m_i for the multiplicity of the root α_i, $i = 1, \ldots, 2k+r$. Let A be the Vandermonde matrix for the numbers $\alpha_1, \ldots, \alpha_n$, so that $\mathrm{Bez}(f) = A^{\mathsf{T}}A$. We start by computing the rank of $\mathrm{Bez}(f)$.

Denote by $\tilde{A}$ the $(2k + r) \times n$ submatrix of A consisting of the first $2k + r$ rows of A. An easy computation shows that

$$\mathrm{Bez}(f) = A^{\mathsf{T}}A = \tilde{A}^{\mathsf{T}} \operatorname{diag}(m_1, \ldots, m_{2k+r})\tilde{A}, \tag{2.14}$$

where $\operatorname{diag}(m_1, \ldots, m_{2k+r})$ is the diagonal matrix with the multiplicities of the roots on the diagonal. Since, $\tilde{A}$ contains a submatrix equal to the Vandermonde matrix for the distinct roots $\alpha_1, \ldots, \alpha_{2k+r}$, it follows by Lemma 2.2.4 that the rows of $\tilde{A}$ are linearly independent. Since the diagonal matrix has full rank, it follows that $\mathrm{Bez}(f)$ has rank $2k + r$.

To finish the proof, we write $A = B + iC$, where B and C are real matrices and i denotes a square root of -1. Since f has real coefficients, $\mathrm{Bez}(f)$ is a real matrix and hence

$$\mathrm{Bez}(f) = B^{\mathsf{T}}B - C^{\mathsf{T}}C + i(C^{\mathsf{T}}B + B^{\mathsf{T}}C) = B^{\mathsf{T}}B - C^{\mathsf{T}}C. \tag{2.15}$$

We have

$$\operatorname{rank}(B) \le r + k, \qquad \operatorname{rank}(C) \le k. \tag{2.16}$$

Indeed, for any pair $\alpha, \overline{\alpha}$ of complex conjugate numbers, the real parts of α^j and $\overline{\alpha}^j$ are equal and the imaginary parts are opposite. Hence B has at most $r + k$ different rows and C has (up to a factor -1) at most k different nonzero rows.

Denote the kernel of $\mathrm{Bez}(f)$, B and C by N, N_B and N_C respectively. Clearly $N_B \cap N_C \subseteq N$. This gives

$$\begin{aligned} \dim N \ge \dim(N_B \cap N_C) &\ge \dim N_B + \dim N_C - n \\ &\ge (n - r - k) + (n - k) - n \\ &= n - r - 2k = \dim N. \end{aligned} \tag{2.17}$$

Hence we have equality throughout, showing that $\dim N_B = n-r-k, \dim N_C = n - k$ and $N_B \cap N_C = N$.

Write $N_B = N \oplus N'_B$ and $N_C = N \oplus N'_C$ as a direct sum of vector spaces. For all nonzero $u \in N'_C$, we have $u^\mathsf{T} C^\mathsf{T} C u = 0$ and $u^\mathsf{T} B^\mathsf{T} B u > 0$ and so $u^\mathsf{T} \mathrm{Bez}(f) u > 0$. This shows that $\mathrm{Bez}(f)$ has at least $\dim N'_C = r + k$ positive eigenvalues (see exercises). Similarly, $u^\mathsf{T} \mathrm{Bez}(f) u < 0$ for all nonzero $u \in N'_B$ so that $\mathrm{Bez}(f)$ has at least $\dim N'_B = k$ negative eigenvalues. Since $\mathrm{Bez}(f)$ has $n - r - 2k$ zero eigenvalues, it has exactly $r + k$ positive eigenvalues and exactly k negative eigenvalues. □

Exercise 2.2.10. Let B be a real $n \times n$ matrix and $x \in \mathbb{R}^n$. Show that $x^\mathsf{T} B^\mathsf{T} B x \ge 0$ and that equality holds if and only if $Bx = 0$.

Exercise 2.2.11. Let A be a real symmetric $n \times n$ matrix. Show that the following are equivalent:

- there exists a linear subspace $V \subseteq \mathbb{R}^n$ of dimension k such that $x^\mathsf{T} A x > 0$ for all nonzero $x \in V$,
- A has at least k positive eigenvalues.

Exercise 2.2.12. Use the previous exercise to show Sylvesters law of inertia: Given a real symmetric $n \times n$ matrix A and an invertible real matrix S, the two matrices A and $S^\mathsf{T} A S$ have the same number of positive, negative and zero eigenvalues. This implies that the signature of A can be easily determined by bringing it into diagonal form using simultaneous row and column operations.

2.3 Exercises

Exercise 2.3.1. Let $f(t) := t^3 + at + b$, where a, b are real numbers.

- Compute $\mathrm{Bez}(f)$.
- Show that $\Delta(f) = -4a^3 - 27b^2$.

- Determine, in terms of a and b, when f has only real roots.

Exercise 2.3.2. Prove the following formulas due to Newton:

$$p_k - s_1 p_{k-1} + \cdots + (-1)^{k-1} s_{k-1} p_1 + (-1)^k k s_k = 0 \tag{2.18}$$

for all $k = 1, \ldots, n$.

Show that for $k > n$ the following similar relation holds:

$$p_k - s_1 p_{k-1} + \cdots + (-1)^n s_n p_{k-n} = 0. \tag{2.19}$$

Hint: Let $f(t) = (1 - tx_1) \cdots (1 - tx_n)$ and compute $f'(t)/f(t)$ in two ways.

Chapter 3

Lecture 3. Multilinear algebra

We review some constructions from linear algebra, in particular the tensor product of vector spaces. Unless explicitly stated otherwise, all our vector spaces are over the field $\mathbb{C}$ of complex numbers.

Definition 3.0.3. Let $V_1, \ldots, V_k, W$ be vector spaces. A map $\phi : V_1 \times \cdots \times V_k \to W$ is called *multilinear* (or *k-linear* or bilinear if $k = 2$ or trilinear if $k = 3$) if for each i and all $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$ the map $V_i \to W,\ v_i \mapsto \phi(v_1, \ldots, v_k)$ is linear.

Let U, V and T be vector spaces and let $\otimes : U \times V \to T$ be a bilinear map. The map $\otimes$ is said to have the *universal property* if for every bilinear map $\phi : U \times V \to W$ there exists a *unique* linear map $f : T \to W$ such that $\phi = f \circ \otimes$.

$$\begin{tikzcd} U \times V \arrow[r, "\phi"] \arrow[d, "\otimes"'] & W \\ T \arrow[ur, "f"'] & \end{tikzcd}$$

We will usually write $u \otimes v := \otimes(u, v)$ for $(u, v) \in U \times V$. Although $\otimes$ will in general not be surjective, the image linearly spans T.

Exercise 3.0.4. Show that if $\otimes : U \times V \to T$ has the universal property, the vectors $u \otimes v, u \in U, v \in V$ span T.

Given U and V, the pair $(T, \otimes)$ is *unique up to a unique isomorphism*. That is, given two bilinear maps $\otimes : U \times V \to T$ and $\otimes' : U \times V \to T'$ that both have the universal property, there is a unique linear isomorphism $f : T \to T'$ such that $f(u \otimes v) = u \otimes' v$ for all $u \in U, v \in V$. This can be seen as follows. Since $\otimes'$ is bilinear, there exists by the universal property of $\otimes$, a unique linear map $f : T \to T'$ such that $\otimes' = f \circ \otimes$. It suffices to show that f is a bijection. By the universal property of $\otimes'$ there is a linear map $f' : T' \to T$ such that $\otimes' = f' \circ \otimes$. Now $\otimes \circ f' \circ f = \otimes$, which implies that $f' \circ f : T \to T$ is the identity since the image of $\otimes$ spans T (or alternatively, by using the universal property of $\otimes$, and the bilinear map $\otimes$ itself). Hence f is injective. Similarly, $f \circ f'$ is the identity on T' and hence f is surjective.

Definition 3.0.5. Let U, V be vector spaces. The *tensor product* of U and V is a vector space T together with a bilinear map $\otimes : U \times V \to T$ having the universal property. The space T, which is uniquely determined by U and V up to an isomorphism, is denoted by $U \otimes V$.

Often we will refer to $U \otimes V$ as the tensor product of U and V, implicitly assuming the map $\otimes : U \times V \to U \otimes V$.

So far, we have not shown that the tensor product $U \otimes V$ exists at all, nor did we gain insight into the dimension of this space in terms of the dimensions of U and V. One possible construction of $U \otimes V$ is as follows.

Start with the vector space F (for *free* or *formal*) formally spanned by pairs (u, v) as u, v run through U, V, respectively. Now take the subspace R (for *relations*) of F spanned by all elements of the form

$$(c_1 u + u', c_2 v + v') - c_1 c_2 (u, v) - c_1 (u, v') - c_2 (u', v) - (u', v') \tag{3.1}$$

with $c_1, c_2 \in \mathbb{C},\ v, v' \in V, u, u' \in U$. Now any map $\phi : U \times V \to W$ factors through the injection $i : U \times V \to F$ and a unique linear map $g : F \to W$. The kernel of g contains R if and only if ϕ is bilinear, and in that case the map g factors through the quotient map $\pi : F \to F/R$ and a unique linear map $f : F/R \to W$. Taking for $\otimes$ the bilinear map $\pi \circ i : (u, v) \mapsto u \otimes v$, the space F/R together with the map $\otimes$ is the tensor product of U and V.

As for the dimension of $U \otimes V$, let $(u_i)_{i \in I}$ be a basis of U. Then by using bilinearity of the tensor product, every element $T \in U \otimes V$ can be written as a $t = \sum_{i \in I} u_i \otimes w_i$ with w_i non-zero for only finitely many i. We claim that the w_i in such an expression are unique. Indeed, for $k \in I$ let ξ_k be the linear function on U determined by $u_i \mapsto \delta_{ik},\ i \in I$. The bilinear map $U \times V \to V,\ (u, v) \to \xi_k(u) v$ factors, by the universal property, through a unique linear map $f : U \otimes V \to V$. This map sends all terms in the expression $\sum_{i \in I} u_i \otimes w_i$ for T to zero except the term with $i = k$, which is mapped to w_k. Hence $w_k = f_k(t)$ and this shows the uniqueness of the w_k.

Exercise 3.0.6. Use a similar argument to show that if $(v_j)_{j \in J}$ is a basis for V, then the set of all elements of the form $u_i \otimes v_j,\ i \in I,\ j \in J$ form a basis of $U \otimes V$.

This exercise may remind you of matrices. Indeed, there is a natural map ϕ from $U \otimes V^*$, where V^* is the dual of V, into the space $\mathrm{Hom}(V, U)$ of linear maps $V \to U$, defined as follows. Given a pair $u \in U$ and $f \in V^*$, $\phi(u \otimes f)$ is the linear map sending v to $f(v)u$. Here we are implicitly using the universal property: the linear map $v \mapsto f(v)u$ depends bilinearly on f and u, hence there is a unique linear map $U \otimes V^* \to \mathrm{Hom}(V, U)$ that sends $u \otimes f$ to $v \mapsto f(v)u$.

Note that if f and u are both non-zero, then the image of $u \otimes f$ is a linear map of rank one.

Exercise 3.0.7.

1. Show that ϕ is injective. Hint: after choosing a basis $(u_i)_i$ use that a general element of $U \otimes V^*$ can be written in a unique way as $\sum_i u_i \otimes f_i$.
2. Show that ϕ is surjective onto the subspace of $\mathrm{Hom}(V, U)$ of linear maps of finite rank, that is, having finite-dimensional image.

Making things more concrete, if $U = \mathbb{C}^m$ and $V = \mathbb{C}^n$ and $u_1, \ldots, u_m$ is the standard basis of U and $v_1, \ldots, v_n$ is the standard basis of V with dual basis $x_1, \ldots, x_n$, then the tensor $u_i \otimes x_j$ corresponds to the linear map with matrix E_{ij}, the matrix having zeroes everywhere except for a 1 in position (i, j).

Remark 3.0.8. A common mistake is to assume that all elements of $U \otimes V$ are of the form $u \otimes v$. The above shows that the latter elements correspond to rank-one linear maps from V^* to U, or to rank-one matrices, while $U \otimes V$ consists of all finite-rank linear maps from V^* to U—a much larger set.

Next we discuss tensor products of linear maps. If $A : U \to U'$ and $B : V \to V'$ are linear maps, then the map $U \times V \to U' \otimes V'$, $(u, v) \mapsto Au \otimes Bv$ is bilinear. Hence, by the universal property of $U \otimes V$ there exists a unique linear map $U \otimes V \to U' \otimes V'$ that sends $u \otimes v$ to $Au \otimes Bv$. This map is denoted $A \otimes B$.

Example 3.0.9. If $\dim U = m, \dim U' = m', \dim V = n, \dim V' = n'$ and if A, B are represented by an $m' \times m$-matrix $(a_{ij})_{ij}$ and an $n' \times n$-matrix $(b_{kl})_{kl}$, respectively, then $A \otimes B$ can be represented by an $m'n' \times mn$-matrix, with rows labelled by pairs (i, k) with $i \in [m'], k \in [n']$ and columns labelled by pairs (j, l) with $j \in [m], l \in [n]$, whose entry at position $((i, k), (j, l))$ is $a_{ij} b_{kl}$. This matrix is called the *Kronecker product* of the matrices $(a_{ij})_{ij}$ and $(b_{kl})_{kl}$.

Exercise 3.0.10. Assume, in the setting above, that $U = U', m' = m$ and $V = V', n' = n$ and A, B are diagonalisable with eigenvalues $\lambda_1, \ldots, \lambda_m$ and $\mu_1, \ldots, \mu_n$, respectively. Determine the eigenvalues of $A \otimes B$.

Most of what we said about the tensor product of two vector spaces carries over *verbatim* to the tensor product $V_1 \otimes \cdots \otimes V_k$ of k. This tensor product can again be defined by a universal property involving k-linear maps, and its dimension is $\prod_i \dim V_i$. Its elements are called k-tensors. We skip the boring details, but do point out that for larger k there is no longer a close relationship with of k-tensors with linear maps—in particular, the *rank* of a k-tensor T, usually defined as the minimal number of terms in any expression of T as a sum of *pure tensors* $v_1 \otimes \cdots \otimes v_k$, is only poorly understood. For instance, *computing* the rank, which for $k = 2$ can be done using Gaussian elimination, is very hard in general. If all V_i are the same, say V, then we also write $V^{\otimes k}$ for $V \otimes \cdots \otimes V$ (k factors).

Given three vector spaces U, V, W, we now have several ways to take their tensor product: $(U \otimes V) \otimes W$, $U \otimes (V \otimes W)$ and $U \otimes V \otimes W$. Fortunately, these tensor products can be identified. For example, there is a unique linear isomorphism $f : U \otimes V \otimes W \to (U \otimes V) \otimes W$ such that $f(u \otimes v \otimes w) = (u \otimes v) \otimes w$ for all $u \in U, v \in V, w \in W$.

Indeed, consider the trilinear map $U \times V \times W \to (U \otimes V) \otimes W$ defined by $(u, v, w) \mapsto (u \otimes v) \otimes w$. By the universal property, there is a unique linear map $f : U \otimes V \otimes W \to (U \otimes V) \otimes W$ such that $f(u \otimes v \otimes w) = (u \otimes v) \otimes w$ for all u, v, w.

Now for fixed $w \in W$, the bilinear map $\phi_w : U \times V \to U \otimes V \otimes W$ defined by $\phi_w(u, v) := u \otimes v \otimes w$ induces a linear map $g_w : U \otimes V \to U \otimes V \otimes W$ such that $u \otimes v$ is mapped to $u \otimes v \otimes w$. Hence the bilinear map $\phi : (U \otimes V) \times W \to U \otimes V \otimes W$ given by $\phi(x, w) := g_w(x)$ induces a linear map $g : (U \otimes V) \otimes W \to U \otimes V \otimes W$ sending $(u \otimes v) \otimes w$ to $u \otimes v \otimes w$. It follows that $f \circ g$ and $g \circ f$ are the identity on $(U \otimes V) \otimes W$ and $U \otimes V \otimes W$ respectively. This shows that f is an isomorphism.

Exercise 3.0.11. Let V be a vector space. Show that for all p, q there is a unique linear isomorphism $V^{\otimes p} \otimes V^{\otimes q} \to V^{\otimes (p+q)}$ sending $(v_1 \otimes \cdots \otimes v_p) \otimes (v_{p+1} \otimes \cdots \otimes v_{p+q})$ to $v_1 \otimes \cdots \otimes v_{p+q}$.

The direct sum $TV := \bigoplus_{k=0}^{\infty} V^{\otimes k}$ is called the *tensor algebra* of V, where the natural linear map $V^{\otimes k} \times V^{\otimes l} \to V^{\otimes k} \otimes V^{\otimes l} = V^{\otimes (k+l)}$ plays the role of (non-commutative but associative) multiplication. We move on to other types of tensors.

Definition 3.0.12. Let V be a vector space. A k-linear map $\omega : V^k \to W$ is called *symmetric* if $\omega(v_1, \ldots, v_k) = \omega(v_{\pi(1)}, \ldots, v_{\pi(k)})$ for all permutations $\pi \in \mathrm{Sym}(k)$.

The k-th *symmetric power* of V is a vector space $S^k V$ together with a symmetric k-linear map $V^k \to S^k V,\ (v_1, \ldots, v_k) \to v_1 \cdots v_k$ such that for all vector spaces W and symmetric k-linear maps $\psi : V^k \to W$ there is a unique linear map $\phi : S^k V \to W$ such that $\psi(u_1, \ldots, u_k) = \phi(u_1 \cdots u_k)$.

Uniqueness of the k-th symmetric power of V can be proved in exactly the same manner as uniqueness of the tensor product. For existence, let R be the subspace of $V^{\otimes k} := V \otimes \cdots \otimes V$ spanned by all elements of the form

$$v_1 \otimes \cdots \otimes v_k - v_{\pi(1)} \otimes \cdots \otimes v_{\pi(k)},\ \pi \in \mathrm{Sym}(k).$$

Then the composition of the maps $V^k \to V^{\otimes k} \to V^{\otimes k}/R$ is a symmetric k-linear map and if $\psi : V^k \to W$ is any such map, then ψ factors through a linear map $V^{\otimes k} \to W$ since it is k-linear, which in turn factors through a unique linear map $V^{\otimes k}/R \to W$ since ψ is symmetric. This shows existence of symmetric powers, and, perhaps more importantly, the fact that they are quotients of tensor powers of V. This observation will be very important in proving the first fundamental theorem for $\mathrm{GL}(V)$.

There is also an analogue of the tensor product of maps: if A is a linear map $U \to V$, then the map $U^k \to S^k V,\ (u_1, \ldots, u_k) \mapsto Au_1 \cdots Au_k$ is multilinear and symmetric. Hence, by the universal property of symmetric powers, it factors through the map $U^k \to S^k U$ and a linear map $S^k U \to S^k V$. This map, which sends $u_1 \cdots u_k$ to $Au_1 \cdots Au_k$, is the k-th symmetric power $S^k A$ of A.

If $(v_i)_{i \in I}$ is a basis of V, then using multilinearity and symmetry every element t of $S^k V$ can be written as a linear combination $\sum_\alpha c_\alpha v^\alpha$ of the elements $v^\alpha := \prod_{i \in I} v_i^{\alpha_i}$—the product order is immaterial—where $\alpha \in \mathbb{N}^I$ satisfies $|\alpha| := \sum_{i \in I} \alpha_i = k$ and only finitely many coefficients c_α are non-zero. We claim that the c_α are unique, so that the v^α, $|\alpha| = k$ a basis of V. Indeed, let $\alpha \in \mathbb{N}^I$ with $|\alpha| = k$. Then there is a unique k-linear map $\psi_\alpha : V^k \to \mathbb{C}$ which on a tuple $(v_{i_1}, \ldots, v_{i_k})$ takes the value 1 if $|\{j \mid i_j = i\}| = \alpha_i$ for all $i \in I$ and zero otherwise. Moreover, ψ_α is symmetric and therefore induces a linear map $\phi_\alpha : S^k V \to \mathbb{C}$. We find that $c_\alpha = \phi_\alpha(t)$, which proves the claim.

This may remind you of polynomials. Indeed, if $V = \mathbb{C}^n$ and $x_1, \ldots, x_n$ is the basis of V^* dual to the standard basis of V, then $S^k V^*$ is just the space of *homogeneous polynomials* in the x_i of degree k. The algebra of all polynomial functions on V therefore is canonically isomorphic to $SV^* := \bigoplus_{k=0}^\infty S^k V^*$. The product of a homogeneous polynomials of degree k and homogeneous polynomials of degree l corresponds to the unique bilinear map $S^k V^* \times S^l V^* \to S^{k+l} V^*$ making the diagram

$$\begin{tikzcd}
(V^*)^{\otimes k} \times (V^*)^{\otimes l} \arrow[r] \arrow[d] & (V^*)^{\otimes k+l} \arrow[d] \\
S^k V^* \times S^l V^* \arrow[r, dashed] & S^{k+l} V^*
\end{tikzcd}$$

commute, and this corresponds to multiplying polynomials of degrees k and l. Thus SV^* is a quotient of the tensor algebra TV (in fact, the maximal commutative quotient).

Above we have introduced $S^k V$ as a *quotient* of $V^{\otimes k}$. This should not be confused with the *subspace* of $V^{\otimes k}$ spanned by all *symmetric tensors*, defined as follows. For every permutation $\pi \in S_k$ there is a natural map $V^k \to V^k$ sending $(v_1, \ldots, v_k)$ to $(v_{\pi^{-1}(1)}, \ldots, v_{\pi^{-1}(k)})$. Composing this map with the natural k-linear map $V^k \to V^{\otimes k}$ yields another k-linear map $V^k \to V^{\otimes k}$, and hence a linear map $V^{\otimes k} \to V^{\otimes k}$, also denoted π. A tensor ω in $V^{\otimes k}$ is called symmetric if $\pi\omega = \omega$ for all $\pi \in S_k$. The restriction of the map $V^{\otimes k} \to S^k V$ to the subspace of symmetric tensors is an isomorphism with inverse determined by $v_1 \cdots v_k \mapsto \frac{1}{k!} \sum_{\pi \in S_k} \pi(v_1 \otimes \cdots v_k)$. (Note that this inverse would not be defined in characteristic less than k.)

Exercise 3.0.13. Show that the subspace of symmetric tensors in $V^{\otimes k}$ is spanned by the tensors $v \otimes v \cdots \otimes v$, where $v \in V$.

3.1 Exercises

Exercise 3.1.1. Let $U \otimes V$ be the tensor product of the vector spaces U and V. Let $u_1, \ldots, u_s$ and $u'_1, \ldots, u'_t$ be two systems of linearly independent vectors in U and let $v_1, \ldots, v_s$ and $v'_1, \ldots, v'_t$ be two systems of linearly independent vectors in V. Suppose that

$$u_1 \otimes v_1 + \cdots + u_s \otimes v_s = u'_1 \otimes v'_1 + \cdots + u'_t \otimes v'_t. \tag{3.2}$$

Show that $s = t$.

Exercise 3.1.2. a) Let $T \in V_1 \otimes V_2 \otimes V_3$ be an element of the tensor product of V_1, V_2 and V_3. Suppose that there exist $v_1 \in V_1$, $v_3 \in V_3$, $T_{23} \in V_2 \otimes V_3$ and $T_{12} \in V_1 \otimes V_2$ such that

$$T = v_1 \otimes T_{23} = T_{12} \otimes v_3. \tag{3.3}$$

Show that there exist a $v_2 \in V_2$ such that $T = v_1 \otimes v_2 \otimes v_3$.

b) Suppose that $T \in V_1 \otimes V_2 \otimes V_3$ can be written as a sum of at most d_1 tensors of the form $v_1 \otimes T_{23}$, where $v_1 \in V_1, T_{23} \in V_2 \otimes V_3$, and also as a sum of at most d_3 tensors of the form $T_{12} \otimes v_3$, where $v_3 \in V_3, T_{12} \in V_1 \otimes V_2$. Show that T can be written as the sum of at most $d_1 d_3$ tensors of the form $v_1 \otimes v_2 \otimes v_3$, where $v_i \in V_i$.

Exercise 3.1.3. Let U, V, W be vector spaces. Denote by $B(U \times V, W)$ the linear space of bilinear maps from $U \times V$ to W. Show that the map $f \mapsto f \circ \otimes$ is a linear isomorphism between $\mathrm{Hom}(U \otimes V, W)$ and $B(U \times V, W)$.

Exercise 3.1.4. Let U, V be vector spaces. Show that the linear map $\phi : U^* \otimes V^* \to (U \otimes V)^*$ given by $\phi(f \otimes g)(u \otimes v) := f(u)g(v)$ is an isomorhism.

Chapter 4

Lecture 4. Representations

Central objects in this course are linear representations of groups. We will only consider representations on complex vector spaces. Recall the folowing definition.

Definition 4.0.5. Let G be a group and let V be a vector space. A *(linear) representation of G on V* is a group homomorphism $\rho : G \to \mathrm{GL}(V)$.

If ρ is a representation of G, then the map $(g, v) \mapsto \rho(g)v$ is an action of G on V. A vector space with an action of G by linear maps is also called a *G-module*. Instead of $\rho(g)v$ we will often write gv.

Definition 4.0.6. Let U and V be G-modules. A linear map $\phi : U \to V$ is called a *G-module morphism* or a *G-linear map* if $\phi(gu) = g\phi(u)$ for all $u \in U$ and $g \in G$. If ϕ is invertible, then it is called an isomorphism of G-modules. The linear space of all G-linear maps from U to V is denoted $\mathrm{Hom}(U, V)^G$.

The multilinear algebra constructions from Section 3 carry over to representations. For instance, if $\rho : G \to \mathrm{GL}(U)$ and $\sigma : G \to \mathrm{GL}(V)$ are representations, then the map $\rho \otimes \sigma : G \to \mathrm{GL}(U \otimes V),\ g \mapsto \rho(g) \otimes \sigma(g)$ is also a representation. Similarly, for any natural number k the map $g \mapsto S^k\rho(g)$ is a representation of G on S^kV. Also, the dual space V^* of all linear functions on V carries a natural G-module structure: for $f \in V^*$ and $g \in G$ we let gf be the linear function defined by $gf(v) = f(g^{-1}v)$. The inverse ensures that the action is a left action rather than a right action: for $g, h \in G$ and $v \in V$ we have

$$(g(hf))(v) = (hf)(g^{-1}v) = f(h^{-1}g^{-1}v) = f((gh)^{-1}v) = ((gh)f)(v),$$

so that $g(hf) = (hg)f$.

Exercise 4.0.7. Show that the set of fixed points in $\mathrm{Hom}(U, V)$ under the action of G is precisely $\mathrm{Hom}(U, V)^G$.

Example 4.0.8. Let V, U be G-modules. Then the space $\mathrm{Hom}(V, U)$ of linear maps $V \to U$ is a G module with the action $(g\phi)(v) := g\phi(g^{-1}v)$. The space

$U \otimes V^*$ is also a G-module with action determined by $g(u \otimes f) = (gu) \otimes (gf)$. The natural map $\Psi : U \otimes V^* \to \mathrm{Hom}(V, U)$ determined by $\Psi(u \otimes f)v = f(v)u$ is a morphism of G-modules. To check this it suffices to observe that

$$\Psi(g(u \otimes f))v = \Psi((gu) \otimes (gf))v = (gf)(v) \cdot gu = f(g^{-1}v) \cdot gu$$

and

$$(g\Psi(u \otimes f))v = g\Psi(u \otimes f)(g^{-1}v) = g(f(g^{-1}v)u) = f(g^{-1}v) \cdot gu.$$

The map Ψ is an G-module isomorphism of $U \otimes V^*$ with the space of finite-rank linear maps from V to U. In particular, if U or V is finite-dimensional, then Ψ is an isomorphism.

Example 4.0.9. Let G be a group acting on a set X. Consider the vectorspace

$$\mathbb{C}X := \{\sum_{x \in X} \lambda_x x \mid \lambda_x \in \mathbb{C} \text{ for all } x \in X \text{ and } \lambda_x = 0 \text{ for almost all } x\} \qquad (4.1)$$

consisting of all formal finite linear combinations of elements from X. The natural action of G given by $g(\sum_x \lambda_x x) := \sum_x \lambda_x gx$ makes $\mathbb{C}X$ into a G module. In the special case where $X = G$ and G acts on itself by multiplication on the left, the module $\mathbb{C}G$ is called the *regular representation of* G.

Definition 4.0.10. A G-*submodule* of a G-module V is a G-stable subspace, that is, a subspace U such that $gU \subseteq U$ for all $g \in G$. The *quotient* V/U then carries a natural structure of G-module, as well, given by $g(v+U) := (gv)+U$.

Definition 4.0.11. A G-module V is called *irreducible* if it has precisely two G-submodules (namely, 0 and V).

Exercise 4.0.12. Show that for finite groups G, every irreducible G-module has finite dimension.

Note that, just like 1 is not a prime number and the empty graph is not connected, the zero module is not irreducible. In this course we will be concerned only with G-modules that are either finite-dimensional or *locally finite*.

Definition 4.0.13. A G-module V is called *locally finite* if every $v \in V$ is contained in a finite-dimensional G-submodule of V.

Proposition 4.0.14. *For a locally finite G-module V the following statements are equivalent.*

1. *for every G-submodule U of V there is a G-submodule W of V such that $U \oplus W = V$;*

2. *V is a (potentially infinite) direct sum of finite-dimensional irreducible G-submodules.*

In this case we call V *completely reducible*; note that we include that condition that V be locally finite in this notion.

Proof. First assume (1). Let $\mathcal{M}$ be the collection of all finite-dimensional irreducible G-submodules of V. The collection of subsets S of $\mathcal{M}$ for which the sum $\sum_{U \in S} U$ is direct satisfies the condition of Zorn's Lemma: the union of any chain of such subsets S is again a subset of $\mathcal{M}$ whose sum is direct. Hence by Zorn's Lemma there exists a maximal subset S of $\mathcal{M}$ whose sum is direct. Let U be its (direct) sum, which is a G-submodule of V. By (1) U has a direct complement W, which is also a G-submodule. If W is non-zero, then it contains a non-zero finite-dimensional submodule (since it is locally finite), and for dimension reasons the latter contains an irreducible G-submodule W'. But then $S \cup \{W'\}$ is a subset of $\mathcal{M}$ whose sum is direct, contradicting maximality of S. Hence $W = 0$ and $V = U = \bigoplus_{M \in S} M$, which proves (2).

For the converse, assume (2) and write V as the direct sum $\bigoplus_{M \in S} M$ of irreducible finite-dimensional G-modules. Let U be any submodule of V. Then the collections of subsets of S whose sum intersects U only in 0 satisfies the condition of Zorn's Lemma. Hence there is a maximal such subset S'. Let W be its sum. We claim that $U + W = V$ (and the sum is direct by construction). Indeed, if not, then some element M of S is not contained in $U + W$. But then $M \cap (U + V) = \{0\}$ by irreducibility of M and therefore the sum of $S' \cup \{M\}$ still intersects U trivially, contradicting the maximality of S'. This proves (1). $\square$

Remark 4.0.15. It is not hard to prove that direct sums, submodules, and quotients of locally finite G-modules are again locally finite, and that they are also completely reducible if the original modules were.

Example 4.0.16. A typical example of a module which is not completely reducible is the following. Let G be the group of invertible upper triangular 2×2-matrices, and let $V = \mathbb{C}^2$. Then the subspace spanned by the first standard basis vector e_1 is a G-submodule, but it does not have a direct complement that is G-stable.

Note that the group in this example is infinite. This is not a coincidence, as the following fundamental results show.

Proposition 4.0.17. *Let G be a finite group and let V be a finite-dimensional G-module. Then there exists a Hermitian inner product $(.|.)$ on V such that $(gu|gv) = (u|v)$ for all $g \in G$ and $u, v \in V$.*

Proof. Let $(.|.)'$ be any Hermitian inner product on V and take

$$(u|v) := \sum_{g \in G} (gu|gv)'.$$

Straightforward computations shows that $(.|.)$ is G-invariant, linear in its first argument, and semilinear in its second argument. For positive definiteness, note that for $v \neq 0$ the inner product $(v|v) = \sum_{g \in G} (gv|gv)$ is positive since every entry is positive. $\square$

Theorem 4.0.18. *For a finite group G any G-module is completely reducible.*

Proof. Let V be a G-module. Then every $v \in V$ lies in the finite-dimensional subspace spanned by its *orbit* $Gv = \{gv \mid g \in G\}$, which moreover is G-stable. Hence V is locally finite. By Zorn's lemma there exists a submodule U of V which is maximal among all direct sums of finite-dimensional irreducible submodules of V. If U is not all of V, then let W be a finite-dimensional submodule of V not contained in U, and let $(.|.)$ be a G-invariant Hermitian form on W. Then $U \cap W$ is a G-submodule of W, and therefore so is the orthogonal complement $(U \cap W)^\perp$ of $U \cap W$ in W—indeed, one has $(gw|U \cap W) = (w|g^{-1}(U \cap W)) \subseteq (w|U \cap W) = \{0\}$ for $g \in G$ and $w \in (U \cap W)^\perp$, so that $gw \in (U \cap W)^\perp$. Let W' be an irreducible submodule of $(U \cap W)^\perp$. Then $U \oplus W'$ is a larger submodule of V which is the direct sum of irreducible submodules of V, a contradiction. Hence $V = U$ is completely reducible. □

4.1 Schur's lemma and isotypic decomposition

The following easy observation due to the German mathematician Issai Schur (1875-1941) is fundamental to representation and invariant theory.

Lemma 4.1.1 (Schur's Lemma). *Let V and U be irreducible finite-dimensional G modules for some group G. Then either V and U are isomorphic and* $\mathrm{Hom}(V, U)^G$ *is one-dimensional, or they are not isomorphic and* $\mathrm{Hom}(V, U)^G = \{0\}$.

Proof. Suppose that $\mathrm{Hom}(V, U)^G$ contains a non-zero element ϕ. Then $\ker \phi$ is a G-submodule of V unequal to all of V and hence equal to $\{0\}$. Also, $\mathrm{im}\, \phi$ is a G-submodule of U unequal to $\{0\}$, hence equal to U. It follows that ϕ is an isomorphism of G-modules. Now suppose that ϕ' is a second element of $\mathrm{Hom}(V, U)^G$. Then $\psi := \phi' \circ \phi^{-1}$ is a G-morphism from U to itself; let $\lambda \in \mathbb{C}$ be an eigenvalue of it. Then $\psi - \lambda I$ is a G-morphism from U to itself, as well, and its kernel is a non-zero submodule, hence all of U. This shows that $\psi = \lambda I$ and therefore $\phi' = \lambda\phi$. Hence $\mathrm{Hom}(V, U)^G$ is one-dimensional, as claimed. □

If G is a group and V is a completely reducible G-module, then the decomposition of V as a direct sum of irreducible G-modules need not be unique. For instance, if V is the direct sum $U_1 \oplus U_2 \oplus U_3$ where the first two are isomorphic irreducible modules and the third is an irreducible module not isomorphic to the other two, then V can also be written as $U_1 \oplus \Delta \oplus U_3$, where $\Delta = \{u_1 + \phi(u_1) \mid u_1 \in U_1\}$ is the diagonal subspace of $U_1 \oplus U_2$ corresponding to an isomorphism ϕ from U_1 to U_2.

However, there is always a coarser decomposition of V into G-modules which *is* unique. For this, let $\{U_i\}_{i \in I}$ be a set of representatives of the isomorphism classes of G-modules, so that every irreducible finite-dimensional G-module is isomorphic to U_i for some unique $i \in I$. For every $i \in I$ let V_i be the (non-direct) sum of all G-submodules of V that are isomorphic to U_i. Clearly each V_i is a G-submodule of V and, since V is a direct sum of irreducible G-modules, $\sum_{i \in I} V_i = V$. Using Zorn's lemma one sees that V_i can also be written as $\bigoplus_{j \in J_i} V_{ij}$ for irreducible submodules $V_{ij},\ j \in J_i$ that are all isomorphic to U_i.

We claim that $V = \bigoplus_{i \in I} V_i$. To see this, suppose that $V_{i_0} \cap \sum_{i \neq i_0} V_i \neq \{0\}$ and let U be an irreducible submodule of this module. Then the projection of U onto some $V_{i_0,j}$ along the direct sum of the remaining direct summands ov V_{i_0} is non-zero, and similarly the projection of U onto $V_{i_1,j}$ for some i_1, j along the remaining summands of V_{i_1} is non-zero. By Schur's lemma U is then both isomorphic to $V_{i_0,j}$ and to $V_{i_1,j}$, a contradiction. Hence $V_{i_0} \cap \sum_{i \neq i_0} V_i$ is zero, as claimed.

The space V_i is called the *isotypic component of V of type U_i*, and it has the following pretty description. The map $\mathrm{Hom}(U_i, V)^G \times U_i \to V,\ (\phi, u) \mapsto \phi(u)$ is bilinear, and therefore gives rise to a linear map $\Psi : \mathrm{Hom}(U_i, V)^G \otimes U_i \to V$. This linear map is a linear isomorphism onto V_i.

Exercise 4.1.2. Let U, V, W be G-modules. Show that $\mathrm{Hom}(U \oplus V, W)^G \cong \mathrm{Hom}(U, W) \oplus \mathrm{Hom}(V, W)$ and $\mathrm{Hom}(W, U \oplus V) \cong \mathrm{Hom}(W, U) \oplus \mathrm{Hom}(W, V)$.

4.2 Exercises

Exercise 4.2.1.

- Let V be a G-module and $\langle , \rangle$ a G-invariant inner product on V. Show that for any two non-isomorphic, irreducible submodules $V_1, V_2 \subset V$ we have $V_1 \perp V_2$, that is, $\langle v_1, v_2 \rangle = 0$ for all $v_1 \in V_1$, $v_2 \in V_2$.
- Give an example where $V_1 \not\perp V_2$ for (isomorphic) irreducible G-modules V_1 and V_2.

Exercise 4.2.2. Let the symmetric group on 3 letters S_3 act on $\mathbb{C}[x_1, x_2, x_3]_2$ by permuting the variables. This action makes $\mathbb{C}[x_1, x_2, x_3]_2$ into a S_3-module. Give a decomposition of this module into irreducible submodules.

Exercise 4.2.3. Let G be an abelian group. Show that every irreducible G-module has dimension 1. Show that G has a *faithful* irreducible representation if and only if G is cyclic. A representation ρ is called faithful if it is injective.

Exercise 4.2.4. Let G be a finite group and V an irreducible G-module. Show that there is a *unique* G-invariant inner product on V, unique up to multiplication by scalars.

Exercise 4.2.5. Let G be a finite group, and let $\mathbb{C}G$ be the regular representation of G and let $\mathbb{C}G = W_1^{m_1} \oplus \cdots \oplus W_k^{m_k}$ be the isotypic decomposition of $\mathbb{C}G$. Show that for every irreducible G-module W, there is an i such that W is isomorphic to W_i and show that $m_i = \dim W_i$. Hint: for all $w \in W$ the linear map $\mathbb{C}G \to W$ given by $\sum_g \lambda_g g \mapsto \sum_g \lambda_g g w$ is a G-linear map.

Chapter 5

Finite generation of the invariant ring

In all examples we have met so far, the invariant ring was generated by a finite number of invariants. In this section we prove Hilbert's theorem that under reasonable conditions, this is always the case. For the proof we will need another theorem by Hilbert.

Recall that for a ring R and a subset $S \subseteq R$, the *ideal generated by S* is defined as

$$(S) := \{r_1 s_1 + \cdots + r_k s_k \mid k \in \mathbb{N}, r_1, \ldots, r_k \in R, s_1, \ldots, s_k \in S\}. \tag{5.1}$$

You may want to check that this indeed defines an ideal in R. An ideal $I \subseteq R$ is called *finitely generated* if there is a finite set S such that $I = (S)$.

Definition 5.0.6. A ring R is called *Noetherian* if every ideal I in R is finitely generated.

Exercise 5.0.7. Show that a ring R is Noetherian if and only if there is no infinite ascending chain of ideals $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots$.

We will be mostly interested in polynomial rings over $\mathbb{C}$ in finitely many indeterminates, for which the following theorem is essential.

Theorem 5.0.8 (Hilbert's Basis Theorem)**.** *The polynomial ring $\mathbb{C}[x_1, \ldots, x_n]$ is Noetherian.*

We will deduce this statement from the following result.

Lemma 5.0.9 (Dixon's Lemma)**.** *If $m_1, m_2, m_3, \ldots$ is an infinite sequence of monomials in the variables $x_1, \ldots, x_n$, then there exist indices $i < j$ such that $m_i | m_j$.*

Proof. We proceed by induction on n. For $n = 0$ all monomials are 1, so we can take any $i < j$. Suppose that the statement is true for $n - 1 \geq 0$. Define the infinite sequences $e_1 \leq e_2 \leq \ldots$ and $i_1 < i_2 < \ldots$ as follows: e_1 is the smallest exponent of x_n in any of the monomials m_i, and i_1 is the smallest index i for which the exponent of x_n in m_i equals e_1. For $k > 1$ the exponent e_k is the smallest exponent of x_n in any of the m_i with $i > i_{k-1}$ and i_k is the smallest index $i > i_{k-1}$ for which the exponent of x_n in m_i equals e_k. Now the monomials in the sequence $m_{i_1}/x_n^{e_1}, m_{i_2}/x_n^{e_2}, \ldots$ do not contain x_n. Hence by induction there exist $j < l$ such that $m_{i_j}/x_n^{e_j} | m_{i_l}/x_n^{e_l}$. As $e_j \leq e_l$ we then also have $m_{i_j} | m_{i_l}$, and of course $i_j < i_l$, as claimed. □

Proof of Hilbert's Basis Theorem. Let $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be an ideal. For any polynomial f in $\mathbb{C}[x_1, \ldots, x_n]$ we denote by $\mathrm{lm}(f)$ the leading monomial of f: the lexicographically largest monomial having non-zero coefficient in f. By Dixon's lemma, the set of $|$-minimal monomials in $\{\mathrm{lm}(f) \mid f \in I\}$ is finite. Hence there exist finitely many polynomials $f_1, \ldots, f_k \in I$ such that for all $f \in I$ there exists an i with $\mathrm{lm}(f_i)|\mathrm{lm}(f)$. We claim that the ideal $J := (f_1, \ldots, f_k)$ generated by the f_i equals I. If not, then take an $f \in I \setminus J$ with the lexicographically smallest leading monomial among all counter examples. Take i such that $\mathrm{lm}(f_i)|\mathrm{lm}(f)$, say $\mathrm{lm}(f) = m\mathrm{lm}(f_i)$. Subtracting a suitable scalar multiple of mf_i, which lies in J, from f gives a polynomial with a lexicographically smaller leading monomial, and which is still in $I \setminus J$. But this contradicts the minimality of $\mathrm{lm}(f)$. □

Remark 5.0.10. More generally, Hilbert showed that for R Noetherian, also $R[x]$ is Noetherian (which you may want to prove yourself!). Since clearly any field is a Noetherian ring, this implies the previous theorem by induction on the number of indeterminates.

With this tool in hand, we can now return to our main theorem of this section.

Theorem 5.0.11 (Hilbert's Finiteness Theorem)**.** *Let G be a group and let W be a finite dimensional G-module with the property that $\mathbb{C}[W]$ is completely reducible. Then $\mathbb{C}[W]^G := \{f \in \mathbb{C}[W] \mid gf = f\}$ is a finitely generated subalgebra of $\mathbb{C}[W]$. That is, there exist $f_1, \ldots, f_k \in \mathbb{C}[W]^G$ such that every G-invariant polynomial on W, is a polynomial in the f_i.*

The proof uses the so-called *Reynolds operator* ρ, which is defined as follows. We assume that the vector space $\mathbb{C}[W]$ is completely reducible. Consider its isotypic decomposition $\mathbb{C}[W] = \bigoplus_{i \in I} V_i$ and let $1 \in I$ correspond to the trivial 1-dimensional G-module, so that $\mathbb{C}[W]^G = V_1$. Now let ρ be the projection from $\mathbb{C}[W]$ onto V_1 along the direct sum of all V_i with $i \neq 1$. This is a G-equivariant linear map. Moreover, we claim that

$$\rho(f \cdot h) = f \cdot \rho(h) \text{for all } f \in V_1, \tag{5.2}$$

where the multiplication is multiplication in $\mathbb{C}[W]$. Indeed, consider the map $\mathbb{C}[W] \to \mathbb{C}[W],\ h \mapsto fh$. This a G-module morphism, since $g(f \cdot h) = (gf) \cdot (gh) = f \cdot (gh)$, where the first equality reflects that G acts by automorphisms on $\mathbb{C}[W]$ and the second equality follows from the invariance of f. Hence if we write h as $\sum_i h_i$ with $h_i \in V_i$, then $fh_i \in V_i$ by Schur's lemma, and therefore the component of $fh = \sum_i (fh_i)$ in V_1 is just fh_1. In other words $\rho(fh) = f\rho(h)$, as claimed.

Exercise 5.0.12. Show that for a finite group G, the Reynolds operator is just $f \mapsto \frac{1}{|G|}\sum_{g\in G} gf$.

Proof of Hilbert's finiteness theorem. Let $I' := \bigoplus_{d>0} \mathbb{C}[W]_d^G$ be the ideal in $\mathbb{C}[W]^G$ consisting of all invariants with zero constant term. Denote by $I := \mathbb{C}[W]I'$ the ideal in $\mathbb{C}[W]$ generated by I'. Since W is finite dimensional, it follows from Hilbert's basis theorem that there exist $f_1, \ldots, f_k \in I$ that generate the ideal I. We may assume that the f_i belong to I'. Indeed, if $f_i \notin I'$, we can write $f_i = \sum_j f_{ij}g_{ij}$ for certain $f_{ij} \in I'$ and $g_{ij} \in \mathbb{C}[W]$ and replace f_i with the f_{ij} to obtain a finite generating set of I with fewer elements in $I \setminus I'$.

We observe that the ideal I' is generated by the f_i. Indeed, let $h \in I' \subseteq I$ and write $h = \sum_i g_i f_i$ for some $g_i \in \mathbb{C}[W]$. Using the Reynolds operator ρ we find: $h = \rho(h) = \sum_i \rho(f_i g_i) = \sum_i f_i \rho(g_i)$.

The proof is now completed by exercise 5.0.13. □

Exercise 5.0.13. Let $A \subseteq \mathbb{C}[W]$ be a subalgebra, and let $A^+ := \oplus_{d\geq 1} A \cap \mathbb{C}[W]_d$ be the ideal of polynomials in A with zero coefficient. Suppose that the ideal A^+ is finitely generated. Show that A is finitely generated as an algebra over $\mathbb{C}$.

5.1 Noethers degree bound

For a *finite* group G, any G-module V is completely reducible as we have seen in the previous lecture. This implies by Hilbert's theorem that for finite groups, the invariant ring is always finitely generated. In this section, we prove a result of Noether stating that for finite groups G, the invariant ring is already generated by the invariants of degree at most $|G|$, which implies a bound on the number of generators needed.

Theorem 5.1.1 (Noether's degree bound)**.** *Let G be a finite group, and let W be a (finite dimensional) G-module. Then the invariant ring $\mathbb{C}[W]^G$ is generated by the homogeneous invariants of degree at most $|G|$.*

Proof. We choose a basis $x_1, \ldots, x_n$ of W^* so that $\mathbb{C}[W] = \mathbb{C}[x_1, \ldots, x_n]$. For any n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers we have an invariant

$$j_\alpha := \sum_{g\in G} g(x_1^{\alpha_1}\cdots x_n^{\alpha_n}) \tag{5.3}$$

homogeneous of degree $|\alpha| := \alpha_1 + \cdots + \alpha_n$. Clearly, the invariants j_α span the vector space $\mathbb{C}[W]^G$, since for any invariant $f = \sum_\alpha c_\alpha x_1^{\alpha_1}\cdots x_n^{\alpha_n}$, we have

$$f = \frac{1}{|G|}\sum_{g\in G} gf = \frac{1}{|G|}\sum_\alpha c_\alpha j_\alpha. \tag{5.4}$$

It will therefore suffice to prove that every j_α is a polynomial in the j_β with $|\beta| <= |G|$.

Let $z_1, \ldots, z_n$ be n new variables and define for $j \in \mathbb{N}$ the polynomials

$$p_j(x_1, \ldots, x_n, z_1, \ldots, z_n) := \sum_{g \in G} (gx_1 \cdot z_1 + \cdots + gx_n \cdot z_n)^j. \tag{5.5}$$

So these are the Newton polynomials (see Lecture 2), where we have substituted the expressions $(gx_1 \cdot z_1 + \cdots + gx_n \cdot z_n)$ for the $|G|$ variables. Expanding p_j and sorting terms with respect to the variables z_i, we see that $p_j = \sum_{|\alpha|=j} f_\alpha z_1^{\alpha_1} \cdots x_n^{\alpha_n}$, where

$$f_\alpha = \binom{j}{\alpha_1, \ldots, \alpha_n} j_\alpha. \tag{5.6}$$

Now let $j > |G|$. Recall that p_j is a polynomial in $p_1, \ldots, p_{|G|}$. This implies that also the coefficients $f_\alpha, |\alpha| = j$ of p_j are polynomials in the coefficients $f_\beta, |\beta| \leq |G|$ of $p_1, \ldots, p_{|G|}$. This finishes the proof, since $\binom{j}{\alpha_1,\ldots,\alpha_n} \neq 0$ when $\alpha_1 + \cdots + \alpha_n = j$. □

Exercise 5.1.2. Show that for all cyclic groups, the bound in the theorem is met in some representation.

5.2 Exercises

For a finite group G, define $\beta(G)$ to be the minimal number m such that for every (finite dimensional) G-module W, the invariantring $\mathbb{C}[W]^G$ is generated by the invariants of degree at most m. By Noether's theorem, we always have $\beta(G) \leq |G|$.

Exercise 5.2.1. Let G be a finite abelian group. We use additive notation. Define the Davenport constant $\delta(G)$ to be the maximum length m of a non-shortable expression $0 = g_1 + \cdots + g_m, g_1, \ldots, g_m \in G$. Non-shortable means that there is no strict non-empty subset I of $\{1, \ldots, n\}$ such that $\sum_{i \in I} g_i = 0$. Show that $\delta(G) = \beta(G)$. Compute $\delta((\mathbb{Z}/2\mathbb{Z})^n)$.

Chapter 6

Affine varieties and the quotient map

6.1 Affine varieties

Definition 6.1.1. An *affine variety* is a subset of some $\mathbb{C}^n$ which is the common zero set of a collection of polynomials in the coordinates $x_1, \ldots, x_n$ on $\mathbb{C}^n$.

Suppose that S is a subset of $\mathbb{C}[\underline{x}] := \mathbb{C}[x_1, \ldots, x_n]$ and let $p \in \mathbb{C}^n$ be a common zero of the elements of S. Then any finite combination $\sum_i a_i f_i$ where the f_i are in S and the a_i are in $\mathbb{C}[x]$ also vanishes on p. The collection of all such polynomials is the ideal generated by S. So the study of affine varieties leads naturally to the study of ideals in the polynomial ring $\mathbb{C}[x_1, \ldots, x_n]$. We have seen in Week 5 that such ideals are always finitely generated.

Exercise 6.1.2. Show that the collection of affine varieties in $\mathbb{C}^n$ satisfy the following three properties:

1. $\mathbb{C}^n$ and $\emptyset$ are affine varieties;
2. the union of two affine varieties is an affine variety; and
3. the intersection of arbitrarily many affine varieties is an affine variety.

These conditions say that the affine varieties in $\mathbb{C}^n$ form the closed subsets in a *topology* on $\mathbb{C}^n$. This topology is called the *Zariski topology*, after the Polish-American mathematician Otto Zariski (1899-1986). We will interchangeably use the terms affine (sub)variety in $\mathbb{C}^n$ and Zariski-closed subset of $\mathbb{C}^n$. Moreover, in the last case we will often just say closed subset; when we mean closed subset in the Euclidean sense rather than in the Zariski-sense, we will explicitly mention that.

Exercise 6.1.3. The Zariski-topology on $\mathbb{C}^n$ is very different from the Euclidean topology on $\mathbb{C}^n$, as the answers to the following problems show:

1. determine the Zariski-closed subsets of $\mathbb{C}$;
2. prove that $\mathbb{R}^n$ is Zariski-dense in $\mathbb{C}^n$ (that is, the smallest Zariski-closed subset of $\mathbb{C}^n$ containing $\mathbb{R}^n$ is $\mathbb{C}^n$ itself); and
3. show that every non-empty *Zariski-open* subset of $\mathbb{C}^n$ (that is, the complement of a Zariski-closed set) is dense in $\mathbb{C}^n$.

On the other hand, in some other aspects the Zariski topology resembles the Euclidean topology:

1. show that Zariski-open subsets of $\mathbb{C}^n$ are also open in the Euclidean topology;
2. determine the image of the map $\phi : \mathbb{C}^2 \to \mathbb{C}^3,\ (x_1, x_2) \to (x_1, x_1x_2, x_1(1+x_2))$, and show that its Zariski closure coincides with its Euclidean closure.

If you solved the last exercise correctly, then you found that the image is some Zariski-closed subset minus some Zariski-closed subset plus some other Zariski closed subset. In general, the subsets of $\mathbb{C}^n$ that are generated by the Zariski-closed sets under (finitely many of) the operations $\cup, \cap$, and complement, are called *constructible sets*. An important result due to the French mathematician Claude Chevalley (1909-1984) says that the image of a constructible set under a polynomial map $\mathbb{C}^n \to \mathbb{C}^m$ is again a constructible set. Another important fact is that the Euclidean closure of a constructible set equals its Zariski closure.

From undergraduate courses we know that $\mathbb{C}$ is an *algebraically closed field*, that is, that every non-constant univariate polynomial $f \in \mathbb{C}[x]$ has a root. The following multivariate analogue of this statement is the second major theorem of Hilbert's that we will need.

Theorem 6.1.4 (Hilbert's weak Nullstellensatz). *Let I be an ideal in $\mathbb{C}[\underline{x}]$ that is not equal to all of $\mathbb{C}[\underline{x}]$. Then there exists a point $\xi = (\xi_1, \ldots, \xi_n)$ such that $f(\xi) = 0$ for all $f \in I$.*

The theorem is also true with $\mathbb{C}$ replaced by any other algebraically closed field. But we will give a self-contained proof that uses the fact that $\mathbb{C}$ is not countable.

Lemma 6.1.5. *Let U, V be vector spaces over $\mathbb{C}$ of countably infinite dimension, let $A(x) : U \otimes \mathbb{C}[x] \to V \otimes \mathbb{C}[x]$ be a $\mathbb{C}[x]$-linear map, and let $v(x) \in V \otimes \mathbb{C}[x]$ be a target vector. Suppose that for all $\xi \in \mathbb{C}$ there is a $u \in U$ such that $A(\xi)u = v(\xi)$. Then there exists a $u(x) \in U \otimes \mathbb{C}(x)$ such that $Au(x) = v(x)$.*

Proof. Suppose, on the contrary, that no such $u(x)$ exists. This means that the image under $A(x)$ of the $\mathbb{C}(x)$-vector space $U \otimes \mathbb{C}(x)$ does not contain $v(x)$. Let $F(x)$ be a $\mathbb{C}(x)$-linear function on $V \otimes \mathbb{C}(x)$ taking the value 0 on $A(U \otimes \mathbb{C}(x))$ and 1 on $v(x)$; such a function exists and is determined by its values $f_1(x), f_2(x), f_3(x), \ldots \in \mathbb{C}(x)$ on a $\mathbb{C}$-basis $v_1, v_2, v_3, \ldots$ of V. Since $\mathbb{C}$ is uncountable there is a value $\xi \in \mathbb{C}$ where all f_i are defined, so that $F(\xi)$

is a well-defined linear function on V. Now we have $F(\xi)A(\xi)u = 0$ for all $u \in U$ but $F(\xi)v(\xi) = 1$, contradicting the assumption that $A(\xi)u = v(\xi)$ has a solution. □

Proof of the weak Nullstellensatz. We proceed by induction on n. For $n = 0$ the statement is just that any proper ideal of $\mathbb{C}$ is 0. Now suppose that $n > 0$ and that the statement is true for $n-1$. By Hilbert's basis theorem, the ideal I is generated by finitely many polynomials $f_1, \ldots, f_k$. If there exists a value $\xi \in \mathbb{C}$ for x_n such that the ideal in $\mathbb{C}[x_1, \ldots, x_{n-1}]$ generated by $f_{1,\xi} := f_1(x_1, \ldots, x_{n-1}, \xi), \ldots, f_{k,\xi} := (x_1, \ldots, x_{n-1}, \xi)$ does not contain 1, then we can use the induction hypothesis and we are done. Suppose therefore that no such ξ exists, that is, that 1 can be written as a $\mathbb{C}[x_1, \ldots, x_{n-1}]$-linear combination

$$1 = \sum_j c_{j,\xi} f_{j,\xi}$$

for every $\xi \in \mathbb{C}$. We will use this fact in two ways. First, note that this means that

$$\sum_j c_{j,\xi} f_j = 1 + (x_n - \xi) g_\xi$$

for some polynomial $g_\xi \in \mathbb{C}[x_1, \ldots, x_n]$. Put differently, $(x_n - \xi)$ has a *multiplicative inverse modulo the ideal* I for each $\xi \in \mathbb{C}$. But then every univariate polynomial in x_n, being a product of linear ones since $\mathbb{C}$ is algebraically closed, has such a multiplicative inverse. Since 1 does not lie in I, this implies that $I \cap \mathbb{C}[x_n] = \{0\}$.

Second, by Lemma 6.1.5 applied to $U = \mathbb{C}[x_1, \ldots, x_{n-1}]^k$, $V = \mathbb{C}[x_1, \ldots, x_{n-1}]$, $x = x_n$, and $A(c_1, \ldots, c_k) = \sum_{i=1}^k c_j f_j$, we can write

$$1 = \sum_{i=1}^k c_j(x_n) f_j,$$

where each $c_j(x_n)$ lies in $\mathbb{C}[x_1, \ldots, x_{n-1}](x_n)$. Letting $D(x_n) \in \mathbb{C}[x_n] \setminus 0$ be a common denominator of the c_j and setting $c'_j := Dc_j \in \mathbb{C}[x_1, \ldots, x_n]$, we find that

$$D(x_n) = \sum_{i=1}^k c'_j f_j \in I.$$

But this contradicts our earlier conclusion that I does not contain non-zero polynomials in x_n only. □

The Nulstellensatz has many applications to combinatorial problems.

Exercise 6.1.6. Let $G = (V, E)$ be a finite, undirected graph with vertex set V and edge set $E \subseteq \binom{V}{2}$. A *proper k-colouring* of G is a map $c : V \to [k]$ with the property that $c(i) \neq c(j)$ whenever $\{i, j\} \in E$. To G we associate the polynomial ring $\mathbb{C}[x_i \mid i \in V]$ and its ideal I generated by the following polynomials:

$$x_i^k - 1 \text{ for all } i \in V; \text{ and } x_i^{k-1} + x_i^{k-2}x_j + \ldots + x_j^{k-1} \text{ for all } \{i,j\} \in E.$$

Prove that G has a proper k-colouring if and only if $1 \notin I$.

Two important maps set up a beautiful duality between geometry and algebra. First, we have the map $\mathcal{V}$ that sends a subset $S \subseteq \mathbb{C}[\underline{x}]$ to the variety $\mathcal{V}(S)$ that it defines; and second, the map $\mathcal{I}$ that sends a subset $X \subseteq \mathbb{C}^n$ to the ideal $\mathcal{I}(X) \subseteq \mathbb{C}[\underline{x}]$ of all polynomials that vanish on all points in X. The following properties are straightforward:

1. if $S \subseteq S'$ then $\mathcal{V}(S) \supseteq \mathcal{V}(S')$;
2. if $X \subseteq X'$ then $\mathcal{I}(X) \supseteq \mathcal{I}(X')$;
3. $X \subseteq \mathcal{V}(\mathcal{I}(X))$;
4. $S \subseteq \mathcal{I}(\mathcal{V}(S))$;
5. $\mathcal{V}(\mathcal{I}(\mathcal{V}(S))) = \mathcal{V}(S)$; and
6. $\mathcal{I}(\mathcal{V}(\mathcal{I}(X))) = \mathcal{I}(X)$.

For instance, in (5) the containment $\supseteq$ follows from (3) applied to $X = \mathcal{V}(S)$ and the containment $\subseteq$ follows from (4) applied to S and then (1) applied to $S \subseteq S' := \mathcal{I}(\mathcal{V}(S))$.

This shows that $\mathcal{V}$ and $\mathcal{I}$ set up an inclusion-reversing bijection between sets of the form $\mathcal{V}(S) \subseteq \mathbb{C}^n$—that is, affine varieties in $\mathbb{C}^n$—and sets of the form $\mathcal{I}(X) \subseteq \mathbb{C}[\underline{x}]$. Sets of the latter form are always ideals, but not all ideals are of this form, as the following example shows.

Example 6.1.7. Suppose that $n = 1$, fix a natural number k, and let I_k be the ideal in $\mathbb{C}[x_1]$ generated by x_1^k. Then $\mathcal{V}(I) = \{0\}$ and $\mathcal{I}(\mathcal{V}(\mathcal{I}))$ is the ideal generated by x_1. So for $k > 1$ the ideal I_k is not of the form $\mathcal{I}(X)$ for any subset of $\mathbb{C}^n$.

This example exhibits a necessary condition for an ideal to be of the form $\mathcal{I}(X)$ for some set X—it must be *radical*.

Definition 6.1.8. The *radical* of an ideal $I \subseteq \mathbb{C}[\underline{x}]$ is the set of all polynomials f of which some positive power lies in I; it is denoted $\sqrt{I}$. The ideal I is called *radical* if $I = \sqrt{I}$.

Indeed, suppose that $I = \mathcal{I}(X)$ and suppose that $f \in \mathbb{C}[\underline{x}]$ has $f^k \in I$ for some $k > 0$. Then f^k vanishes on X and hence so does f, and hence $f \in \mathcal{I}(X) = I$. This shows that I is radical.

Exercise 6.1.9. Show that, for general ideals I, $\sqrt{I}$ is an ideal containing I.

The second important result of Hilbert's that we will need is that the condition that I be radical is also *sufficient* for I to be the vanishing ideal of some set X.

Theorem 6.1.10 (Hilbert's Nullstellensatz)**.** *Suppose that $I \subseteq \mathbb{C}[\underline{x}]$ is a radical ideal. Then $\mathcal{I}(\mathcal{V}(I)) = I$.*

Proof. This follows from the weak Nullstellensatz using *Rabinowitsch's trick* from 1929. Let g be a polynomial vanishing on all common roots of the polynomials in I. Introducing an auxilliary variable t, we have that the ideal in $\mathbb{C}[\underline{x}, t]$ generated by I and $tg-1$ does not have any common zeroes. Hence by the weak Nullstellensatz 1 can be written as

$$1 = a(tg-1) + \sum_{i=1}^{k} c_j(\underline{x}, t) f_j, \;\; a, c_j \in \mathbb{C}[\underline{x}, t], \;\; f_j \in I.$$

Replacing t on both sides by $1/g$ we have

$$1 = \sum_j c_j(\underline{x}, 1/g) f_j.$$

Multiplying both sides with a suitable power g^d eliminates g from the denominators and hence expresses g^d as a $\mathbb{C}[\underline{x}]$-linear combination of the f_j. Hence $g^d \in I$ and therefore $f \in I$ since I is radical. □

We have thus set up an inclusion-reversing bijection between closed subsets of $\mathbb{C}^n$ and radical ideals in $\mathbb{C}[\underline{x}]$. It is instructive to see what this bijection does with the smallest closed subsets consisting of a single point $p = (p_1, \ldots, p_n) \in \mathbb{C}^n$. The ideal $\mathcal{I}(p) := \mathcal{I}(\{p\})$ of polynomials vanishing on p is generated by $x_1 - p_1, \ldots, x_n - p_n$ (check this). This is a *maximal ideal* (that is, an ideal which is maximal among the *proper* ideals of $\mathbb{C}[x_1, \ldots, x_n]$), since the quotient by it is the field $\mathbb{C}$. This follows from the fact that, by definition, $\mathcal{I}$ is the kernel of the homomorphism of $\mathbb{C}$-algebras $\mathbb{C}[x_1, \ldots, x_n] \to \mathbb{C}, \; f \mapsto f(p)$ and that this homomorphism is surjective. Conversely, suppose that I is a maximal ideal. Then it is radical—indeed, if the radical were strictly larger than I, it would contain 1 by maximality, but then some power of 1 would be in I, a contradiction. Hence by the Nullstellensatz there exists a non-empty subset X of $\mathbb{C}^n$ such that $I = I(X)$. But then for any point p in X we have that $\mathcal{I}(p)$ is a radical ideal containing I, hence equal to I by maximality. We have thus proved the following corollary of the Nullstellensatz.

Corollary 6.1.11. *The map sending p to $\mathcal{I}(p)$ is a bijection between points in $\mathbb{C}^n$ and maximal ideals in $\mathbb{C}[\underline{x}]$.*

6.2 Regular functions and maps

Definition 6.2.1. Let X be an affine variety in $\mathbb{C}^n$. Then a *regular function* on X is by definition a $\mathbb{C}$-valued function of the form $f|_X$ where $f \in \mathbb{C}[\underline{x}]$.

Regular functions form a commutative $\mathbb{C}$-algebra with 1, denoted $\mathbb{C}[X]$ (or sometimes $\mathcal{O}(X)$) and sometimes called the *coordinate ring* of X. By definition, $\mathbb{C}[X]$ is the image of the restriction map $\mathbb{C}[\underline{x}] \to \{\mathbb{C}-\text{valued functions on } X\}$, $f \mapsto f|_X$. Hence it is isomorphic to the quotient algebra $\mathbb{C}[\underline{x}]/\mathcal{I}(X)$.

Example 6.2.2.

1. If X is a d-dimensional subspace of $\mathbb{C}^n$, then $\mathcal{I}(X)$ is generated by the space $X^0 \subseteq (\mathbb{C}^n)^*$ of linear functions vanishing on X. If $y_1, \ldots, y_d \in (\mathbb{C}^n)^*$ span a vector space complement of X^0, then modulo $\mathcal{I}(X)$ every polynomial in the x_i is equal to a unique polynomial in the y_j. This shows that $\mathbb{C}[X] = \mathbb{C}[y_1, \ldots, y_d]$ is a polynomial ring in d variables. In terminology to be introduced below, X is *isomorphic* to the variety $\mathbb{C}^d$.

2. Consider the variety X of $(m+1) \times (m+1)$-matrices of the shape

$$\begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}$$

with x an $m \times m$-matrix and y a complex number satisfying $\det(x)y = 1$. Then $\mathbb{C}[X] = \mathbb{C}[(x_{ij})_{ij}, y]/(\det(x)y - 1)$. The map $y \mapsto 1/\det(x)$ sets up an isomorphism of this algebra with the algebra of rational functions in the variables x_{ij} generated by the x_{ij} and $1/\det(x)$. We therefore also write $\mathbb{C}[X] = \mathbb{C}[(x_{ij})_{ij}, 1/\det(x)]$. Note that X is a group with respect to matrix multiplication, isomorphic to GL_n. This is the fundamental example of an *algebraic group*; here *algebraic* refers to the variety structure of X.

3. Consider the variety $X = M_{k,m}^{\leq l}$ of all $k \times m$-matrices all of whose $(l+1) \times (l+1)$-minors (that is, determinants of $(l+1) \times (l+1)$-submatrices) vanish. Elementary linear algebra shows that X consists of all matrices of rank at most l, and that such matrices can always be written as AB with $A \in M_{k,l},\ B \in M_{l,m}$.

Remark 6.2.3. In these notes a $\mathbb{C}$-*algebra* is always a vector space A over $\mathbb{C}$ together with an associative, bilinear multiplication $A \times A \to A$, such that A contains an element 1 for which $1a = a = a1$ for all $a \in A$. A *homomorphism* from A to a $\mathbb{C}$-algebra B is a $\mathbb{C}$-linear map $\phi : A \to B$ satisfying $\phi(1) = 1$ and $\phi(a_1 a_2) = \phi(a_1)\phi(a_2)$ for all $a_1, a_2 \in A$. Most algebras that we will encounter are commutative.

Just like group homomorphisms are the natural maps between groups and continuous maps are the natural maps between topological spaces, *regular maps* are the natural maps between affine varieties.

Definition 6.2.4. A *regular map* from an affine variety X to $\mathbb{C}^m$ is a map $\phi : X \to \mathbb{C}^m$ of the form $\phi : x \mapsto (f_1(x), \ldots, f_m(x))$ with $f_1, \ldots, f_m$ regular functions on X. If $Y \subseteq \mathbb{C}^m$ is an affine variety containing the image of ϕ, then we also call ϕ a regular map from X to Y.

Exercise 6.2.5. If ψ is a regular map from Y to a third affine variety Z, then $\psi \circ \phi$ is a regular map from X to Z.

Lemma 6.2.6. *If $X \subseteq \mathbb{C}^n$ and $Y \subseteq \mathbb{C}^m$ are affine varieties, and if $\phi : X \to Y$ is a regular map, then the map $\phi^* : f \mapsto f \circ \phi$ is a homomorphism of $\mathbb{C}$-algebras from $\mathbb{C}[Y]$ $\mathbb{C}[X]$.*

Proof. Suppose that ϕ is given by regular functions $(f_1, \ldots, f_m)$ on X. Then ϕ^* sends the regular function $h|_Y \in \mathbb{C}[Y]$, where h is a polynomial in the coordinates $y_1, \ldots, y_m$ on $\mathbb{C}^m$, to the function $h(f_1, \ldots, f_m)$, which is clearly a regular function on $\mathbb{C}[X]$. This shows that ϕ^* maps $\mathbb{C}[Y]$ to $\mathbb{C}[X]$. One readily verifies that ϕ^* is an algebra homomorphism. $\square$

Note that if $\psi : Y \to Z$ is a second regular map, then $\phi^* \circ \psi^* = (\psi \circ \phi)^*$.

Definition 6.2.7. If $X \subseteq \mathbb{C}^n$ and $Y \subseteq \mathbb{C}^m$ are affine varieties, then an *isomorphism* from X to Y is a regular map whose inverse is also a regular map. The varieties X and Y are called *isomorphic* if there is an isomorphism from X to Y.

Lemma 6.2.8. *If $X \subseteq \mathbb{C}^n$ and $Y \subseteq \mathbb{C}^m$ are isomorphic varieties, then $\mathbb{C}[X]$ and $\mathbb{C}[Y]$ are isomorphic $\mathbb{C}$-algebras.*

Proof. If $\phi : X \to Y,\ \psi : Y \to X$ are a regular maps such that $\psi \circ \phi = \mathrm{id}_X$ and $\phi \circ \psi = \mathrm{id}_Y$, then $\phi^* \circ \psi^* = \mathrm{id}_{\mathbb{C}[X]}$ and $\psi^* \circ \phi^* = \mathrm{id}_{\mathbb{C}[Y]}$, hence these two algebras are isomorphic. $\square$

Example 6.2.9. The affine variety $X = \mathbb{C}^1$ and the affine variety $Y = \{(x, y) \in \mathbb{C}^2 \mid y - x^2 = 0\}$ are isomorphic, as the regular maps $\phi : X \to Y,\ t \mapsto (t, t^2)$ and $\psi : Y \to X,\ (x, y) \mapsto x$ show.

Exercise 6.2.10. Prove that $X = \mathbb{C}^1$ is not isomorphic to the variety $Z = \{(x, y) \in \mathbb{C}^2 \mid xy - 1 = 0\}$.

6.3 The quotient map

Let G be a group and let W be a finite-dimensional G-module such that $\mathbb{C}[W] = \bigoplus S^k W^*$ is a completely reducible G-module. By Hilbert's finiteness theorem, we know that the algebra $\mathbb{C}[W]^G$ of G-invariant polynomials is a finitely generated algebra. Let $f_1, \ldots, f_k$ be a generating set of this algebra. Then we have a polynomial map

$$\pi : W \to \mathbb{C}^k,\ w \mapsto (f_1(w), \ldots, f_k(w)).$$

This map is called the *quotient map*, because in some sense, which will become clear below, the image of this map parameterises G-orbits in W.

Example 6.3.1. Take $G := \{-1, 1\}$ with its action on $W := \mathbb{C}^2$ where -1 sends (x, y) to $(-x, -y)$. The invariant ring $\mathbb{C}[W]^G$ is generated by the polynomials $f_1 := x^2,\ f_2 := xy,\ f_3 := y^2$ (check this). Thus the quotient map is $\pi : \mathbb{C}^2 \to \mathbb{C}^3,\ (x, y) \mapsto (x^2, xy, y^2)$. Let u, v, w be the standard coordinates on $\mathbb{C}^3$, and note that the image of π is contained in the affine variety $Z \subseteq \mathbb{C}^3$ with equation $v^2 - uw$. Indeed, π is surjective onto Z: let u, v, w be complex numbers such that $v^2 = uw$. Let $x \in \mathbb{C}$ be a square root of u. If $x \neq 0$, then set $y := v/x$ so that $v = xy$ and $w = v^2/u = x^2y^2/x^2 = y^2$. If $x = 0$, then let y be a square root of w. In both cases $\pi(x, y) = (u, v, w)$. Note that the fibres of π over every point (u, v, w) are orbits of G.

Example 6.3.2. Take $G := S_n$ with its standard action on $W := \mathbb{C}^n$. Recall that the invariants are generated by the elementary symmetric polynomials $\sigma_1(x) = \sum_i x_i, \ldots, \sigma_n(x) = \prod_i x_i$. This gives the quotient map

$$\pi : \mathbb{C}^n \to \mathbb{C}^n,\ (x_1, \ldots, x_n) \mapsto (\sigma_1(x), \ldots, \sigma_n(x)).$$

We claim that the image of this map is all of $\mathbb{C}^n$. Indeed, for any n-tuple $(c_1, \ldots, c_n) \in \mathbb{C}^n$ consider the polynomial

$$f(t) := t^n - c_1 t^{n-1} + \ldots + (-1)^n c_n.$$

As $\mathbb{C}$ is algebraically closed, this polynomial has n roots (counted with multiplicities), so that we can also write

$$f(t) = (t - x_1) \cdots (t - x_n);$$

expanding this expression for f and comparing with the above gives $\pi(x_1, \ldots, x_n) = (c_1, \ldots, c_n)$. Note furthermore that any other n-tuple $(x'_1, \ldots, x'_n)$ with this property is necessarily some permutation of the x_i, since the roots of f are determined by $(c_1, \ldots, c_n)$.

This means that we can think of $\pi : \mathbb{C}^n \to \mathbb{C}^n$ as follows: every S_n-orbit on the first copy of $\mathbb{C}^n$ is "collapsed" by π to a single point on the second copy of $\mathbb{C}^n$, and conversely, the fibre of π above any point in the second copy of $\mathbb{C}^n$ consists of a single S_n-orbit. Thus *the second copy of $\mathbb{C}^n$ parameterises S_n-orbits on $\mathbb{C}^n$.*

Example 6.3.3. Let the group $G := \mathrm{SL}_2$ act on the $W := M_2(\mathbb{C})$ of 2×2-matrices by left multiplication. If A has rank 2, then left-multiplying by a suitable $g \in \mathrm{SL}_2$ gives

$$gA = \begin{bmatrix} 1 & 0 \\ 0 & \det(A) \end{bmatrix}.$$

Now det is a polynomial on $M_2(\mathbb{C})$ which is SL_2-invariant. We claim that it generates the invariant ring. Indeed, suppose $f \in \mathbb{C}[a_{11}, a_{12}, a_{21}, a_{22}]$ is any invariant. Then for A of rank two we find that

$$f(A) = f(gA) = f(1, 0, 0, \det(A)) =: h(\det A)$$

where h is a polynomial in 1 variable. Since both f and h are continous, and since the rank-two matrices are dense in M_2, we find that this equality actually holds for all matrices. Hence $f(A) = h(\det A)$ for all A, so f is in the algebra generated by det.

In this case the quotient map $\pi : M_2(\mathbb{C}) \to \mathbb{C}$ is just the map $A \mapsto \det(A)$. The fibre above a point $d \in \mathbb{C}^*$ is just the set of 2×2-matrices of determinant d, which is a Zariski-closed set and a single SL_2-orbit. The fibre above $0 \in \mathbb{C}$ is the set of all matrices of rank ≤ 1. This is, of course, also a Zariski closed set, but *not* a single orbit—indeed, it consists of the closed orbit consisting of the zero matrix and the non-closed orbit consisting of all matrices of rank 1. Note that the latter orbit has 0 in its closure.

These two examples illustrate the general situation: for *finite* G the fibres of the quotient map are precisely the orbits of G, while for *infinite* G they are certain G-stable closed sets.

Theorem 6.3.4. *Let Z denote the Zariski closure of $\pi(W)$, that is, the set of all points in $\mathbb{C}^m$ that satisfy all polynomial relations that are satisfied by the invariants $f_1, \ldots, f_k$. The quotient map π has the following properties:*

1. *$\pi(gw) = \pi(w)$ for all $g \in G,\ w \in W$;*
2. *the fibres of π are G-stable, Zariski-closed subsets of W;*
3. *for any regular (polynomial) map $\psi : W \to \mathbb{C}^m$ that satisfies $\psi(gw) = \psi(w)$ for all $g \in G$ there exists a unique regular map $\phi : Z \mapsto \mathbb{C}^m$ such that $\phi \circ \pi = \psi$.*
4. *π is surjective onto Z;*

Proof.

1. $\pi(gw) = (f_1(gw), \ldots, f_k(gw)) = (f_1(w), \ldots, f_k(w))$ because the f_i are invariant.
2. If $w \in \pi^{-1}(z)$, then $\pi(gw) = \pi(w) = z$, so $gw \in \pi^{-1}(z)$.
3. Let $y_1, \ldots, y_m$ be the standard coordinates on $\mathbb{C}^m$. Then $y_i \circ \psi$ is a G-invariant polynomial on W for all i. As the f_j generate these polynomials, we may write $y_i \circ \psi$ as $g_i(f_1, \ldots, f_k)$ for some k-variate polynomial g_i. Now the regular map $\phi : Z \mapsto U,\ z \mapsto (g_1(z), \ldots, g_m(z))$ has the required property. Notice that the g_i need not be unique. However, the map $Z \to \mathbb{C}^m$ with the required property *is* unique: if ϕ_1, ϕ_2 both have the property, then necessarily $\phi_1(\pi(w)) = \phi_2(\pi(w)) = \psi(w)$ for all $w \in W$, so that ϕ_1 and ϕ_2 agree on the subset $\operatorname{im} \pi$ of Z. Since Z is the Zariski closure of this set, ϕ_1 and ϕ_2 need to agree everywhere. (In fact $Z = \operatorname{im} \pi$ as we will see shortly.)
4. Let $z \in Z$. This means that the coordinates of z satisfy all polynomial relations satisfied by $f_1, \ldots, f_k$, hence there exists a homomorphism $\phi : \mathbb{C}[W]^G = \mathbb{C}[f_1, \ldots, f_k] \to \mathbb{C}$ of $\mathbb{C}$-algebras sending f_i to z_i. The kernel of this homomorphism is a maximal ideal M_z in $\mathbb{C}[W]^G$. We claim that there exists a maximal ideal M' in $\mathbb{C}[W]$ whose intersection with $\mathbb{C}[W]^G$ is M_z. Indeed, let I be the ideal in $\mathbb{C}[W]$ generated by M_z. We only need to show that $I \neq \mathbb{C}[W]$; then the axiom of choice implies the existence of

a maximal ideal containing I. Suppose, on the contrary, that $I \ni 1$, and write

$$1 = \sum_{i=1}^{l} a_i h_i \text{ with all } a_i \in \mathbb{C}[W], h_i \in M_z.$$

Since $\mathbb{C}[W]$ is completely reducible as a G-module, there exists a Reynolds operator ρ. Applying ρ to both sides of the equality yields

$$1 = \sum_{i=1}^{l} \rho(a_i) h_i,$$

where the $\rho(a_i)$ are in $\mathbb{C}[W]^G$. But this means that 1 lies in M_z, a contradiction to the maximality of the latter ideal. This proves the claim that such an M' exists. The maximal ideal M' is the kernel of evaluation at some point $w \in W$ by the discussion of maximal ideals after the Nullstellensatz. Thus we have found a point $w \in W$ with the property that evaluating f_i at w gives z_i. Thus $\pi(w) = z$ and we are done.

□

Remark 6.3.5. By (3) Z is independent of the choice of generators of $\mathbb{C}[W]^G$ in the following sense: any other choice of generators of $\mathbb{C}^G$ yields a variety Z' with a G-invariant map $\pi' : W \to Z'$, and (3) shows that there exist regular maps $\phi : Z \to Z'$ and $\phi' : Z' \to Z$ such that $\phi \circ \phi' = \mathrm{id}_Z$ and $\phi' \circ \phi = \mathrm{id}_{Z'}$.

Exercise 6.3.6. Let $G = \mathbb{C}^*$ act on $W = \mathbb{C}^4$ by $t(x_1, x_2, y_1, y_2) = (tx_1, tx_2, t^{-1}y_1, t^{-1}y_2)$.

1. Find generators $f_1, \ldots, f_k$ of the invariant ring $\mathbb{C}[W]^G$.
2. Determine the image Z of the quotient map $\pi : W \to \mathbb{C}^k,\ w \mapsto (f_1(w), \ldots, f_k(w))$.
3. For every $z \in Z$ determine the fibre $\pi^{-1}(z)$.

Chapter 7

The null-cone

Let G be a group and let W be a finite-dimensional G-module. We have seen in Example 6.3.3 that G-orbits on W cannot always be separated by invariant polynomials on W. Here is another example of this phenomenon.

Example 7.0.7. Let $G = \mathrm{GL}_n(\mathbb{C})$ act on $W = M_n(\mathbb{C})$ by conjugation. We have seen in week 1 that the invariant ring is generated by the coefficients of the characteristic polynomial χ. This means that the map π sending A to χ_A is the quotient map. By exercise 1.5.5 each fibre $\pi^{-1}(p)$ of π, where p is a monic univariate polynomial of degree n, contains a unique conjugacy class of diagonalisable matrices. This conjugacy class is in fact Zariski closed, since it is given by the additional set of equations

$$(A - \lambda_1) \cdots (A - \lambda_k) = 0$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of p. We claim that all other (non-diagonalisable) conjugacy classes are not Zariski closed. Indeed, if A is not diagonalisable, then after a conjugation we may assume that A is of the form $D+N$ with D diagonal and N strictly upper triangular (e.g., move A to Jordan normal form). Conjugating $A = D + N$ with a diagonal matrix of the form $\mathrm{diag}(t^{n-1}, \ldots, t^0)$, $t \in \mathbb{C}^*$ multiplies the (i, j)-entry of A with t^{j-i}. Hence for $i \geq j$ the entries of A do not change, while for $i < j$ they are multiplied by a positive power of t. Letting t tend to 0 we find that the result tends to D. Hence D lies in the Euclidean closure of the conjugacy class of A, hence also in the Zariski closure.

Note in particular the case where $D = 0$, i.e., where A is nilpotent. Then the characteristic polynomial of A is just x^n, i.e., all invariants with zero constant term vanish on A, and the argument above shows that 0 lies in the closure of the orbit of A. This turns out to be a general phenomenon.

Definition 7.0.8. The *null-cone* N_W of the G-module W is the set of all vectors $w \in W$ on which all G-invariant polynomials with zero constant term vanish.

Thus, the null-cone of the module $M_n(\mathbb{C})$ with the conjugation action of $\mathrm{GL}_n(\mathbb{C})$ on $M_n(\mathbb{C})$ consists of all nilpotent matrices, and the null-cone of the module $M_n(\mathbb{C})$ with the action of $\mathrm{SL}_n(\mathbb{C})$ by left multiplication consists of all singular matrices.

Exercise 7.0.9. Check the second statement by verifying that the invariant ring is generated by det.

Remark 7.0.10. Suppose that the invariant ring $\mathbb{C}[W]^G$ is generated by finitely many invariant functions $f_1, \ldots, f_k$, each with zero constant term. Let π be the corresponding quotient map $W \to \mathbb{C}^k$. Prove that the null-cone N_W is just the fibre $\pi^{-1}(0)$ above $0 \in \mathbb{C}^k$.

Exercise 7.0.11. Show that if G is finite, then the null-cone consists of 0 alone.

We want to describe the structure of the null-cone for one class of groups, namely, *tori.* The resulting structure theorem actually carries over *mutatis mutandis* to general *semisimple algebraic groups*, and we will see some examples of that fact later.

Definition 7.0.12. The *n-dimensional torus* T_n is the group $(\mathbb{C}^*)^n$.

For any n-tuple $\alpha = (a_1, \ldots, a_n) \in \mathbb{Z}^n$ we have a homomorphism

$$\rho_\alpha : \mathrm{T}_n \to \mathbb{C}^*, \ t = (t_1, \ldots, t_n) \mapsto t^\alpha := t_1^{a_1} \cdots t_n^{a_n},$$

and hence a one-dimensional representation of T_n. Let W be a finite direct sum of m such one-dimensional representations of T_n, so W is determined by a sequence $A = (\alpha_1, \ldots, \alpha_m)$ of lattice points in $\mathbb{Z}^n$ (possibly with multiplicities). Relative to a basis consisting of one vector for each α_i, the representation $\mathrm{T}_n \to \mathrm{GL}(W)$ is just the matrix representation

$$t \mapsto \begin{bmatrix} t^{\alpha_1} & & \\ & \ddots & \\ & & t^{\alpha_m} \end{bmatrix}.$$

We think of $\alpha_i = (a_{i1}, \ldots, a_{in})$ as the i-th row of the $m \times n$-matrix A. Let $x = (x_1, \ldots, x_m)$ denote the corresponding coordinate functions on W. Then $(t_1, \ldots, t_n)$ acts on the variable x_i by $\prod_{j=1}^n t_j^{-a_{i,j}}$—recall that the action on functions involves taking an inverse of the group element—and hence on a monomial x^u, $u \in \mathbb{N}^m$ by

$$\prod_{i=1}^m \prod_{j=1}^n t_j^{-u_i a_{i,j}} = \prod_{j=1}^n t_j^{(-uA)_j},$$

where uA is the row vector obtained by left-multiplying A by u. This implies two things: first, all monomials appearing in any T_n-invariant polynomial on W are themselves invariant, so that $\mathbb{C}[W]^{\mathrm{T}_n}$ is spanned by monomials in the x_i, and second, the monomial x^u is invariant if and only if $uA = 0$.

Definition 7.0.13. For $w \in W$ let $\mathrm{supp}(w)$, the *support* of w, be the set of $\alpha_i \in \mathbb{Z}^n$ for which $x_i(w) \neq 0$.

Theorem 7.0.14. *The null-cone of the* T_n*-module* W *consists of all vectors* w *such that* 0 *does not lie in the convex hull of* $\mathrm{supp}(w) \subseteq \mathbb{Z}^n \subseteq \mathbb{R}^n$.

Proof. Suppose first that 0 does not lie in that convex hull. Then there exists a vector $\beta = (b_1, \ldots, b_n) \in \mathbb{Z}^n$ such that $\beta \cdot \alpha > 0$ for all $\alpha \in \mathrm{supp}(w)$; here $\cdot$ is the dot product. This means that the vector

$$\lambda(t) = (t^{b_1}, \ldots, t^{b_n}),\ t \in \mathbb{C}^*$$

acts by a strictly positive power of t on all non-zero components of w. Hence for $t \to 0$ the vector $\lambda(t)w$ tends to 0. Hence each T_n-invariant polynomial f on W satisfies

$$f(w) = f(\lambda(t)w) \to f(0),\ t \to 0;$$

here the equality follows from the fact that f is invariant and the limit follows from the fact that f is continuous. Hence w is in the null-cone N_V.

Conversely, suppose that 0 lies in the convex hull of the support of w. Then we may write 0 as $u_1\alpha_1 + \ldots + u_m\alpha_m$ where the u_i are natural numbers and not all zero and where $u_i > 0$ implies that α_i lies in the support of w. Then $uA = 0$, so x^u is a non-constant invariant monomial, which moreover does not vanish on w since the only variables x_i appearing in it have $x_i(w) \neq 0$. Hence w does not lie in the null-cone of T_n on W. □

Exercise 7.0.15. Let T be the group of invertible diagonal $n \times n$-matrices, and let T act on $M_n(\mathbb{C})$ by conjugation, that is,

$$t \cdot A := tAt^{-1},\ t \in T,\ A \in M_n(\mathbb{C}).$$

Prove that A lies in the null-cone of T on $M_n(\mathbb{C})$ if and only if there exists a permutation matrix P such that PAP^{-1} is strictly upper triangular.

Exercise 7.0.16. Let T be the group of diagonal 3×3-matrices with determinant 1. Let $U = \mathbb{C}^3$ be the standard T-module with action

$$\mathrm{diag}(t_1, t_2, t_3)(x_1, x_2, x_3) := (t_1x_1, t_2x_2, t_3x_3),$$

and consider the 9-dimensional T-module $W = U^{\otimes 2}$.

1. Show that $T \cong \mathrm{T}_2$.

2. Determine the irreducible T-submodules of W.

3. Draw the vectors α_i for W in the plane, and determine all possible supports of vectors in the null-cone of T on W.

Exercise 7.0.17. Let G be the group $\mathrm{SL}_2(\mathbb{C})$ acting on the space $U = \mathbb{C}^2$ and let W be the space S^2U with the standard action of G given by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} (x_{11}e_1^2 + x_{12}e_1e_2 + x_{22}e_2^2) = x_{11}(ae_1 + ce_2)^2 + x_{12}(ae_1 + ce_2)(be_1 + de_2) + x_{22}(be_1 + de_2)^2,$$

where e_1, e_2 are the standard basis of $\mathbb{C}^2$.

1. Determine the invariant ring $\mathbb{C}[W]^G$.
2. Determine the fibres of the quotient map.
3. Determine the null-cone.

Chapter 8

Molien's theorem and self-dual codes

Let $W \bigoplus_{d=0}^{\infty} W_d$ be a direct sum of finite dimensional (complex) vector spaces W_d. The *Hilbert series* (or *Poincaré series*) $H(W,t)$ is the formal power series in t defined by

$$H(V,t) := \sum_{d=0}^{\infty} \dim(V_d) t^d, \tag{8.1}$$

and encodes in a convenient way the dimensions of the vector spaces W_d. In this lecture, W will usually be the vector space $\mathbb{C}[V]^G$ of polynomial invariants with respect to the action of a group G, where W_d is the subspace of invariants homogeneous of degree d.

Example 8.0.18. Taking the polynomial ring in one variable, the Hilbert series is given by $H(\mathbb{C}[x],t) = 1 + t + t^2 + \cdots = \frac{1}{1-t}$. Similarly, $H(\mathbb{C}[x_1,\ldots,x_n]) = \frac{1}{(1-t)^n}$.

Exercise 8.0.19. Let $f_1,\ldots,f_k \in \mathbb{C}[x_1,\ldots,x_n]$ be algebraically independent homogeneous polynomials, where f_i has degree d_i. Show that the Hilbert series of the subalgebra generated by the f_i is given by

$$H(\mathbb{C}[f_1,\ldots,f_k],t) = \frac{1}{\prod_{i=1}^{k}(1-t^{d_i})}. \tag{8.2}$$

Example 8.0.20. Consider the action of the group G of order 3 on $\mathbb{C}[x,y]$ induced by the linear map $x \mapsto \zeta_3 x$, $y \mapsto \zeta_3^{-1} y$, where ζ_3 is a third root of unity. Clearly, x^3, y^3 and xy are invariants and $\mathbb{C}[x,y]^G = \mathbb{C}[x^3,y^3,xy]$. In fact, x^3 and y^3 are algebraically independent, and

$$\mathbb{C}[x,y]^G = \mathbb{C}[x^3,y^3] \oplus \mathbb{C}[x^3,y^3]xy \oplus \mathbb{C}[x^3,y^3](xy)^2. \tag{8.3}$$

Since $H(\mathbb{C}[x^3,y^3],t) = \frac{1}{(1-t^3)^2}$, we obtain $H(\mathbb{C}[x,y]^G,t) = \frac{1+t^2+t^4}{(1-t^3)^2}$.

Exercise 8.0.21. Compute the Hilbert series of $\mathbb{C}[x^2,y^2,xy]$.

8.1 Molien's theorem

For finite groups G, it is possible to compute the Hilbert series directly, without prior knowlegde about the generators. This is captured in the following beautiful theorem of Molien.

Theorem 8.1.1 (Molien's Theorem)**.** *Let $\rho : G \to \mathrm{GL}(V)$ be a representation of a finite group on a finite dimensional vector space V. Then the Hilbert series is given by*

$$H(\mathbb{C}[V]^G, t) = \frac{1}{|G|}\sum_{g\in G}\frac{1}{\det(I-\rho(g)t)}. \tag{8.4}$$

Proof. Consider the action of G on $\mathbb{C}[V]$ induced by the representation ρ. Denote for $g \in G$ and $d \in \mathbb{N}$ by $L_d(g) \in \mathrm{GL}(\mathbb{C}[V]_d)$ the linear map corresponding to the action of $g \in G$ on the homogeneous polynomials of degree d. So $L_1(g) = \rho^*(g)$.

The linear map $\pi_d := \frac{1}{|G|}\sum_{g\in G} L_d(g)$ is a projection onto $\mathbb{C}[V]_d^G$. That is, $\pi_d(p) \in \mathbb{C}[V]_d^G$ for all $p \in \mathbb{C}[V]_d$ and π_d is the identity on $\mathbb{C}[V]_d^G$. It follows that $\mathrm{tr}(\pi_d) = \dim(\mathbb{C}[V]_d^G)$. This gives:

$$H(\mathbb{C}[V]^G, t) = \frac{1}{|G|}\sum_{g\in G}\sum_{d=0}^{\infty}\mathrm{tr}(L_d(g)). \tag{8.5}$$

Now lets fix an element $g \in G$ and compute the inner sum $\sum_{d=0}^{\infty}\mathrm{tr}(L_d(g))$. Pick a basis $x_1, \ldots, x_n$ of V^* that is a system of eigenvectors for $L_1(g)$, say $L_1(g)x_i = \lambda_i x_i$. Then the monomials in $x_1, \ldots, x_n$ of degree d for a system of eigenvectors of $L_d(g)$ with eigenvalues given by:

$$L_d(g)\cdot x_1^{d_1}\cdots x_n^{d_n} = \lambda_1^{d_1}\cdots\lambda_n^{d_n}\cdot x_1^{d_1}\cdots x_n^{d_n} \tag{8.6}$$

for all $d_1 + \cdots + d_n = d$. It follows that

$$\begin{aligned}\sum_{d=0}^{\infty} t^d\mathrm{tr}(L_d(g)) &= (1+\lambda_1 t+\lambda_1^2t^2+\cdots)\cdots(1+\lambda_n t+\lambda_n t^2+\cdots)\\ &= \frac{1}{1-\lambda_1 t}\cdots\frac{1}{1-\lambda_n t} = \frac{1}{\det(I-L_1(g)t)}.\end{aligned} \tag{8.7}$$

Using the fact that for every g the equality $\det(I-L_1(g)t) = \det(I-\rho(g^{-1})t)$ holds and combining equations (8.5) and (8.7), we arrive at

$$\begin{aligned}H(\mathbb{C}[V]^G, t) &= \frac{1}{|G|}\sum_{g\in G}\sum_{d=0}^{\infty}\mathrm{tr}(L_d(g))\\ &= \frac{1}{|G|}\sum_{g\in G}\frac{1}{\det(I-\rho(g^{-1})t)}\\ &= \frac{1}{|G|}\sum_{g\in G}\frac{1}{\det(I-\rho(g)t},\end{aligned} \tag{8.8}$$

where the last equality follows by changing the order in which we sum over G. This completes the proof. $\square$

Exercise 8.1.2. Let $U \subset V$ be finite dimensional vector spaces and let $\pi : V \to U$ be the identity on U. Show that $\mathrm{tr}(\pi) = \dim(U)$. Hint: write π as a matrix with respect to a convenient basis.

Example 8.1.3. Consider again the action of the group $G = \mathbb{Z}/3\mathbb{Z}$ on $\mathbb{C}[x, y]$ induced by the linear map $x \mapsto \zeta x$, $y \mapsto \zeta^{-1} y$, where ζ is a third root of unity. Using Molien's theorem, we find

$$H(\mathbb{C}[x,y]^G, t) = \frac{1}{3}\left(\frac{1}{(1-t)(1-t)} + \frac{1}{(1-\zeta t)(1-\zeta^2 t)} + \frac{1}{(1-\zeta^2 t)(1-\zeta t)}\right). \tag{8.9}$$

A little algebraic manipulation and the fact that $(1-\zeta t)(1-\zeta^2 t) = (1+t+t^2)$ shows this to be equal to

$$\frac{(1-t+t^2)(1+t+t^2)}{(1-t)^2(1-\zeta t)^2(1-\zeta^2 t)^2} = \frac{1+t^2+t^4}{(1-t^3)^2}. \tag{8.10}$$

Since this is equal to the Hilbert series of $\mathbb{C}[x^3, y^3, xy]$, we obtain as a by-product that the invariant ring is indeed generated by the three invariants x^3, y^3 and xy.

Exercise 8.1.4. Let G be the matrix group generated by $A, B \in \mathrm{GL}_2(\mathbb{C})$ given by

$$A := \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad B := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{8.11}$$

- Use Molien's theorem to prove that the Hilbert series of $\mathbb{C}[x, y]^G$ is given by $H(\mathbb{C}[x, y]^G, t) = \frac{1+t^6}{(1-t^4)^2}$.
- Find algebraically independent invariants f_1, f_2 of degree 4 and a third invariant f_3 of degree 6, such that $\mathbb{C}[x, y]^G = \mathbb{C}[f_1, f_2] \oplus \mathbb{C}[f_1, f_2] f_3$.

8.2 Linear codes

A *linear code* is a linear subspace $C \subseteq \mathbb{F}_q^n$, where $\mathbb{F}_q$ is the field of q elements. The number n is called the *length* of the code. In the following, we will only consider *binary* codes, that is, $q = 2$. The *weight* $w(u)$ of a word $u \in \mathbb{F}_2^n$ is the number of nonzero positions in u, that is, $w(u) := |\{i \mid u_i = 1\}|$. The *Hamming distance* $d(u, v)$ between two words is defined as the number of positions in which u and v, differ: $d(u, v) = w(u - v)$.

A code $C \subseteq \mathbb{F}_2^n$ is called an $[n, k, d]$*-code* if the dimension of C is equal to k and the smallest Hamming distance between two distinct codewords is equal to d. In the setting of error correcting codes, messages are transmitted using words from the set of 2^k codewords. If at most $(d-1)/2$ errors are introduced (by noise) into a codeword, the original can still be recovered by finding the word in C at minimum distance from the distorted word. The higher d, the more errors can be corrected and the higher k, the higher the information rate.

Much information about a code, including the parameters d and k, can be read of from its *weight enumerator* W_C. This is the polynomial in x, y and homogeneous of degree n, defined by

$$W_C(x,y) := \sum_{i=0}^{n} A_i y^i x^{n-i}, \quad A_i := |\{u \in C \mid w(u) = i\}|. \tag{8.12}$$

Observe that the coefficient of x^n in W_C is always equal to 1, since C contains the zero word. The number 2^k of codewords equals the sum of the coefficients $A_0, \ldots, A_n$ and d is the smallest positive index i for which $A_i > 0$.

For a code $C \subseteq \mathbb{F}_2^n$, the *dual code* $C^\perp$ is defined by

$$C^\perp := \{u \in \mathbb{F}_2^n \mid u \cdot c = 0 \text{ for all } c \in C\}, \quad \text{where } u \cdot c := u_1 c_1 + \cdots + u_n c_n. \tag{8.13}$$

Exercise 8.2.1. Check that the dimensions of a code $C \subseteq \mathbb{F}_2^n$ and its dual $C^\perp$ sum to n.

The MacWilliams identity relates the weight enumerator of a code C and that of its dual $C^\perp$.

Proposition 8.2.2. *Let $C \subseteq \mathbb{F}_2^n$ be a code. The weight enumerator of $C^\perp$ satisfies*

$$W_{C^\perp}(x,y) = \frac{1}{|C|} W_C(x+y, x-y). \tag{8.14}$$

Exercise 8.2.3. Prove the MacWilliams identity. Hint: let

$$f(u) := \sum_{v \in \mathbb{F}_2^n} x^{n-w(v)} y^{w(v)} (-1)^{u \cdot v}, \tag{8.15}$$

and compute $\sum_{c \in C} f(c)$ in two ways.

A code is called *self-dual* if $C = C^\perp$. This implies that n is even and the dimension of C equals $n/2$. Furhermore, we have for every $c \in C$ that $c \cdot c = 0$ so that $w(c)$ is even. If every word in C has weight divisible by 4, the code is called *even*.

Exercise 8.2.4. An example of an even self-dual code is the *extended Hamming code* spanned by the rows of the matrix

$$\begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{8.16}$$

That this code is self-dual follows from the fact that it has dimension 4 and any two rows of the given matrix have dot product equal to 0. To see that it is an even code, observe that the rows have weights divisible by 4 and that for any two words u, v with weights divisible by four and $u \cdot v = 0$, also $u+v$ has weight divisible by four.

Consider an even, self-dual code C. Then its weight enumerator must satisfy

$$W_C(x,y) = W_C(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}), \quad W_C(x,y) = W_C(x, iy). \tag{8.17}$$

Here the first equality follows from Proposition 8.2.2 and the fact that $|C| = (\sqrt{2})^n$. The second equality follows from the fact that all weights are divisible by 4. But this means that W_C is invariant under the group G generated by the matrices

$$A := \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}, \quad B := \begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix}, \tag{8.18}$$

a group of 192 elements!

Exercise 8.2.5. Let $\zeta = e^{\frac{2\pi i}{8}}$ be a primitive 8-th root of unity. Show that the group G defined above is equal to the set of matrices

$$\zeta^k\begin{pmatrix}1 & 0\\ 0 & \alpha\end{pmatrix}, \quad \zeta^k\begin{pmatrix}0 & 1\\ \alpha & 0\end{pmatrix}, \quad \zeta^k\frac{1}{\sqrt{2}}\begin{pmatrix}1 & \beta\\ \alpha & \alpha\beta\end{pmatrix}, \tag{8.19}$$

where $\alpha, \beta \in \{1, i, -1, -i\}$ and $k = 0, \ldots, 7$.

What can we say about the invariant ring $\mathbb{C}[x,y]^G$? Using Molien's theorem, we can find the Hilbert series. A (slightly tedious) computation gives

$$H(\mathbb{C}[x,y]^G) = \frac{1}{(1-t^8)(1-t^{24})}. \tag{8.20}$$

This suggests that the invariant ring is generated by two algebraically independent polynomials f_1, f_2 homogeneous of degrees 8 and 24 respectively. This is indeed the case, just take $f_1 := x^8 + 14x^4y^4 + y^8$ and $f_2 := x^4y^4(x^4 - y^4)^4$. So the invariant ring is generated by f_1 and f_2, which implies the following powerful theorem on the weight enumerators of even self-dual codes.

Theorem 8.2.6 (Gleason)**.** *The weight enumerator of an even self-dual code is a polynomial in $x^8 + 14x^4y^4 + y^8$ and $x^4y^4(x^4 - y^4)$.*

Exercise 8.2.7. The Golay code is an even self-dual $[24, 12, 8]$-code. Use Theorem 8.2.6 to show that the weight enumerator of the Golay code equals

$$x^{24} + 759x^{16}y^8 + 2576x^{12}y^{12} + 759x^8y^{16} + y^{24}. \tag{8.21}$$

Exercise 8.2.8. There exists an even self-dual code $C \subseteq \mathbb{F}_2^{40}$, that contains no words of weight 4. How many words of weight 8 does C have?

Exercise 8.2.9. Let G be the group generated by the matrices $\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$. Use Molien's theorem to compute the Hilbert series of $\mathbb{C}[x,y]^G$ and find a set of algebraically independent generators.

Chapter 9

Algebraic groups

9.1 Definition and examples

Definition 9.1.1. A *linear algebraic group* is a Zariski closed subgroup of some GL_n.

In more down-to-earth terms, a linear algebraic group is a subgroup G of GL_n (with matrix product as group operation) which moreover is the zero set of some polynomial equations in the matrix entries. The *algebra of regular functions* $\mathbb{C}[G]$ on G is therefore $\mathbb{C}[x_{11}, x_{12}, \ldots, x_{nn}, \frac{1}{\det(x)}]/I(G)$ where $I(G)$ is the ideal of all regular functions on GL_n that vanish on G. We will usually drop the adjective *linear* and just say *algebraic group*. (Their theory, however, is very different from the theory of elliptic curves and other Abelian varieties, which are algebraic groups in other contexts.)

Example 9.1.2.

1. GL_n itself is a linear algebraic group, with zero defining ideal.
2. Often we will find it convenient not to specify a basis, and work with $\mathrm{GL}(V)$ (and its subgroups), where V is an n-dimensional complex vector space, rather than with GL_n. There is a basis-independent description of the algebra of regular functions on $\mathrm{GL}(V)$, as well: it is the algebra of functions on $\mathrm{GL}(V)$ generated by $\mathrm{End}(V)^*$ and $1/\det$.
3. $\mathrm{O}_n := \{g \in \mathrm{GL}_n \mid g^T g = 1\}$ is a linear algebraic group, called the *orthogonal group*. Its ideal turns out to be generated by the quadratic polynomials of the form $\sum_i x_{ij}x_{il} - \delta_{jl}$ for $j, l = 1, \ldots, n$ (it is obvious that these polynomials are contained in the ideal of O_n, but not that they generate a radical ideal).
4. The following is a basis-free version of the orthogonal group: let β be a non-degenerate symmetric bilinear form on an n-dimensional vector space V. Then $\mathrm{O}(\beta) := \{g \in \mathrm{GL}(V) \mid \beta(gv, gw) = \beta(v, w) \text{ for all } v, w \in V\}$ is isomorphic, in the sense defined below, to O_n.

5. $\mathrm{SL}_n := \{g \in \mathrm{GL}_n \mid \det(g) = 1\}$ is a linear algebraic group, whose ideal is generated by $\det(x) - 1$.

6. $\mathrm{G_a} := \left\{ \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \mid b \in \mathbb{C} \right\}$ is an algebraic group, called the *additive group*. Its ideal is generated by the polynomials $x_{11} - 1, x_{22} - 1, x_{21}$, and the quotient by this ideal is isomorphic to $\mathbb{C}[x_{12}]$.

7. Every finite group G "is" a linear algebraic group, because it can be realised as a finite (and hence Zariski-closed) subgroup of $\mathrm{GL}_{|G|}$ using the left regular representation.

8. Let A be a finite-dimensional $\mathbb{C}$-algebra. Then the set of all automorphisms of A is a linear algebraic subgroup of $\mathrm{GL}(A)$. Indeed, choosing a basis $1 = a_1, \ldots, a_n$ of A, the condition that a linear map $g \in \mathrm{GL}(A)$ is an automorphism of A is that $ga_1 = a_1$ and $(ga_i)(ga_j) = g(a_i a_j)$; this gives (at most) quadratic equations for the entries of the matrix of g relative to the basis $a_1, \ldots, a_n$.

9. The group of nonsingular diagonal $n \times n$-matrices is a linear algebraic group, called the n-dimensional *torus* T_n. The torus T_1 is also called the *multiplicative group* $\mathrm{G_m}$ or denoted GL_1.

The name torus stems from the following relation with the real n-dimensional torus.

Exercise 9.1.3. Let G be the subset of T_n of all matrices whose diagonal entries lie on the unit circle in $\mathbb{C}$. Show that G (an n-dimensional real torus) is a subgroup which is dense in T_n in the Zariski topology.

Exercise 9.1.4. Prove directly that the polynomials $\sum_j x_{ij}x_{kj} - \delta_{ik}$ with $i, k = 1, \ldots, n$ are contained in the ideal generated by the polynomials $\sum_i x_{ij}x_{il} - \delta_{il}$ with $j, l = 1, \ldots, k$.

Remark 9.1.5. The map $(\mathbb{C}, +) \to \mathrm{G_a}$ sending b to the matrix in the definition of $\mathrm{G_a}$ is an isomorphism of abstract groups, which moreover is given by a regular map whose inverse is also regular. We will therefore also consider $(\mathbb{C}, +)$ an algebraic group. This is consistent with a more abstract definition of (affine) algebraic groups as affine varieties with a compatible group structure.

Exercise 9.1.6. Prove that $\det(x) - 1$ is an irreducible polynomial in the entries of x. (This shows that the ideal of SL_n is, indeed, generated by $\det(x) - 1$.)

Definition 9.1.7. Let $G \subseteq \mathrm{GL}_n$ and $H \subseteq \mathrm{GL}_m$ be algebraic groups. An *algebraic group homomorphism* from G to H is a group homomorphism $G \to H$ which moreover is a regular map.

The latter condition means, very explicitly, that the homomorphism is given by m^2 functions $\phi_{ij} : \mathrm{G} \to \mathrm{GL}_m$ which are restrictions to G of polynomials in the matrix entries x_{ij} of GL_n and $1/\det(x)$.

Example 9.1.8. Consider the map $\phi : \mathrm{GL}_n \to \mathrm{GL}(M_n)$ that sends g to the linear map $\phi(g) : A \mapsto gAg^{-1}$. This is an algebraic group homomorphism. Indeed, it is clearly a group homomorphism, so we need only verify that it is a regular map. This means that the $n^2 \times n^2$-matrix of $\phi(g)$ relative to the basis of elementary matrices E_{ij} (with zeroes everywhere but for a 1 on position (i,j)) of M_n depends polynomially on the entries and the inverse of the determinant of g. But this follows from

$$(gE_{pq}g^{-1})_{ij} = \prod_k \prod_l g_{ik}\delta_{kp}\delta_{ql}(g^{-1})_{lj}$$

and Cramer's rule that expresses the entries of g^{-1} in those of g.

We claim that the image of ϕ is itself an algebraic group, that is, that it is a Zariski-closed subgroup of $\mathrm{GL}(M_n)$. Note that it is certainly a group, isomorphic to $\mathrm{GL}_n/\ker\phi$, where $\ker\phi$ consists of the scalar matrices; this group is called the *projective (general) linear group* and denoted PGL_n. Note also that $\operatorname{im}\phi$ is contained in the group $\mathrm{Aut}(M_n)$ of automorphisms of the n^2-dimensional algebra M_n. We prove that $\operatorname{im}\phi = \mathrm{Aut}(M_n)$. For this, note that the matrices $E_{ij},\ i,j = 1,\dots,n$ satisfy

$$\sum_i E_{ii} = I \text{ and } E_{ij}E_{kl} = \delta_{jk}E_{il}.$$

Now if α is any automorphism of M_n, then the matrices $F_{ij} := \alpha(E_{ij}),\ i,j = 1,\dots,n$ will satisfy the same relations. Let V_i be the image of F_{ii}. Then $V_i \neq 0$ since $F_{ii} \neq 0$, and the relations $\sum_i F_{ii} = I$ and $F_{ii}F_{ii} = F_{ii}$ and $F_{ii}F_{jj} = 0,\ i \neq j$ imply that

$$\mathbb{C}^n = \bigoplus_{i=1}^n V_i.$$

Hence each V_i must be one-dimensional. Let v_1 be a basis of V_1, and set $v_i := F_{i1}v_1,\ i = 2,\dots,n$. Then $v_i \neq 0$ since $0 \neq F_{i1} = F_{i1}F_{11}$, and v_i spans V_i since $F_{ii}v_i = F_{ii}F_{i1}v_1 = F_{i1}v_1 = v_i$. Now there is a unique $g \in \mathrm{GL}_n$ such that $ge_i = v_i$, and we claim that $\phi(g) = \alpha$. Indeed, on the one hand we have $gE_{ij}e_k = \delta_{jk}v_i$, and on the other hand we have

$$F_{ij}ge_k = F_{ij}v_k = F_{ij}F_{k1}v_1 = \delta_{jk}F_{i1}v_1 = \delta_{jk}v_i,$$

which shows that $gE_{ij}g^{-1} = F_{ij}$ for all $i,j = 1,\dots,n$, so that $\alpha = \phi(g)$, as claimed.

Remark 9.1.9. In fact, the image of *any* homomorphism of algebraic groups is Zariski-closed, and hence an algebraic group itself. This is a fundamental property of algebraic groups, which for instance Lie groups do not share. To prove this property, however, one needs slightly more algebraic geometry than we have at our avail here.

Definition 9.1.10. An *algebraic group isomorphism* is an algebraic group homomorphism from G to H that has an inverse which is also an algebraic group homomorphism.

Exercise 9.1.11. Determine all algebraic group automorphisms of $\mathrm{G_a}$ and of T_n.

Remark 9.1.12. If there exists an algebraic group isomorphism between algebraic groups $G \subseteq \mathrm{GL}_n$ and $H \subseteq \mathrm{GL}_m$, then one can use this isomorphism to prove that $\mathbb{C}[G] = \mathbb{C}[\mathrm{GL}_n]/I(G)$ and $\mathbb{C}[H] = \mathbb{C}[\mathrm{GL}_m]/I(H)$ are isomorphic, as well. Just like with affine varieties discussed earlier, this allows one to think of an algebraic group abstractly, without thinking of one particular closed embedding into some matrix group. We will sometimes implicitly adopt this more abstract point of view.

Definition 9.1.13. A *finite-dimensional rational representation* of an algebraic group G is an algebraic group homomorphism $\rho : G \to \mathrm{GL}(V)$ for some finite-dimensional vector space V, which is then called a *rational G-module*.

A *locally finite rational representation* of G is a group homomorphism ρ from G into the group $\mathrm{GL}(V)$ of bijective linear maps from some potentially infinite vector space V into itself with the following additional property: for each $v \in V$ there is a finite-dimensional subspace U of V containing v which is $\rho(G)$-stable and for which the induced homomorphism $\rho : G \to \mathrm{GL}(U)$ is an algebraic group homomorphism, that is, regular.

In both cases we will write gv instead of $\rho(g)v$ when ρ is clear from the context. We will also use the word *module* for the space V equipped with its linear G-action.

Note that we insist that a rational representation ρ be defined by regular functions defined everywhere on G. We nevertheless follow the tradition of using the adjective *rational*, which refers to the fact that the defining functions may involve $1/\det(x)$ in addition to the matrix entries x_{ij}.

Example 9.1.14.

1. The identity map $\mathrm{GL}_n \to \mathrm{GL}_n = \mathrm{GL}(\mathbb{C}^n)$ makes $V = \mathbb{C}^n$ into a rational GL_n-module. The second tensor power $V^{\otimes 2}$ is also a rational GL_n-module. To verify this we note that the matrix entries of $g^{\otimes 2}$ relative to the basis $e_j \otimes e_l,\ j, l = 1, \ldots, n$ of $V^{\otimes 2}$, where the e_i are the standard basis of $\mathbb{C}^n$, depend polynomially (in fact, quadratically) on the matrix entries of g: $(g^{\otimes 2})(e_j \otimes e_l) = (\sum_i g_{ij} e_i) \otimes (\sum_k g_{kl} e_k) = \sum_{i,k} g_{ij} g_{kl} e_i \otimes e_k$.

2. For any integer k the map $\mathrm{GL}_n \to \mathrm{G_m},\ g \mapsto \det(g)^k$ is a one-dimensional rational representation. It restricts to a rational representation of any Zariski closed subgroup G of GL_n. For $k = 0$ this one-dimensional representation is called the *trivial representation* of G.

3. The group homomorphism $\mathrm{G_a} \to \mathrm{GL}_1$ sending $\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$ to $\exp(b)$ is *not* given by regular functions, hence *not* a rational representation. This reflects a difference between the theories of algebraic groups and of Lie groups.

4. For any algebraic group G given as a Zariski closed subgroup of $\mathrm{GL}(V)$ the space V is a rational G-module, which we will the *defining module* of G.

Exercise 9.1.15. Determine all the one-dimensional rational representations of the torus T_n. Hint: $\mathbb{C}[\mathrm{T}_n] \cong \mathbb{C}[x_{11}^{\pm 1}, \ldots, x_{nn}^{\pm 1}]$, and a one-dimensional rational representation $\rho : \mathrm{T}_n \to \mathrm{GL}_1 = \mathrm{T}_1$ is an element ρ of $\mathbb{C}[\mathrm{T}_n]$ that, apart from being multiplicative, does not vanish anywhere on T_n.

Exercise 9.1.16. Show that $\mathrm{G_a}$ does not have any non-trivial one-dimensional rational representations. Show also that the $\mathrm{G_a}$-submodule $\mathbb{C}e_1$ in the defining module $\mathbb{C}^2$ does not have a $\mathrm{G_a}$-stable vector space complement in $\mathbb{C}^2$.

Remark 9.1.17. Like the second exterior power above, (multi-)linear algebra constructions transform locally finite rational G-modules into others. In particular, if $\rho : G \to \mathrm{GL}(V)$ is a locally finite rational representation, then

1. if U is a $\rho(G)$-stable subspace of V, then the induced maps $G \to \mathrm{GL}(U)$ and $G \to \mathrm{GL}(V/U)$ are locally finite rational representations;

2. if W is a second locally finite rational G-module, then $V \oplus W$ and $V \otimes W$ are also locally finite rational G-modules;

3. if k is a natural number, then $S^k V$ is a locally finite rational G-module—indeed, it is a quotient of $V^{\otimes k}$, and hence locally finite and rational by the previous two constructions; etc.

We will never consider other representations of algebraic groups than locally finite rational ones. We will often drop these adjectives.

9.2 The algebra of regular functions as a representation

To any algebraic group $G \subseteq \mathrm{GL}_n$ we have associated its algebra $\mathbb{C}[G]$ of regular functions. Consider the map $\lambda : G \to \mathrm{GL}[\mathbb{C}[x_{ij}, 1/\det(x)])$ defined by

$$(\lambda(g)f)(h) := f(g^{-1}h) \text{ for all } g \in G,\ h \in \mathrm{GL}_n.$$

As f is a polynomial function in the x_{ij} and $1/\det(x)$, the expression $f(g^{-1}h)$ is a polynomial in the matrix entries and the inverse determinant of $g^{-1}h$ and hence in the matrix entries $g_{ij}, h_{ij},\ i, j = 1, \ldots, n$ and the inverse determinants $\det(g)^{-1}, \det(h)^{-1}$. In particular, for fixed g, the function $\lambda(g)f$ is a regular function on GL_n.

The map λ satisfies $\lambda(1)f = f$ and $\lambda(g_1g_2)f = \lambda(g_1)\lambda(g_2)f$ and $\lambda(g)(f_1 + f_2) = \lambda(g)f_1 + \lambda(g)f_2$ and $\lambda(g)(cf) = c\lambda(g)f$ and $\lambda(g)(f_1f_2) = (\lambda(g)f_1)(\lambda(g)f_2)$. In other words, λ furnishes a representation of G by means of automorphisms on $\mathbb{C}[\mathrm{GL}_n]$. We claim that it is locally finite and rational. To see this let $f \in \mathbb{C}[\mathrm{GL}_n]$. Then the expression $f(g^{-1}h)$ can be expanded as a finite sum

$\sum_i f_i(g)f_i'(h)$, where the f_i, f_i' are polynomial functions in the entries and inverse determinant of g, h, respectively. This shows that for all $g \in G$ the function $\lambda(g)f$ lies in the finite-dimensional subspace of $\mathbb{C}[\mathrm{GL}_n]$ spanned by the f_i'. Now let $U \subseteq \mathbb{C}[\mathrm{GL}_n]$ be the linear span of all $\lambda(g)f,\ g \in G$. This is a G-stable space, and finite-dimensional since it is contained in in the span of the f_i'. This proves that $\mathbb{C}[\mathrm{GL}_n]$ is a locally finite G-module. Finally, let $g_1, \ldots, g_k$ be such that the $f_j'' := \lambda(g_j)f$ form a basis of U. Choose any projection π from the span of the f_i' onto U. Then we have, for all $g \in G$,

$$\lambda(g)f_j'' = \pi(\sum_i f_i(gg_j)f_i').$$

The right-hand side is a polynomial expression in g and its inverse determinant, hence $G \to \mathrm{GL}(U)$ is a regular map, so that $\mathbb{C}[\mathrm{GL}_n]$ is a rational G-module, as claimed.

Exercise 9.2.1. Show that the ideal $I(G) \subseteq \mathrm{GL}_n$ of the algebraic group G is stable under $\lambda(G)$.

As a consequence of the above, and of the exercise, the algebra $\mathbb{C}[G] = \mathbb{C}[\mathrm{GL}_n]/I(G)$ of regular functions on G is a locally finite, rational G-module.

Remark 9.2.2. The above concerns the action of G by left translation on GL_n and on itself. A similar construction can, of course, be carried out for its action by right translation.

Proposition 9.2.3. *Let V be any finite-dimensional rational module for the algebraic group G whose dual V^* is generated, as a G-module, by a single element. Then there exists a G-equivariant embedding of V into $\mathbb{C}[G]$.*

Proof. Suppose that f generates V^*. Take the map $\psi : V \to \mathbb{C}[G],\ v \mapsto f_v$, where $f_v \in \mathbb{C}[G]$ is defined by $f_v(h) = f(h^{-1}v),\ h \in G$—this is a regular map since V is a rational G-module. The map ψ is G-equivariant since

$$f_{gv}(h) = f(h^{-1}gv) = f_v(g^{-1}h) = (gf_v)(h)$$

for all $g, h \in G$ and $v \in V$. Moreover, ψ is linear and has trivial kernel—indeed, $f_v = 0$ means that $(hf)(v) = f(h^{-1}v) = f_v(h) = 0$ for all $h \in G$ so that, since the G-orbit of f spans V^*, v must be zero. □

Exercise 9.2.4.

1. Use the argument of the proof to show that every finite-dimensional rational G-module can be embedded, as a G-module, into a direct sum of finitely many copies of $\mathbb{C}[G]$.
2. Prove that every finite-dimensional rational T_n-module is a direct sum of one-dimensional T_n-modules. Hint: analyse the T_n-module $\mathbb{C}[\mathrm{T}_n]$.

We already knew that finite-dimensional representations of *finite* groups are completely reducible, and the preceding exercise shows that the same is true for finite-dimensiona rational representations of T_n. Next week we will prove that this is true for a larger class of groups.

Exercise 9.2.5. Let A be the algebra of (complexified) quaternions, that is, $A = \mathbb{C}I \oplus \mathbb{C}J \oplus \mathbb{C}K \oplus \mathbb{C}L$ where I is the identity element and the (associative) multiplication is determined by $J^2 = K^2 = L^2 = -I$ and $JK = L$. Prove that the automorphism group of A is isomorphic, as an algebraic group, to the subgroup SO_3 of O_3 consisting of the matrices with determinant 1.

Chapter 10

Reductiveness

We want to extend techniques and theory that we first derived for finite groups—such as the Reynolds operator and Hilbert's finiteness results—to algebraic groups. In general this is not so easy, but for the class of *linearly reductive groups* the theory is very rich and deep.

Recall the following definition from the lecture on representations.

Definition 10.0.6. A locally finite representation of a group G is called *completely reducible* if it is the direct sum of finite dimensional irreducible representations of G.

Exercise 10.0.7. Show that a locally finite G-module V is completely reducible if every finite dimensional submodule of V is completely reducible.

Definition 10.0.8. An algebraic group G is called *linearly reductive* if each finite-dimensional rational representation of G is completely reducible.

It turns out that linearly reductive groups G can be characterised in a more intrinsic manner, namely, that the *unipotent radical* of G is trivial. Groups with this property can been classified by combinatorial data involving objects such as *root systems* and *lattices*. However, we are not going to do so. In this course it will suffice to have a collection of examples of linearly reductive groups, such as the classical groups $\mathrm{GL}_n, (\mathrm{S})\mathrm{O}_n, \mathrm{SL}_n, \mathrm{Sp}_{2n}$ and the tori T_n, in addition to *finite* groups, which are always linearly reductive.

In the case of finite groups linear reductiveness follows from the fact that the group leaves invariant a Hermitian inner product on any finite-dimensional representation. This is not true for the classical groups.

Exercise 10.0.9. Show that there is no Hermitian inner product on $\mathbb{C}^n$ that satisfies $(gv|gw) = (v|w)$ for all $g \in \mathrm{O}_n$ (we take non-degenerateness as part of the definition of a Hermitian inner product).

However, a weaker statement is true, which still suffices to prove linear reductiveness of the classical groups.

Theorem 10.0.10. *Suppose that $G \subseteq \mathrm{GL}_n$ has the property that with each element g also the Hermitian conjugate $g^* = \overline{g^T}$ is in G. Then G is linearly reductive.*

For the proof it will be convenient to give a name to the property that the defining representation in this theorem has. Given a group G and a map $g \mapsto g^*$ from G into itself, a $*$-*representation* is by definition a finite-dimensional representation $\rho : G \to \mathrm{GL}(V)$ together with a Hermitian inner product $(.|.)$ on V such that for each $g \in G$ we have $\rho(g^*) = \rho(g)^*$, where the latter map is the Hermitian conjugate determined by $(\rho(g)v|w) = (v|\rho(g)^*)$ for all $v, w \in V$. We also call V a $*$-*module* for G.

From the $*$-module V for G we can construct other $*$-modules. First, if U is a G-submodule of V, then $U^\perp$ is also a G-submodule of V. Indeed, for $g \in G$ and $v \in U^\perp$ we have

$$(\rho(g)v|U) = (v|\rho(g)^*U) = (v|\rho(g^*)U) = (v|U) = \{0\}). \qquad (10.1)$$

Now both U and $U^\perp$ inherit Hermitian inner products from V, and one readily checks that $(\rho(g)|_U)^* = \rho(g)^*|_U = \rho(g^*)|_U$. This shows that U is a $*$-module for G, and so is $U^\perp$. Furthermore, the natural map $U^\perp \to V/U$ is an isomorphism of G-modules, by means of which the Hermitian inner product on $U^\perp$ can be transferred to a Hermitian inner product on V/U making the latter into a $*$-module for G.

Second, we have an inner product on $\mathrm{End}(V)$ defined by $(A|B) = \mathrm{tr}(AB^*)$, for $A, B \in \mathrm{End}(V)$. Now for $g \in G$ and all $A, B \in \mathrm{End}(V)$ we have

$$\begin{aligned}(\rho(g)A\rho(g)^{-1}|B) &= \mathrm{tr}(\rho(g)A\rho(g)^{-1}B^*) \\ &= \mathrm{tr}(A(\rho(g)^*B(\rho(g)^*)^{-1})^*) = (A|\rho(g^*)B\rho((g^*)^{-1})).\end{aligned} \qquad (10.2)$$

This shows that $\mathrm{End}(V)$ is a $*$-module with the action of G by conjugation on it.

Third, if U is another $*$-module for G, then $U \oplus V$ is a $*$-module for G with respect to the direct sum inner product; and $U \otimes V$ is a $*$-module for G with respect to the inner product determined by $(u \otimes v|u' \otimes v') = (u|u') \cdot (v|v'),\ u, u' \in U,\ v, v' \in V$.

Exercise 10.0.11. Let V be a $*$-module for G with respect to the inner product $(\cdot|\cdot)$ on V.

- Let $\phi : V \to V^*$ be the antilinear map given by $\phi(v) = (u \mapsto (u|v))$. Check that the G-module V^* with the inner product $(\phi(v)|\phi(w)) := (w|v)$ is a $*$-module for G.

- Consider the $*$-module $\mathrm{End}(V)$. By the previous item, the dual module $\mathrm{End}(V)^*$ is also a $*$-module for G. Check that the inner product and G-action on $\mathrm{End}(V)^*$ coincide with those induced by the natural linear bijection $\mathrm{End}(V) \to \mathrm{End}(V)^*$ given by $A \mapsto (B \mapsto \mathrm{tr}(AB))$.

Finally, a symmetric power $S^k(\mathrm{End}(V)^*) \cong S^k\mathrm{End}(V)$ is the quotient of the tensor power $\mathrm{End}(V)^{\otimes k}$ of V by some G-submodules, and hence, by the constructions above, a $*$-module for G. This shows that every finite-dimensional G-submodule of $\mathbb{C}[\mathrm{End}(V)] = \mathbb{C}[(x_{ij})_{ij}]$ is a $*$-module for G. Now we are in a position to prove the theorem.

Proof of Theorem 10.0.10. Let $V := \mathbb{C}^n$ be the standard G-module. By assumption it is a $*$-module for G with respect to the standard Hermitian inner product $(v|w) = \sum_i v_i\overline{w_i}$ on $\mathbb{C}^n$ and the map $g \mapsto g^*$ that sends a matrix to its standard Hermitian conjugate. Hence, by the above construction, all finite-dimensional G-submodules of $\mathbb{C}[\mathrm{End}(V)]$ are $*$-modules for G. In particular, every finite-dimensional G-submodule of $\mathbb{C}[\mathrm{End}(V)] = \mathbb{C}[(x_{ij})_{ij}]$ is completely reducible—this follows by induction from the argument with U and $U^\perp$ above. But then (every finite-dimensional G-submodule of) the slightly larger algebra $\mathbb{C}[\mathrm{GL}(V)] = \mathbb{C}[(x_{ij})_{ij}, 1/\det(x)]$ is also completely reducible; see the exercise below. As a consequence, also the quotient $\mathbb{C}[\mathrm{GL}(V)]/I(G) = \mathbb{C}[G]$ and direct sums of finitely many copies of $\mathbb{C}[G]$ are completely reducible into finite-dimensional irreducible rational G-modules. By Exercise 9.2.4 *every* finite-dimensional rational G-module W is a submodule of some direct sum of copies of $\mathbb{C}[G]$, so W is completely reducible, as well. □

Exercise 10.0.12. Let G be a closed subgroup of $\mathrm{GL}(V)$ and suppose that W is a finite-dimensional G-submodule of $\mathbb{C}[\mathrm{GL}(V)]$. Show that for the integer k larger enough, $W' := \det(x)^k W \subseteq \mathbb{C}[(x_{ij})_{ij}]$, and that W is completely reducible if and only if W' is.

As an application of the ideas presented above, we classify the (finite dimensional) irreducible rational representations of $\mathrm{SL}_2(\mathbb{C})$.

Proposition 10.0.13. *Let $V = \mathbb{C}^2$ be the standard $\mathrm{SL}_2(\mathbb{C})$-module. Then for each nonnegative integer k, the rational $\mathrm{SL}_2(\mathbb{C})$-module $S^k(V)$ is irreducible and every irreducible rational $\mathrm{SL}_2(\mathbb{C})$-module is isomorphic to exactly one $S^k(V)$.*

Proof. We will first show that the $S^k(V)$ are irreducible. Denote by x, y the standard basis $\mathbb{C}^2$ and by $(x^iy^{k-i})_{i=0}^k$ the induced basis of $S^d(V)$. Hence we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} x^iy^{k-i} = (ax+cy)^i(bx+dy)^{k-i}. \tag{10.3}$$

Let $U \subseteq S^k(V)$ be an irreducible submodule. Take $u = \sum_{i=0}^k c_ix^iy^{k-i} \in U$, with maximal support (i.e. the coefficients c_i that are nonzero). We claim that *all* coefficients c_i are nonzero.

Indeed, $c_0 \neq 0$, since otherwise we can replace u by

$$\begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} u = \sum_i c_i(x+\lambda y)^iy^{k-i} = \sum_i d(\lambda)x^iy^{k-i}.$$

Since d_0 and every d_i with $c_i \neq 0$ are nonzero polynomials in λ, we can take any λ not a root of these nonzero polynomials and obtain a vector in U with larger support. Now since $c_0 \neq 0$, the vector $\left(\begin{smallmatrix} 1 & \lambda \\ 0 & 1 \end{smallmatrix}\right) u = \sum_i c_i x^i (y + \lambda x)^{k-i}$ has full support for all but a finite number of λ.

Next, take $\mu_0, \ldots, \mu_k \in \mathbb{C}$ such that the numbers $\lambda_i := \mu_i^2$ are distinct and nonzero. Then for every i the vector $\mu^k \begin{pmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{pmatrix} u = \sum_i \lambda^i c_i x^i y^{k-i}$ belongs to U and these $k+1$ vectors are linearly independent because the Vandermonde matrix $(\lambda_i^j)_{i,j=0}^k$ has nonzero determinant. This shows that $U = S^k(V)$.

To prove that we have found *all* irreducible rational $\mathrm{SL}_2(\mathbb{C})$-modules, it suffices by Exercise 9.2.4 to show that $\mathbb{C}[\mathrm{SL}_2(\mathbb{C})] = \mathbb{C}[\mathrm{End}(V)]/(det - 1)$ decomposes into copies of the $S^k(V)$. Since $\mathbb{C}[\mathrm{End}(V)] = \bigoplus_d S^d(\mathrm{End}(V))$ is completely reducible, it suffices to consider each $S^d(\mathrm{End}(V))$. Because $\mathrm{End}(\mathbb{C}^2) \cong S^2(\mathbb{C}^2) \oplus \mathbb{C}$, we have

$$S^d(\mathrm{End}(V)) = S^d(\mathrm{End}(S^2(V) \oplus \mathbb{C})) = \bigoplus_{i=0}^{d} S^i(S^2(V)). \tag{10.4}$$

By exercise 10.0.14, we can decompose each $S^i(S^2(V))$ as required, finishing the proof. $\square$

Exercise 10.0.14. Let $V = \mathbb{C}^2$ be the standard $\mathrm{SL}_2(\mathbb{C})$ module. Show that for every positive integer d, we have the following decomposition of the $\mathrm{SL}_2(\mathbb{C})$ module $S^d(S^2(V))$:

$$S^d(S^2(V)) \cong \bigoplus_{k=0}^{\lfloor d/2 \rfloor} S^{2d-4k}(V). \tag{10.5}$$

Hint: Let x, y be a basis of V and $X = x^2, Y = y^2, Z = xy$ a basis of $S^2(V)$. Show that the kernel of the homomorphism $\phi : S^d(S^2(V)) \to S^{2d}(V)$ defined by $\phi(X^a Y^b Z^c) := (x^2)^a (y^2)^b (xy)^c$, is isomorphic to $S^{d-2}(S^2(V))$.

Chapter 11

First Fundamental Theorems

11.1 Schur-Weyl Duality

Let $V := \mathbb{C}^n$, $G := \mathrm{GL}(V)$, and $\rho : G \to \mathrm{GL}(V^{\otimes k})$ the diagonal representation of G. The latter space also furnishes a representation $\lambda : S_k \to \mathrm{GL}(V^{\otimes k})$ of the symmetric group, given by

$$\lambda(\pi)(v_1 \otimes \cdots \otimes v_k) = v_{\pi^{-1}(1)} \otimes \cdots \otimes v_{\pi^{-1}(k)},\ \pi \in S_k,\ v_1, \ldots, v_k \in V.$$

Clearly $\rho(g)$ and $\lambda(\pi)$ commute for all $g \in G$ and $\pi \in S_k$. Let R be the associative algebra generated by all $\rho(g)$, $g \in G$ and let S be the associative algebra generated by all $\lambda(\pi)$, $\pi \in S_k$. Then every element of R commutes with every element of S. The following theorem, due to Schur, shows that this also *characterises* R in terms of S and vice versa.

Theorem 11.1.1 (Schur-Weyl Duality)**.** *We have*

$$R = \{a \in \mathrm{End}(V^{\otimes k}) \mid as = sa \textit{ for all } s \in S\}$$

and

$$S = \{a \in \mathrm{End}(V^{\otimes k}) \mid ar = ra \textit{ for all } r \in R\}.$$

Proof. We first prove the inclusion

$$R \supseteq \{a \in \mathrm{End}(V^{\otimes k}) \mid as = sa \text{ for all } s \in S\} \tag{11.1}$$

(the other inclusion being obvious), and then deduce the second statement from the *Double Commutant Theorem.* Note that the right-hand side is just $\mathrm{End}(V^{\otimes k})^{S_k}$, where S_k acts by conjugation via λ on $\mathrm{End}(V^{\otimes k})$. The map

$$\mathrm{End}(V)^k \to \mathrm{End}(V^{\otimes k}),\ (a_1, \ldots, a_k) \mapsto a_1 \otimes \cdots \otimes a_k$$

is k-linear and hence induces a linear map

$$\mathrm{End}(V)^{\otimes k} \to \mathrm{End}(V^{\otimes k}).$$

It is not hard to check that this latter map is bijective and S_k-equivariant. Hence it identifies the space $\mathrm{End}(V^{\otimes k})^{S_k}$ with $(\mathrm{End}(V)^{\otimes k})^{S_k}$ of symmetric tensors in $E^{\otimes k}$ where $E := \mathrm{End}(V)$.

On the other hand, the image $\rho(g)$ of $g \in G$ in $E^{\otimes k}$ is just $g^{\otimes k}$, and the left-hand side, R, of (11.1) contains all these elements. Since R is closed and G is dense in E and taking k-th tensor powers is continuous, R contains all tensors of the form $a^{\otimes k}$ with $a \in E$. Hence we have to prove that the symmetric tensors in $E^{\otimes k}$ are spanned by the tensors of the form $a^{\otimes k}$.

For this we forget that E is an algebra of linear maps, and claim that for any finite-dimensional vector space E the subspace of $E^{\otimes k}$ consisting of symmetric tensors is spanned by all elements of the form $a^{\otimes k} := a \otimes \cdots \otimes a$. We prove this indirectly: suppose that f is a linear function on the space of symmetric tensors in $E^{\otimes k}$ that vanishes identically on $a^{\otimes k}$ for all $a \in E$. Expanding a as $\sum_{i=1}^m t_i a_i$, where $a_1, \ldots, a_m$ is a basis of E, we find that

$$\sum_{\alpha \in \mathbb{N}^m,\ \sum_i \alpha_i = k} t^\alpha f\Big(\sum_{i_1,\ldots,i_k} a_{i_1} \otimes \cdots \otimes a_{i_k}\Big) = 0,$$

where the second sum is over all k-tuples $i_1, \ldots, i_k \in [m]$ in which exactly α_i copies of i occur for all i, and where the vanishing is identical in $t_1, \ldots, t_m$. But then f must vanish on each such sum $\sum_{i_1,\ldots,i_k} a_{i_1} \otimes \cdots \otimes a_{i_k}$, and as α runs through all multi-indices with $\sum_i \alpha_i = k$ these sums run through a basis of the space of symmetric tensors in $E^{\otimes k}$. This proves the claim, and hence the first statement. Applying the following theorem to $H = S_n$ and R and S as above, we obtain the second statement. □

Theorem 11.1.2 (Double Commutant Theorem)**.** *Let H be a group and let $\lambda : H \to \mathrm{GL}(W)$ be a completely reducible representation of H. Let S be the linear span of $\lambda(H)$; this is a subalgebra of* $\mathrm{End}(W)$*. Set*

$$R := \{a \in \mathrm{End}(W) \mid as = sa \text{ for all } s \in S\}.$$

Then, conversely, we have

$$S = \{a \in \mathrm{End}(W) \mid ar = ra \text{ for all } r \in R\}.$$

Proof. Let $w_1, \ldots, w_p$ be a basis of W. Since W is a completely reducible H-module, so is W^p, the Cartesian product of p copies of W. Let M be the submodule of W^p generated by $(w_1, \ldots, w_p)$; it consists of all elements of the form $(sw_1, \ldots, sw_p)$ where s runs through S. As W^p is completely reducible, M has a direct complement U that is stable under H, and the map $\phi : W^p \to W^p,\ m + u \mapsto m,\ m \in M, u \in U$ is H-equivariant. This means that each component ϕ_{ij}, obtained by precomposing ϕ with the inclusion of the j-th copy of W and postcomposing with the projection onto the i-th copy of W, is an element of R.

Now let $a \in \mathrm{End}(W)$ be an element that commutes with all elements of R. Then, since all ϕ_{ij} are elements of R, the p-tuple $(a, \ldots, a) : W^p \to W^p$ commutes with ϕ. Hence

$$\phi(aw_1, \ldots, aw_p) = (a, \ldots, a)\phi(w_1, \ldots, w_p) = (aw_1, \ldots, aw_p)$$

where the last equality follows from $(w_1, \ldots, w_p) \in M$, so that ϕ leaves it invariant. The left-hand side is an element of $\operatorname{im} \phi = M$, hence equal to $(sw_1, \ldots, sw_p)$ for some $s \in S$. As s and a agree on the basis $w_1, \ldots, w_p$, we have $s = a$. □

11.2 The First Fundamental Theorem for GL_n

One of the fundamental contributions of German mathematician Hermann Weyl to the field of invariant theory concerns the invariants of GL_n on "p covectors and q vectors". To set the stage, set $G := \mathrm{GL}_n$ and let $V = \mathbb{C}^n$ denote the standard G-module. Consider the G-module $U := (V^*)^p \oplus V^q$, i.e., the direct sum of p copies of V^* and q copies of the dual V. We want to determine $\mathbb{C}[U]^G$, that is, the algebra of G-invariant regular functions on U. Now for every $i = 1, \ldots, p$ and $j = 1, \ldots, q$ the function

$$\phi_{ij} : U \to \mathbb{C},\ (f_1, \ldots, f_p, v_1, \ldots, v_q) \mapsto f_i(v_j)$$

is G-invariant by definition of the G-action on V^*: $(gf_i)(gv_j) = f_i(v_j)$.

Theorem 11.2.1 (Weyl's First Fundamental Theorem (FFT) for GL_n)**.** *The algebra $\mathbb{C}[U]^G$ is generated by the functions ϕ_{ij}.*

Before proving this theorem we will give a geometric interpretation. In terms of matrices U is isomorphic to $M_{p,n} \times M_{n,q}$, where the action of $g \in G$ is given by $g(A, B) = (Ag^{-1}, gB)$—the rows of A are the p *covectors* and the columns of B the q vectors. Clearly the *product* $A \cdot B$ equals the product $Ag^{-1} \cdot gB$, and therefore the compositions of the $p \cdot q$ entry functions on $M_{p,q}$ with the product map $M_{p,n} \times M_{n,q} \to M_{p,q}$ are invariant regular functions. The theorem implies that they generate the algebra of invariants on $M_{p,n} \times M_{n,q}$, so that *the product map $M_{p,n} \times M_{n,q} \to M_{p,q}$ is a quotient map in the sense of Chapter 6.*

Proof. The theorem is a relatively easy consequence of Schur-Weyl duality and the complete reducibility of rational represations of GL_n. However, you need a "tensor flexible mind" to follow the proof; here it goes. We want to determine the G-invariant homogeneous polynomials of degree d on U. Those are the elements of $(S^d U^*)^G$. By complete reducibility, the map $((U^*)^{\otimes d})^G \to (S^d U^*)^G$ is surjective, so we concentrate on the invariant tensors in $(U^*)^{\otimes d} = (V^p \oplus (V^*)^q)^{\otimes d}$. Denote the p copies of V by $U_1, \ldots, U_p$ and the q copies of V^* by $U_{1'}, \ldots, U_{q'}$. The tensor product $(V^p \oplus (V^*)^q)^{\otimes d}$ decomposes, as a G-module, into the direct sum of tensor products $U_{i_1} \otimes \cdots \otimes U_{i_d}$ over all d-tuples $(i_1, \ldots, i_d) \in$

$\{1,\ldots,p,1',\ldots,q'\}$. Each such product is isomorphic to $V^{\otimes k}\otimes(V^*)^{\otimes d-k}$ for some k. Now by Exercise 11.2.2 below the space of G-invariants in this space is zero if $k\neq d-k$. If $d=k$, on the other hand, then this space is just $\mathrm{End}(V)^{\otimes k}$, and Schur-Weyl duality tells us that the space of G-invariant elements is spanned by the image of the symmetric group.

Finally, we need to reinterpret these invariant tensors in terms of the ϕ_{ij}. This is best done by an example: suppose that $p=2,q=3,d=4,i_1=i_2=1$ and $i_3=1'$ and $i_4=3'$. Then $U_1\otimes U_1\otimes U_{1'}\otimes U_{3'}=V\otimes V\otimes V^*\otimes V^*$. Now the elements of S_2 correspond to the two ways of pairing the copies of V with the copies of V^*, and both invariant tensors project to the polynomial function $\phi_{11}\phi_{13}$. On the other hand, if we take $i_1=1,\ i_2=2$, and leave the remaining values unchanged, then we obtain exactly the two polynomial functions $\phi_{11}\phi_{23}$ and $\phi_{13}\phi_{21}$. □

Exercise 11.2.2. Show that $V^{\otimes k}\otimes(V^*)^{\otimes l}$ has no non-zero G-invariant elements unless $k=l$.

Exercise 11.2.3. By the proof above, the space $[V^{\otimes k}\otimes(V^*)^{\otimes k}]^{\mathrm{GL}(V)}$ has dimension at most $k!$. Show that for k fixed and n sufficiently large this dimension is, indeed, $k!$.

Remark 11.2.4. Weyl's *Second Fundamental Theorem* for GL_n describes the ideals of relations among the functions ϕ_{ij}. The image of the quotient map consists of all $p\times q$-matrices of the form AB with $A\in M_{p,n}$ and $B\in M_{n,q}$. If $n\geq\min\{p,q\}$, then all $p\times q$-matrices are of this form, and the ϕ_{ij} are algebraically independent. If $n<p,q$, then only the matrices of rank at most n are of the form AB. In this latter cases the $n\times n$-subdeterminants (or *minors*) of the matrix $(\phi_{ij})_{ij}$ are elements of the ideal. Weyl's theorem says that these minors generate the ideal.

11.3 First Fundamental Theorem for $\mathrm{O_n}$

For applications to graph invariants we will need an analogue of Theorem 11.2.1 for the orthogonal group. So consider the defining O_n-module $V=\mathbb{C}^n$, and let $(.|.)$ be the standard symmetric bilinear form on V defined by $(v|w)=\sum_{i=1}^n v_iw_i$. By definition we have $(gv|gw)=(v|w)$ for all $v,w\in V$ and $g\in G:=\mathrm{O}_n$. Now it is unnecessary to distinguish between vectors and co-vectors, because the map $v\mapsto(.|v)$ is an isomorphism of G-modules from V to V^*. Hence consider the G-module $U:=V^p$. We want to know the G-invariant polynomials on U. For all $i,j=1,\ldots,p$ define $\psi_{ij}:U\to\mathbb{C}$ by

$$\psi_{ij}(v_1,\ldots,v_p)=(v_i|v_j).$$

This is a G-invariant function.

Theorem 11.3.1 (FFT for O_n, polynomial version)**.** *The functions ψ_{ij}, $1\leq i\leq j\leq p$ generate the invariant ring $\mathbb{C}[U]^{\mathrm{O}_n}$.*

Remark 11.3.2. Geometrically, this means that the map sending a matrix $A \in M_{n,p}$ to $A^T A$ is the quotient map for the action of O_n on $M_{n,p}$ by multiplication from the right. The image consists of all symmetric $p \times p$-matrices of rank at most n. Weyl's *Second Fundamental Theorem* for O_n states that the ideal of this image is generated by the $(n+1) \times (n+1)$-minors, in addition to the equations expressing that the matrix be symmetric.

Like in the case of GL_n, this first fundamental theorem follows from the following tensor variant. Let $\beta \in V \otimes V$ be the element corresponding to the identity under the O_n-isomorphism $V \otimes V \to V \otimes V^*$ given by the bilinear form $(.|.)$; equivalently, β equals $\sum_{i=1}^n e_i \otimes e_i$ for any orthonormal basis $e_1, \ldots, e_n$ of V. Then $\beta^{\otimes k} \in V^{\otimes 2k}$ is an O_n-invariant tensor. Moreover, for any permutation $\pi \in S_{2k}$ acting on the $2k$ factors in $V^{\otimes 2k}$ the tensor $\pi\beta^{\otimes k}$ is O_n-invariant, as well.

Theorem 11.3.3 (FFT for O_n, tensor version)**.** *Let d be a natural number. The space $(V^{\otimes d})^{\mathrm{O}_n}$ is trivial if d is odd, and equal to the span of all tensors of the form $\pi\beta^{\otimes k}$, $\pi \in S_{2k}$ if $d = 2k$.*

Exercise 11.3.4. Prove the polynomial version of the FFT for O_n, assuming the tensor version.

The tensor version, in turn, follows from a double commutant argument, where the so-called *Brauer algebra* plays the role of the algebra of linear maps from $V^{\otimes k}$ to itself commuting with all elements of O_n. The arguments are slightly technical, and we do not treat them here.

Exercise 11.3.5. Prove directly, without using the FFT, that $\mathbb{C}[V]^{\mathrm{O}_n}$ does not contain non-zero polynomials of odd degrees.

Exercise 11.3.6. Let m, n be natural numbers. Prove that the dimension of $((S^2(\mathbb{C}^n))^{\otimes m})^{\mathrm{O}_n}$ is at most the number of undirected graphs on the vertices $1, \ldots, m$ in which every vertex has valency 2, i.e., is incident to exactly 2 edges (here a loop on a vertex is allowed and counts as valency 2), and that equality holds for n sufficiently large.

Exercise 11.3.7. In the context of this exercise, a *graph* is a pair $G = (V, E)$ where V ("vertices") is a set and E ("edges") is a multi-set of unordered pairs $\{u, v\}$ with $u, v \in V$. Here u is allowed to equal v ("loops are allowed"), and two vertices may be connected by multiple edges (whence the term "multi-set"). Isomorphisms of such graphs are defined in the obvious manner. The graph G is said to be *k-regular* if for every vertex v the number of edges containing v is k; here the loops from v to itself count twice. See Figure 11.1 for an example.

Let m, n be any natural numbers. Prove that the number of isomorphism classes of k-regular graphs on m vertices is an upper bound to the dimension of the space $(S^m(S^k\mathbb{C}^n))^{\mathrm{O}_n}$ of O_n-invariants in $S^m(S^k\mathbb{C}^n)$.

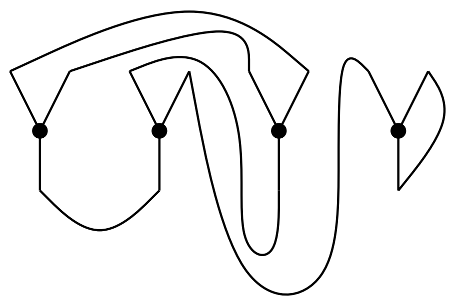

Figure 11.1: A 3-regular graph on 4 vertices.