Invariant Theory

Notes by Bernd Sturmfels
for the lecture on June 19, 2018, in the
IMPRS Ringvorlesung *Introduction to Nonlinear Algebra*

We fix the polynomial ring $K[\mathbf{x}] = K[x_1, \ldots, x_n]$ over a field K of characteristic zero. The group $\mathrm{GL}(n, K)$ of invertible $n \times n$ matrices acts on K^n. This induces an action by G on the ring of polynomial functions on K^n. Namely, if $\sigma = (\sigma_{ij})$ is a matrix in $\mathrm{GL}(n, K)$ and f is a polynomial in $K[\mathbf{x}]$ then σf is the polynomial that is obtained from f by replacing the variable x_i by the linear form $\sum_{j=1}^n \sigma_{ij} x_j$ for $i = 1, \ldots, n$.

Let G be a subgroup of $\mathrm{GL}(n, K)$. A polynomial $f \in K[\mathbf{x}]$ is an *invariant* of the group G if $\sigma f = f$ for all $\sigma \in G$. We write $K[\mathbf{x}]^G$ for the set of all such invariants. This set is a subring because the sum of two invariants is again an invariant, and same for the product.

In this lecture we discuss two scenarios. First we consider finite groups G, and later we consider representations of nice infinite groups like $\mathrm{SL}(d, K)$ and $\mathrm{SO}(d, K)$. Such groups are called *reductive*. A celebrated theorem of Hilbert shows that the invariant ring is finitely generated in this case. After two initial examples, we begin by proving this for finite groups G.

Example 1. Let G be the group of $n \times n$ permutation matrices. The invariant ring $K[\mathbf{x}]^G$ consists of all polynomials f that are invariant under permuting the coordinates, i.e.

$$f(x_{\pi_1}, x_{\pi_2}, \ldots, x_{\pi_n}) \,=\, f(x_1, x_2, \ldots, x_n) \qquad \text{for all permutations } \pi \text{ of } \{1, 2, \ldots, n\}.$$

The invariant ring $K[\mathbf{x}]^G$ is generated by the n elementary symmetric polynomials $E_1, \ldots, E_n$. These are the coefficients of the following auxiliary polynomial in one variable z:

$$(z + x_1)(z + x_2) \cdots (z + x_n) \quad = \quad z^n \,+\, \sum_{i=1}^{n} E_i(\mathbf{x}) z^{n-i}. \tag{1}$$

We also set $E_0 = 1$. Alternatively, $K[\mathbf{x}]^G$ can also be generated by the power sums

$$P_j(\mathbf{x}) \,=\, x_1^j + x_2^j + \cdots + x_n^j \qquad \text{for } j = 1, 2, \ldots, n.$$

The formulas which translate between the E_i and the P_j are known as *Newton's Identities*:

$$\begin{aligned} kE_k &= \textstyle\sum_{i=1}^{k} (-1)^{i-1} E_{k-i} P_i \\ \text{and} \qquad P_k &= (-1)^{k-1} k E_k \,+\, \textstyle\sum_{i=1}^{k-1} (-1)^{k-1-i} E_{k-i} P_i \qquad \text{for } 1 \le k \le n. \end{aligned} \tag{2}$$

Invariants are polynomial functions that are constant along G-orbits on K^n. They offer an algebraic view on the space of orbits. Namely, we think of the spectrum of $K[\mathbf{x}]^G$ as a quotient space $K^n//G$, whose points are these orbits. This interpretation is only informal, as the details are very subtle. Making it all precise is the aim of *Geometric Invariant Theory*.

Example 2. For $n = 2$, consider the following representation of the *cyclic group of order* 4:

$$G \;=\; \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\}. \tag{3}$$

These are the rotational symmetries of the square. Its invariant ring is generated by

$$I_1 = x_1^2 + x_2^2\,, \quad I_2 = x_1^2 x_2^2\,, \quad I_3 = x_1^3 x_2 - x_1 x_2^3.$$

These three invariants are algebraically dependent. Using their relation we can write

$$K[x_1, x_2]^G \;=\; K[I_1, I_2, I_3] \;\simeq\; K[y_1, y_2, y_3]/\langle y_1^2 y_2 - 4y_2^2 - y_3^2 \rangle. \tag{4}$$

The spectrum of the ring (4) is the cubic surface in K^3 defined by the equation $y_1^2 y_2 = 4y_2^2 + y_3^2$. The points on this surface are in one-to-one correspondence with the G-orbits on K^2.

In what follows, we assume that G is a finite subgroup of $\mathrm{GL}(n, K)$. One can create invariants by averaging polynomials. The *Reynolds operator*, denoted by a star, is defined as

$$* \,:\, K[\mathbf{x}] \,\to\, K[\mathbf{x}]^G\,, \quad f \,\mapsto\, f^* := \frac{1}{|G|} \sum_{\sigma \in G} \sigma f. \tag{5}$$

Each of the following properties of the Reynolds operator is easily verified:

Lemma 3. *The Reynolds operators $*$ has the following three properties:*

(a) The map $*$ is K-linear, i.e. $(\lambda f + \nu g)^* = \lambda f^* + \nu g^*$ for all $f, g \in K[\mathbf{x}]$ and $\lambda, \nu \in K$.

(b) The map $*$ restricts to the identity on $K[\mathbf{x}]^G$, i.e. $I^* = I$ for all invariant polynomials I.

(c) The map $*$ is a $K[\mathbf{x}]^G$-module homomorphism, i.e. $(fI)^* = f^* I$ for all $f \in K[\mathbf{x}]$ and $I \in K[\mathbf{x}]^G$.

The following result from 1890 is seen as the beginning of modern Commutative Algebra.

Theorem 4 (Hilbert's Finiteness Theorem)**.** *The invariant ring $K[\mathbf{x}]^G$ of any finite matrix group $G \subset \mathrm{GL}(n, K)$ is finitely generated as a K-algebra.*

We present the proof under the hypothesis that K has characteristic zero. However, the result holds for every field K. For a proof see [3]. This is known as *modular invariant theory*.

Proof. Let $\mathcal{I}_G = \langle K[\mathbf{x}]^G_+ \rangle$ be the ideal in $K[\mathbf{x}]$ that is generated by all homogeneous invariants of positive degree. By Lemma 3 (a), every invariant is a K-linear combination of symmetrized monomials $(\mathbf{x}^{\mathbf{a}})^*$. These homogeneous invariants are the images of monomials under the Reynolds operator. Thus $\mathcal{I}_G$ is generated by the set $\{\,(\mathbf{x}^{\mathbf{a}})^* \,:\, \mathbf{a} \in \mathbb{N}^n \backslash \{0\}\}$. By Hilbert's Basis Theorem, the ideal $\mathcal{I}_G$ is finitely generated, so that a finite subset of $\mathbf{a}$ in $\mathbb{N}^n$ suffices. In conclusion, there exist invariants $I_1, I_2, \ldots, I_m$ such that $\mathcal{I}_G = \langle I_1, I_2, \ldots, I_m \rangle$.

We claim that these m invariants generate the invariant ring $K[\mathbf{x}]^G$ as a K-algebra. Suppose the contrary, and let I be a homogeneous element of minimal degree in $K[\mathbf{x}]^G \backslash K[I_1, I_2, \ldots, I_m]$. Since $I \in \mathcal{I}_G$, we have $I = \sum_{j=1}^m f_j I_j$ for some homogeneous polynomials $f_j \in K[\mathbf{x}]$ whose degrees are all strictly less than $\deg(I)$.

Applying the Reynolds operator on both sides of the equation $I = \sum_{j=1}^m f_j I_j$, we obtain

$$I \,=\, I^* \,=\, \big(\sum_{j=1}^m f_j I_j\,\big)^* \,=\, \sum_{j=1}^m f_j^* I_j.$$

Here we are using the properties (b) and (c) in Lemma 3. The new coefficients f_j^* are homogeneous invariants whose degrees are less than $\deg(I)$. From the minimality assumption on the degree of I, we get $f_j^* \in K[I_1, \ldots, I_m]$ for $j = 1, \ldots, m$. This implies $I \in K[I_1, \ldots, I_m]$, which is a contradiction to our assumption. This completes the proof of Theorem 4. $\square$

Hilbert's Finiteness Theorem also holds for an infinite group $G \subset \mathrm{GL}(n, K)$ that has a Reynolds operator $*$ satisfying the properties (a), (b) and (c) in Lemma 3. Such G are known as *reductive groups*. We shall return to this important point later in the lecture.

Corollary 5. *Consider any reductive group G of $n \times n$-matrices. If $\{g_1, g_2, \ldots, g_m\}$ is any collection of homogeneous polynomials that generates the ideal $\mathcal{I}_G$ then its image $\{g_1^*, g_2^*, \ldots, g_m^*\}$ under the Reynolds operator generates the invariant ring $K[\mathbf{x}]^G$ as a K-algebra.*

Proof. Let $M = \langle x_1, \ldots, x_n \rangle$ be the homogeneous maximal ideal in $K[\mathbf{x}]$, and consider the finite-dimensional vector space $\mathcal{I}_G / M\mathcal{I}_G$. It has a basis of invariants since $\mathcal{I}_G$ is generated by invariants. This means that the Reynolds operator acts as the identity on $\mathcal{I}_G / M\mathcal{I}_G$. The images of $g_1, g_2, \ldots, g_m$ also span $\mathcal{I}_G / M\mathcal{I}_G$ as a vector space, and hence so do the invariants $g_1^*, g_2^*, \ldots, g_m^*$. By Nakayama's Lemma, we find that $g_1^*, g_2^*, \ldots, g_m^*$ generate the ideal $\mathcal{I}_G$. As in the proof of Theorem 4, we conclude that $g_1^*, g_2^*, \ldots, g_m^*$ generate the K-algebra $K[\mathbf{x}]^G$. $\square$

Theorem 6 (Noether's Degree Bound)**.** *If G is finite and* $\mathrm{char}(K) = 0$ *then the invariant ring $K[\mathbf{x}]^G$ is generated by homogeneous invariants of degree at most $|G|$.*

Proof. Let $\mathbf{u} = (u_1, \ldots, u_n)$ be new variables. For any $d \in \mathbb{N}$, we consider the expression

$$\begin{aligned} S_d(\mathbf{u}, \mathbf{x}) \;&=\; \big[\,(u_1 x_1 + \cdots + u_n x_n)^d\,\big]^* \\ &=\; \tfrac{1}{|G|} \textstyle\sum_{\sigma \in G} \big[u_1(\sigma x_1) + \cdots + u_n(\sigma x_n)\big]^d. \end{aligned}$$

This is a polynomial in $\mathbf{u}$ whose coefficients are polynomials in $\mathbf{x}$. Up to a multiplicative constant, they are the invariants $(\mathbf{x}^{\mathbf{a}})^*$ where $|\mathbf{a}| = d$. All polynomials in $\mathbf{u}$ are fixed under $*$.

Consider the $|G|$ expressions $u_1(\sigma x_1)+\cdots+u_n(\sigma x_n)$, one for each group element $\sigma \in G$. The polynomial $S_d(\mathbf{u},\mathbf{x})$ is the dth power sum of these expressions. The power sums for $d > |G|$ are polynomials in the first $|G|$ power sums. Such a representation is derived from Newton's Identities (2). It implies that all $\mathbf{u}$-coefficients of $S_d(\mathbf{u},\mathbf{x})$ for $d > |G|$ are polynomial functions in the $\mathbf{u}$-coefficients of $S_d(\mathbf{u},\mathbf{x})$ for $d \leq |G|$. Hence all invariants $(\mathbf{x}^{\mathbf{a}})^*$ with $|\mathbf{a}| > |G|$ are polynomial functions (over K) in the invariants $(\mathbf{x}^{\mathbf{b}})^*$ with $|\mathbf{b}| \leq |G|$. This proves the claim. □

We note that Example 2 attains Noether's degree bound. The cyclic group in that example has order 4, and the invariant ring requires a generator of degree 4.

Our next theorem is a useful tool for constructing the invariant ring. It says that we can count invariants by averaging the reciprocal characteristic polynomials of the group elements.

Theorem 7 (Molien series)**.** *The Hilbert series of the invariant ring $K[\mathbf{x}]^G$ equals*

$$\sum_{d=0}^{\infty} \dim_K\bigl(K[\mathbf{x}]_d^G\bigr) z^d \quad = \quad \frac{1}{|G|}\sum_{\sigma\in G}\frac{1}{\det(\mathrm{id}-z\sigma)}. \tag{6}$$

The coefficient of z^d in this series is the number of linearly independent invariants of degree d.

Proof. See [4, Theorem 2.2.1]. □

Example 8. Consider the cyclic group $G = \mathbf{Z}_4$ in Example 2. For the four matrices σ in Example 2, the quadratic polynomials $\det(\mathrm{id} - z\sigma)$ are $(1-z)^2$, $1+z^2$, $(1+z)^2$ and $1+z^2$. Adding up their reciprocals and dividing by $|G| = 4$, we see that the Hilbert series of $K[\mathbf{x}]^G$ is

$$\frac{1+z^4}{(1-z^2)(1-z^4)} \quad = \quad 1+z^2+3z^4+3z^6+5z^8+\cdots. \tag{7}$$

This agrees with the Hilbert series of the ring on the right in (4), where $\deg(y_1) = 2$ and $\deg(y_2) = \deg(y_3) = 4$. Indeed, the principal ideal exhibits a Noether normalization. We see that the ring is a free module with basis $\{1, y_3\}$ over $K[y_1, y_2]$. This explains the numerator and denominator on the left of (7), and it proves that I_1, I_2, I_3 do indeed generate $K[\mathbf{x}]^G$.

Classical invariant theory was primarily concerned with the case when G is a representation of the group $\mathrm{SL}(d, K)$ of $d \times d$-matrices with determinant 1. Here d is an integer that is usually much smaller than n and K is a field of characteristic zero. This means that G is the image of a group homomorphism $\mathrm{SL}(d, K) \to \mathrm{GL}(n, K)$. It is known that $\mathrm{SL}(d, K)$ is a reductive group, i.e. there also exists an averaging operator $* : K[\mathbf{x}] \to K[\mathbf{x}]^G$ which has the same formal properties as the averaging operator of a finite group, stated in Lemma 3.

That Reynolds operator $*$ can be realized either by integration or by differentiating. In the first realization, one replaces the sum in (5) by an integral. Namely, one takes $K = \mathbb{C}$ and one integrates over the compact subgroup $\mathrm{SU}(d, \mathbb{C})$ with respect to Haar measure. The same kind of integral also works in Theorem 7. If $G = \mathrm{SL}(d, \mathbb{C})$ then one can compute the Hilbert series of the invariant ring by averaging reciprocal characteristic polynomials.

An alternative to integrating with respect to Haar measure on $\mathrm{SU}(d,\mathbb{C})$ is a certain differential operator known as *Cayley's Ω-process*. This process, which is explained in [4, Section 4.3], can also be used to transform arbitrary polynomials into invariants.

A third method for computing invariants is plain old linear algebra. Indeed, suppose we fix an integer $d \in \mathbb{N}$ and we seek a basis for the space $K[\mathbf{x}]_d^G$ of homogeneous invariants of degree d. We then pick a general polynomial f of degree d with unknown coefficients, and we examine the equations $\sigma f = f$ for $\sigma \in G$. Each of these translates into a linear system of equations in the unknown coefficients of f. By taking enough matrices σ, we obtain a linear system of equations whose solutions are precisely the invariants of degree d. In the case when G is a connected Lie group, like $\mathrm{SL}(d,\mathbb{C})$, one can replace the condition $\sigma f = f$ by requiring that f is annihilated by the associated *Lie algebra*. Setting up these linear equations and solving them is usually quite efficient on small examples. See [4, Section 4.5].

In what follows we take the matrix group to be an n-dimensional polynomial representation of $G = \mathrm{SL}(d,K)$ for some $d, n \in \mathbb{N}$. Each of these representations is a direct sum of irreducible representations, one for each integer partition, as seen in the previous lecture.

Example 9. Let $U = (K^d)^m$ be the space of $d \times m$-matrices. Thus U is the direct sum of m copies of the defining representation of G. The group G acts on U by matrix multiplication on the left. This induces an action on the ring $K[U]$ of polynomials in the entries of a $d \times m$ matrix of variables. If $m < d$ then this action has no non-constant invariants. If $m \geq d$ then the $\binom{m}{d}$ maximal minors of the $d \times m$ matrix are invariants. This invariance holds because the determinant of the product of two $d \times d$-matrices is the product of the determinants. It is known that the invariant ring $K[U]^G$ is generated by these $\binom{m}{d}$ determinants. This result is known as the First Fundamental Theorem of Invariant Theory; cf [4, Section 3.2].

Note that we already encountered the ring $K[U]^G$ in the fourth lecture. It is the coordinate ring of the Grassmannian of d-dimensional subspaces in K^m. Thus, $K[U]^G$ is isomorphic to a polynomial ring in $\binom{m}{d}$ variables, modulo the ideal of quadratic Plücker relations.

Arguably, the most important irreducible representations of the group $G = \mathrm{SL}(d,K)$ are the p-th symmetric powers of the defining representation K^d, where $p \in \mathbb{N}$. We denote such a symmetric power by $V = K[u_1,\ldots,u_d]_p = \mathrm{Sym}_p(K^d)$. Its elements are homogeneous polynomials of degree p in d variables. The G-module V has dimension $n = \binom{p+d-1}{p}$. The monomials form a basis. The action of G on V is simply by linear change of coordinates.

Example 10 (d=2, p=3)**.** Consider the 4-dimensional space $V = \mathrm{Sym}_3(K^2)$ of binary cubics

$$f(u_1,u_2) \quad = \quad x_1u_1^3 + x_2u_1^2u_2 + x_3u_1u_2^2 + x_4u_2^3. \tag{8}$$

The coefficients x_i are the coordinates on $V \simeq K^4$. The way we set things up, the group $\mathrm{SL}(2,K)$ acts on this space by left multiplication, in its guise as the group G of 4×4-matrices

$$\phi(\sigma) \quad = \quad \begin{pmatrix} \sigma_{11}^3 & \sigma_{11}^2\sigma_{12} & \sigma_{11}\sigma_{12}^2 & \sigma_{12}^3 \\ 3\sigma_{11}^2\sigma_{21} & \sigma_{11}^2\sigma_{22}+2\sigma_{11}\sigma_{12}\sigma_{21} & \sigma_{12}^2\sigma_{21}+2\sigma_{11}\sigma_{12}\sigma_{22} & 3\sigma_{12}^2\sigma_{22} \\ 3\sigma_{11}\sigma_{21}^2 & \sigma_{12}\sigma_{21}^2+2\sigma_{11}\sigma_{21}\sigma_{22} & \sigma_{11}\sigma_{22}^2+2\sigma_{12}\sigma_{21}\sigma_{22} & 3\sigma_{12}\sigma_{22}^2 \\ \sigma_{21}^3 & \sigma_{21}^2\sigma_{22} & \sigma_{21}\sigma_{22}^2 & \sigma_{22}^3. \end{pmatrix}. \tag{9}$$

For $\sigma \in G = \mathrm{SL}(2,K)$, the determinant of this 4×4-matrix equals $(\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21})^6 = 1$. The G-action on V is given by $x \mapsto \phi(\sigma)x$ where x is the column vector $(x_1, x_2, x_3, x_4)^T$.

One invariant under this action is the discriminant of the binary cubic $f(u_1, u_2)$, which is

$$\Delta \;=\; 27x_1^2x_4^2 \,-\, 18x_1x_2x_3x_4 \,+\, 4x_1x_3^3 \,+\, 4x_2^3x_4 \,-\, x_2^2x_3^2. \tag{10}$$

It turns out that the discriminant generates the ring of all invariants, i.e. $K[\mathbf{x}]^G = K[\Delta]$.

Invariants of binary forms ($d = 2$) are a well-studied subject in invariant theory. Complete lists of generators for the invariant ring are known up to degree $p = 10$. For $p = 2$, there is also only the discriminant $\Delta = 4x_2 - x_1x_3$. For $p = 4$, we have two generating invariants of degree 2 and 3 respectively. For $p = 10$, the invariant ring has 104 minimal generators.

According to Felix Klein, invariant theory plays a fundamental role for geometry. Namely, a polynomial in the coordinates of a space is invariant under the group of interest if and only if that polynomial expresses a geometric property. For instance, consider the space V of binary cubics f in Example 10. The hypersurface defined by f in $\mathbb{P}^1$ consists of three points. The vanishing of the invariant Δ means that these three points are not all distinct.

In geometric invariant theory, one considers the variety $\mathcal{V}(\mathcal{I}_G)$ defined by all homogeneous invariants of positive degree. This variety is known as the *nullcone*. Its points are known as *unstable points*. For a finite group G, the nullcone consists just of the origin, $V(\mathcal{I}_G) = \{0\}$. For $G = \mathrm{SL}(d,K)$ the situation is more interesting, and the geometry of the nullcone is very important for understanding the invariant ring $K[\mathbf{x}]^G$. Corollary 5 says, more or less, that computing $K[\mathbf{x}]^G$ is equivalent to finding polynomial equations that define the nullcone.

Example 11 (d=p=3)**.** Consider the 10-dimensional space $V = \mathrm{Sym}_3(K^3)$ of *ternary cubics*

$$f(\mathbf{u}) \;=\; x_1u_1^3+x_2u_2^3+x_3u_3^3+x_4u_1^2u_2+x_5u_1^2u_3+x_6u_2^2u_1+x_7u_2^2u_3+x_8u_3^2u_1+x_9u_3^2u_2+x_0u_1u_2u_3.$$

The group $G = \mathrm{SL}(3,K)$ acts on V by linear change of coordinates. The corresponding invariant ring is generated by two invariants I_4 and I_6 of degrees 4 and 6 respectively. In symbols, $K[\mathbf{x}]^G = K[I_4, I_6]$. The degree 4 invariant is the following sum of 25 monomials:

$$\begin{aligned} I_4 \;=\;\; & x_0^4 - 8x_0^2x_4x_9 - 8x_0^2x_5x_7 - 8x_0^2x_6x_8 - 216x_0x_1x_2x_3 + 24x_0x_1x_7x_9 + 24x_0x_2x_5x_8 \\ & + 24x_0x_3x_4x_6 + 24x_0x_4x_7x_8 + 24x_0x_5x_6x_9 + 144x_1x_2x_8x_9 + 144x_1x_3x_6x_7 \\ & - 48x_1x_6x_9^2 - 48x_1x_7^2x_8 + 144x_2x_3x_4x_5 - 48x_2x_4x_8^2 - 48x_2x_5^2x_9 - 48x_3x_4^2x_7 \\ & - 48x_3x_5x_6^2 + 16x_4^2x_9^2 - 16x_4x_5x_7x_9 - 16x_4x_6x_8x_9 + 16x_5^2x_7^2 - 16x_5x_6x_7x_8 + 16x_6^2x_8^2. \end{aligned}$$

The degree 6 invariant is also unique up to scaling. It is a sum of 103 monomials:

$$I_6 \;=\; x_0^6 - 12x_0^4x_4x_9 - 12x_0^4x_5x_7 - 12x_0^4x_6x_8 + 540x_0^3x_1x_2x_3 + \,\cdots\, + 96x_5x_6^2x_7x_8^2 - 64x_6^3x_8^3.$$

The invariant I_4 is the *Aronhold invariant*. This plays an important role in the theory of tensor decomposition. Indeed, we can regard f as a symmetric $3 \times 3 \times 3$-tensor. A random tensor f has rank 4. The Aronhold invariant f vanishes for those tensors of rank ≤ 3. In other words, $I_4 = 0$ holds if and only if f is a sum of three cubes of linear forms, or can be approximated by a sequence of such. See the discussion of ranks of tensors two weeks ago.

On the geometric side, we identify f with the cubic curve $V(f)$ it defines in the projective plane $\mathbb{P}^2$. To a number theorist, this is an *elliptic curve*. An important invariant of this curve is the *discriminant* Δ. This invariant has degree 12 and its explicit formula equals

$$\Delta \quad = \quad I_4^3 \,-\, I_6^2. \tag{11}$$

This expression vanishes if and only if the curve $V(f)$ has a singular point. Typically, this singularity is a *node*. In the special case when both I_4 and I_6 vanish, that singular point is a *cusp*. Thus, for ternary cubics, the nullcone $\mathcal{V}(\mathcal{I}_G)$ is given by plane cubics that have a cusp. The moduli space of elliptic curves is parametrized by the *j-invariant*, which equals I_4^3/Δ.

We now present a general-purpose algorithm, due to Harm Derksen, for computing the invariant ring of a reductive algebraic group G that acts polynomially on a vector space $V = K^n$. The group G can represented as an algebraic variety inside $\mathrm{GL}(n,K)$, that is, by polynomial equations in the entries of an unknown $n \times n$-matrix. This works for both finite groups and for polynomial representations of $\mathrm{SL}(d,K)$, such as the ones discussed about. As before, we use the notation $\sigma \mapsto \phi(\sigma)$ to write the representation of G on $V = K^n$ explicitly.

The product $G \times V \times V$ is an algebraic variety, with coordinates $(\sigma, \mathbf{x}, \mathbf{y})$. Inside its coordinate ring $K[\sigma, \mathbf{x}, \mathbf{y}]$, let $\mathcal{J}_G$ be the ideal generated by the n entries of the vector $\mathbf{y} - \phi(\sigma)\mathbf{x}$. This ideal is radical, and it is prime when G is a connected group like $\mathrm{SL}(d,K)$. Its variety describes the action of the group. The elimination ideal $\mathcal{J}_G \cap K[\mathbf{x},\mathbf{y}]$ is also radical (resp. prime). Its variety consists of pairs of points in V that lie in the same G-orbit.

Theorem 12 (Derksen's Algorithm)**.** *The ideal $\mathcal{I}_G$ of the nullcone is the image in $K[\mathbf{x}]$ of the elimination ideal $\mathcal{J}_G \cap K[\mathbf{x},\mathbf{y}]$ under the substitution $\mathbf{y} = 0$. From any finite list of ideal generators of $\mathcal{I}_G$, algebra generators for the invariant ring $K[\mathbf{x}]^G$ are found via Corollary 5.*

Proof. Let I be any homogeneous invariant of positive degree. Then $I(\mathbf{x}) \equiv I(\phi(\sigma)\mathbf{x}) \equiv I(\mathbf{y})$ modulo the ideal $\mathcal{J}_G$ that defines the group action. Therefore, $I(\mathbf{x}) - I(\mathbf{y})$ lies in the elimination ideal $\mathcal{J}_G \cap K[\mathbf{x},\mathbf{y}]$, and we find $I(\mathbf{x})$ in the ideal that is obtained by substituting $\mathbf{y} = 0$. This proves that $\mathcal{I}_G$ is contained in the ideal that is computed by Derksen's Algorithm. For the converse direction, we refer to the argument given in the proof of [2, Theorem 3.1]. $\square$

Example 13 (p=d=2)**.** Consider the 3-dimensional space $V = \mathrm{Sym}_2(K^2)$ of binary quadrics

$$f(u_1,u_2) \quad = \quad x_1 u_1^2 \,+\, x_2 u_1 u_2 \,+\, x_3 u_2^2.$$

The coordinate ring of the variety $\mathrm{SL}(2,K) \times V \times V$ is the polynomial ring in 10 variables,

$$K[\sigma,\mathbf{x},\mathbf{y}] \quad = \quad K\big[\,\sigma_{11},\sigma_{12},\sigma_{21},\sigma_{22},\, x_1,x_2,x_3,\, y_1,y_2,y_3\,\big],$$

modulo the principal ideal $\langle \sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21} - 1 \rangle$. The ideal that encodes our action equals

$$\begin{aligned}\mathcal{J}_G \;=\; \big\langle\, & \sigma_{11}^2 x_1 + \sigma_{11}\sigma_{21}x_2 + \sigma_{21}^2 x_3 \,-\, y_1,\, \sigma_{12}^2 x_1 + \sigma_{12}\sigma_{22}x_2\sigma_{22}^2 x_3 - y_3, \\ & 2\sigma_{11}\sigma_{12}x_1 + (\sigma_{11}\sigma_{22}+\sigma_{12}\sigma_{21})x_2 + 2\sigma_{21}\sigma_{22}x_3 \,-\, y_2\,\big\rangle\end{aligned}$$

Elimination of the four variables for the group elements yields the principal ideal

$$\mathcal{J}_G \cap K[\mathbf{x},\mathbf{y}] \quad = \quad \langle\, 4x_1x_3 - x_2^2 \,-\, 4y_1y_3 + y_2^2 \,\rangle.$$

We now set $y_1 = y_2 = y_3 = 0$. The result is the familiar discriminant $\Delta = 4x_1x_3 - x_2^2$. In this manner, Derksen's Algorithm finds the invariant ring for binary quadrics $K[\mathbf{x}]^G = K[\Delta]$.

In Example 10, we determined the invariant ring for $\mathrm{SL}(2,K)$ acting on $2 \times 2 \times 2$ tensors that are symmetric. In what follows, we extend this computation to non-symmetric tensors. Thus, we present case study in invariant theory for $d = 2$ and $n = 8$. We identify K^8 with the space $(K^2)^{\otimes 3}$ of $2 \times 2 \times 2$-tensors. The corresponding polynomial ring is denoted by

$$K[\mathbf{x}] \,=\, K[x_{111}, x_{112}, x_{121}, x_{122}, x_{211}, x_{212}, x_{221}, x_{222}].$$

The group $G = \mathrm{SL}(2,K)$ acts on K^2 by matrix-vector multiplication. This action extends naturally to the triple tensor product of K^2. Explicitly, if $\sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$ is a 2×2-matrix in G then σ acts by performing the following substitution in each polynomial on $K[\mathbf{x}]$:

$$x_{ijk} \,\mapsto\, \sum_{r=1}^{2} \sum_{s=1}^{2} \sum_{t=1}^{2} x_{rst} \sigma_{ri} \sigma_{sj} \sigma_{tk}. \tag{12}$$

Here are two nice polynomials that are invariant under this action:

Example 14. Up to scaling, there is a unique polynomial of degree 2 that is invariant under $G = \mathrm{SL}(2,K)$. That invariant is the following quadric, which we call the *hexagon invariant*:

$$\mathrm{Hex}(\mathbf{x}) \,=\, x_{112}x_{122} - x_{122}x_{121} + x_{121}x_{221} - x_{221}x_{211} + x_{211}x_{212} - x_{212}x_{112}.$$

Our second nice invariant is homogeneous of degree four. This is the *hyperdeterminant*

$$\begin{aligned} \mathrm{Det}(\mathbf{x}) \;=\; & x_{221}^2x_{112}^2 + x_{211}^2x_{122}^2 + x_{121}^2x_{212}^2 + x_{111}^2x_{222}^2 + 4x_{111}x_{221}x_{122}x_{212} + 4x_{121}x_{211}x_{112}x_{222} \\ & - 2x_{211}x_{221}x_{112}x_{122} - 2x_{121}x_{221}x_{112}x_{212} - 2x_{121}x_{211}x_{122}x_{212} \\ & - 2x_{111}x_{221}x_{112}x_{222} - 2x_{111}x_{211}x_{122}x_{222} - 2x_{111}x_{121}x_{212}x_{222}. \end{aligned}$$

One checks by computation that the substitution (12) maps the hexagon invariant $\mathrm{Hex}(\mathbf{x})$ to itself times the third power of $\det(\sigma) = \sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}$. Similary, the hyperdeterminant $\mathrm{Det}(\mathbf{x})$ transforms to itself times $\det(\sigma)^6$. Hence both are invariant when $\det(\sigma) = 1$.

Invariants can be used to test whether two tensors lie in the same orbit. Here is a concrete example. We write our $2 \times 2 \times 2$ tensors as vectors in $\mathbb{R}^8$ as follows: $\mathbf{c} = (c_{111}, c_{112}, c_{121}, c_{122}, c_{211}, c_{212}, c_{221}, c_{222})$. The following two tensors appear in the theory of signatures of paths. It is of interest to know whether their G-orbits agree up to scaling:

$$\mathbf{c}_{\mathrm{axis}} \,=\, \left(\frac{1}{6}\,,\,\frac{1}{2}\,,\,0\,,\,\frac{1}{2}\,,\,0\,,\,0\,,\,0\,,\,\frac{1}{6}\right) \quad \text{and} \quad \mathbf{c}_{\mathrm{mono}} \,=\, \left(\frac{1}{6}\,,\,\frac{1}{4}\,,\,\frac{1}{6}\,,\,\frac{4}{15}\,,\,\frac{1}{12}\,,\,\frac{2}{15}\,,\,\frac{1}{10}\,,\,\frac{1}{6}\right).$$

The two polynomials in Example 14 are relative invariants of the GL(2) action on the tensor space $\mathbb{R}^8$. The following rational function is an absolute invariant. It is homogeneous of degree zero, so it represents an invariant rational function on the projective space $\mathbb{P}^7$:

$$\frac{\mathrm{Hex}(\mathbf{x})^2}{\mathrm{Det}(\mathbf{x})}. \tag{13}$$

We find that the invariant (13) evaluates to 81 on $\mathbf{c}_{\mathrm{axis}}$, and it evaluates to 45 on $\mathbf{c}_{\mathrm{mono}}$. Hence the orbit closures of our two special core tensors of format $2 \times 2 \times 2$ are disjoint in $\mathbb{P}^7$.

We now come to determination of the full ring of invariants for the G-action on the space K^8 of $2 \times 2 \times 2$ tensors. Using *Derksen's Algorithm*, we derive the following result.

Theorem 15. *The invariant ring $K[\mathbf{x}]^{\mathrm{SL}(2)}$ of $2 \times 2 \times 2$ tensors has Krull dimension five. It is minimally generated by 13 invariants, namely the hexagon invariant of degree two, eight invariants of degree four (including the hyperdeterminant), and four invariants of degree six.*

In addition to the hyperdeterminant, there are three additional invariants of degree four that deserve special attention. Each has 17 terms when expanded. One of these invariants is

$$\begin{aligned}&(x_{111}x_{222} - x_{212}x_{121})^2 \,+\, x_{121}x_{222}x_{112}^2 + x_{111}x_{212}x_{122}^2 + x_{121}x_{222}x_{211}^2 + x_{111}x_{212}x_{221}^2 \\ &-(x_{122}+x_{221})(x_{112}+x_{211})(x_{111}x_{222}+x_{212}x_{121}) \,+\, 2x_{111}x_{122}x_{212}x_{221} + 2x_{112}x_{121}x_{211}x_{222}.\end{aligned} \tag{14}$$

The other two invariants in this family are obtained by permuting indices.

Corollary 16. *The three quartics in (14) together with* Hex *and* Det *form an algebraically independent system of five primary invariants. All other invariants in $K[\mathbf{x}]^{\mathrm{SL}(2)}$ are integral over the polynomial subring generated by these five. The five primary invariants cut out the null cone $\mathcal{V}\big(K[\mathbf{x}]_+^{\mathrm{SL}(2)}\big)$, which is a variety of dimension four and degree 12 in $\mathbb{P}^7$.*

It is instructive to restrict the 13 generating invariants in Theorem 15 to the 4-dimensional subspace $\mathrm{Sym}_3(K^2)$ of symmetric $2\times2\times2$ tensors, seen in Example 10. We do this by setting

$$x_{111} = x_1\,,\;\; x_{112} = x_{121} = x_{211} = \frac{1}{3}x_2\,,\;\; x_{122} = x_{212} = x_{221} = \frac{1}{3}x_3\,,\;\; x_{222} = x_4.$$

The resulting symmetric tensors correspond to binary cubics (8). The hyperdeterminant and five other generators of degree four specialize to the *discriminant* Δ of the binary cubic. The other eight generators of $K[\mathbf{x}]^{\mathrm{SL}(2)}$, including the hexagon invariant, specialize to zero. In this manner, the invariant ring in Theorem 15 maps onto the invariant ring of binary cubics.

Exercises

1. Let G be the symmetry group of the square $[-1,1]^2$ in the plane $\mathbb{R}^2$. This is an order 8 subgroup in $\mathrm{GL}(2,\mathbb{R})$. List all eight matrices. Determine the invariant ring $\mathbb{R}[x_1,x_2]^G$.

2. Let G be the symmetry group of the regular 3-cube, as a subgroup of $\mathrm{GL}(3,\mathbb{R})$. How many matrices are in G, and what are their characteristic polynomials? Determine the Molien series (7) of this group. What does it tell you about the invariant ring?

3. Fix $n = 5$. Let $\psi(j)$ denote the number of monomials in the expansion of the power sum P_j in terms of the elementary symmetric functions E_1, E_2, E_3, E_4, E_5. Compute $\psi(j)$ for some small values, say $j \leq 20$. Guess a formula for $\psi(j)$. Can you prove it?

4. Show that Noether's Degree Bound is always tight for finite cyclic groups.

5. Find a subgroup of $\mathrm{GL}(4, K)$ that has order 15. Compute the invariant ring.

6. Let T be the group of 3×3 diagonal matrices with determinant 1, acting on the space $V = \mathrm{Sym}_3(K^3)$ of ternary cubics. This group is the torus $T \simeq (K^*)^2$. Determine the invariant ring $K[V]^T$. Do you see any relationship to the invariants in Example 11?

7. Let $G = A_n$ be the *alternating group* of order $n!/2$. Its elements are the even permutation matrices. Determine the invariant ring $K[\mathbf{x}]^G$.

8. List all 103 monomials of the invariant I_6 of ternary cubics in Example 11. Give an explicit formula, in terms of $x_1, x_2, \ldots, x_9, x_0$, for the discriminant and the j-invariant.

9. Consider the action of $\mathrm{SL}(3, K)$ on the space $\mathrm{Sym}_2(K^3) \simeq K^6$ of symmetric 3×3-matrices. The entries of the 6×6 matrix $\phi(\sigma)$ are quadratic forms in $\sigma_{11}, \sigma_{12}, \ldots, \sigma_{33}$. Write this matrix explicitly, similarly to (9). What is the invariant ring?

10. Using Derksen's Algorithm, determine the invariants for binary quartics ($d = 2, p = 4$).

11. The rotation group $\mathrm{SO}(2, \mathbb{R})$ acts by left multiplication on the space of 2×2-matrices. Determine the invariant ring.

12. Is the invariant ring of every matrix group $G \subset \mathrm{GL}(n, K)$ finitely generated?

References

[1] D. Cox, J. Little and D. O'Shea: *Ideals, Varieties, and Algorithms*. An introduction to computational algebraic geometry and commutative algebra, Third edition, Undergraduate Texts in Mathematics, Springer, New York, 2007.

[2] H. Derksen: *Computation of invariants of reductive groups*, Advances in Mathematics **141** (1999) 366-384.

[3] H. Derksen and G. Kemper: *Computational Invariant Theory*. Invariant Theory and Algebraic Transformation Groups, I. Encyclopaedia of Mathematical Sciences, **130**, Springer-Verlag, Berlin, 2002.

[4] B. Sturmfels: *Algorithms in Invariant Theory*, Texts and Monographs in Symbolic Computation, Springer-Verlag, Vienna, 1993.