Introduction to Invariant Theory and Binary Forms

Carol Wu

18 January 2026

§1 Introduction (10 min.)

Invariant theory studies how mathematical objects behave under transformations; More specifically, it studies which properties of the objects remain unchanged under these transformations.

This idea is extremely general, and leads to many different branches of invariant theory depending on the group and the space under consideration.

Example 1.1 (Euclidean Geometry)

In the Euclidean plane, the natural transformations consist of rotations, reflections, and translations. Under these transformations, distance and angles are preserved. For a finite set of points, the squared distances

$$D_{ij} = (x_i - x_j)^2 + (y_i - y_j)^2$$

are invariants, and in fact they generate all polynomial invariants on these points.

If we enlarge the symmetry group by allowing uniform scalings, then distances are no longer meaningful, since they are no longer invariant. Angles, however, remain unchanged (This illustrates a principle of Klein's *Erlangen Program*: a geometry is determined by its symmetry group, and its fundamental notions are precisely the invariants of that group).

In modern invariant theory, we study a group Γ acting on a space V, and we ask for functions on V that do not change under the action of Γ.

Example 1.2

Consider the set

$$M_n(\mathbb{C}) := \{ \text{ all } n \times n \text{ complex matrices}\}$$

and the group

$$\mathrm{SL}_n(\mathbb{C}) := \{\pi \in M_n(\mathbb{C}) : \det \pi = 1\}.$$

$\mathrm{SL}_n(\mathbb{C})$ acts on the set $M_n(\mathbb{C})$ by left multiplication:

$$\begin{aligned} \pi : M_n(\mathbb{C}) &\to M_n(\mathbb{C}) \\ A &\mapsto \pi \cdot A. \end{aligned}$$

The determinant is an invariant of this action:

$$\det(\pi A) = \det(A) \quad \forall \pi \in \mathrm{SL}_n(\mathbb{C}).$$

This action decomposes $M_n(\mathbb{C})$ into *orbits*, where two matrices lie in the same orbit if one can be obtained from the other by multiplying by an element of $\mathrm{SL}_n(\mathbb{C})$. An invariant is exactly a function that is constant on each orbit.

Thus invariant theory can be rephrased as the study of functions on the space of orbits.

Example 1.3

Let the rotations of a square $\mathbb{Z}_4$ act on the plane $\mathbb{R}^2$. The invariants are precisely the functions on $\mathbb{R}^2$ that take the same value at all points in each $\mathbb{Z}_4$-orbit.

Equivalently, if we form the quotient space (orbifold) $\mathbb{R}^2/\mathbb{Z}_4$, then invariant functions on $\mathbb{R}^2$ are exactly the ordinary functions on $\mathbb{R}^2/\mathbb{Z}_4$.

The ring of invariants thus plays the role of the coordinate ring (the ring of polynomial functions) on an orbit variety. Consequently, if one wishes to study the ring of polynomial functions on an algebraic variety that can be realized as a quotient by a group action, invariant theory provides a natural tool to do so.

As observed from the multitude of examples above, invariant theory has many different flavours. The classical setup, which we are interested in, is as follows.

§1.1 Setup

Let $\Gamma \subseteq \mathrm{GL}_n(\mathbb{C})$. Γ acts on polynomials $f \in \mathbb{C}[x_1, \dots, x_n]$ by linear change of variables:

$$(\pi \cdot f)(x_1, \dots, x_n) = f(\pi(x_1, \dots, x_n)), \quad \pi \in \Gamma.$$

Example 1.4

Consider the binary form $f(x_1, x_2) = x_1^2 + x_1 x_2$ and the linear map $\pi = \begin{pmatrix} 3 & 5 \\ 4 & 7 \end{pmatrix}$. Then,

$$(\pi \cdot f)(x_1, x_2) = f(3x_1 + 5x_2, 4x_1 + 7x_2) = (3x_1 + 5x_2)^2 + (3x_1 + 5x_2)(4x_1 + 7x_2).$$

We are interested in the *invariant subring*

$$\mathbb{C}[x_1, \dots, x_n]^\Gamma := \{f \in \mathbb{C}[x_1, \dots, x_n] : f = \pi \cdot f \quad \forall \pi \in \Gamma\}$$

of all polynomials which are invariant under the action of Γ.

In particular, we aim to solve the following problems.

1. Find a set of *fundamental invariants* $\{I_1, \dots, I_m\}$ which generate $\mathbb{C}[x_1, \dots, x_n]^\Gamma$.
2. Find all *syzygies* (algebraic relations) amongst $I_1, \dots, I_m$.
3. Give an algorithm for rewriting arbitrary invariants $I \in \mathbb{C}[x_1, \dots, x_n]^\Gamma$ as a polynomial in $I_1, \dots, I_m$.

Gröbner bases provide computational tools for both finding syzygies (Problem 2) and rewriting invariants (Problem 3).

§2 Symmetric Polynomials

Definition 2.1. $f \in \mathbb{C}[x_1, \dots, x_n]$ is *symmetric* if

$$f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n) \quad \forall \sigma \in \S_n.$$

The most important examples of symmetric polynomials are:

Example 2.2 (Elementary Symmetric Polynomials)

$$\begin{aligned} e_1 &= x_1 + x_2 + \cdots + x_n \\ e_2 &= x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n = \sum_{1 \le i < j \le n} x_i x_j \\ e_k &= \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k} \\ e_n &= x_1 x_2 \cdots x_n \end{aligned}$$

Theorem 2.3 (Fundamental Theorem of Symmetric Polynomials)
The ring of symmetric polynomials is generated by the elementary symmetric polynomials:

$$\mathbb{C}[x_1, \dots, x_n]^{S_n} = \mathbb{C}[e_1, \dots, e_n].$$

Moreover, the e_i are algebraically independent (no syzygies).

Example 2.4
For $n = 2$, we can express any symmetric polynomial in terms of $e_1 = x_1 + x_2$ and $e_2 = x_1x_2$:

$$x_1^3 + x_2^3 = (x_1 + x_2)^3 - 3(x_1x_2)(x_1 + x_2) = e_1^3 - 3e_1e_2.$$

Example 2.5 (Power Sums)
Another generating set is the power sums:

$$p_k = x_1^k + x_2^k + \cdots + x_n^k, \quad k = 1, \dots, n.$$

We have $\mathbb{C}[x]^{S_n} = \mathbb{C}[p_1, \dots, p_n]$ as well.

Example 2.6

For $n = 3$, we can express any symmetric polynomial in terms of $p_1 = x_1 + x_2 + x_2$, $p_2 = x_1^2 + x_2^2 + x_3^2$, and $p_3 = x_1^3 + x_2^3 + x_3^3$:

$$x_1x_2x_3 = \frac{1}{6}p_1^3 - \frac{1}{2}p_1p_2 + \frac{1}{3}p_3.$$

§2.1 Schur Polynomials

Definition 2.7. A **partition** of d is a weakly decreasing sequence $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0)$ with $\sum_{i=1}^n \lambda_i = d$.

Definition 2.8. For a partition λ, define the *alternating polynomial*:

$$a_\lambda(x_1, \ldots, x_n) = \det \begin{pmatrix} x_1^{\lambda_1+n-1} & x_2^{\lambda_1+n-1} & \cdots & x_n^{\lambda_1+n-1} \\ x_1^{\lambda_2+n-2} & x_2^{\lambda_2+n-2} & \cdots & x_n^{\lambda_2+n-2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{\lambda_n} & x_2^{\lambda_n} & \cdots & x_n^{\lambda_n} \end{pmatrix}.$$

Example 2.9

For $\lambda = (1, 1)$ and $n = 2$:

$$a_\lambda(x_1, x_2) = \det \begin{pmatrix} x_1^2 & x_2^2 \\ x_1 & x_2 \end{pmatrix} = x_1^2x_2 - x_1x_2^2.$$

This polynomial is **alternating**: $a_\lambda(x_2, x_1) = -a_\lambda(x_1, x_2)$.

These alternating polynomials a_λ are essentially antisymmetrizations of monomials. They form a basis for the space of alternating polynomials (those satisfying $f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \operatorname{sgn}(\sigma)f(x_1, \ldots, x_n)$).

Definition 2.10. The *Schur polynomial* associated to λ is:

$$s_\lambda(x_1, \ldots, x_n) = \frac{a_\lambda(x_1, \ldots, x_n)}{a_\rho(x_1, \ldots, x_n)}$$

where $\rho = (n-1, n-2, \ldots, 1, 0)$ is the staircase partition, and

$$a_\rho(x_1, \ldots, x_n) = \det \begin{pmatrix} x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \\ x_1^{n-2} & x_2^{n-2} & \cdots & x_n^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^0 & x_2^0 & \cdots & x_n^0 \end{pmatrix} = \prod_{1 \leq i < j \leq n} (x_i - x_j)$$

is the Vandermonde determinant.

Theorem 2.11

The set $\{s_\lambda : \lambda \text{ partition of } d \text{ with at most } n \text{ parts}\}$ forms a basis for the degree-d homogeneous symmetric polynomials:

$$\mathbb{C}[x_1, \ldots, x_n]_d^{S_n} = \mathbb{C}[\{s_\lambda : \lambda = (\lambda_1 \geq \ldots \geq \lambda_n), |\lambda| = d\}].$$

§3 Binary Forms

§3.1 Setup

Definition 3.1. *A binary form of degree n* is a homogeneous polynomial

$$f(x,y)=\sum_{k=0}^{n}\binom{n}{k}a_k x^k y^{n-k}\in\mathbb{C}[x,y].$$

The binomial coefficients $\binom{n}{k}$ are included mainly for technical convenience—they make many formulas cleaner.

The group $\mathrm{GL}_2(\mathbb{C})$ acts on binary forms by linear changes of variables. Given $\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22}\end{pmatrix}\in\mathrm{GL}_2(\mathbb{C})$, we define new coordinates:

$$x=c_{11}\bar{x}+c_{12}\bar{y},\quad y=c_{21}\bar{x}+c_{22}\bar{y}.$$

Equivalently in vector form,

$$\begin{pmatrix}\bar{x}\\ \bar{y}\end{pmatrix}=\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22}\end{pmatrix}^{-1}\begin{pmatrix}x\\ y\end{pmatrix}.$$

Substitution into f yields the new binary form

$$\bar{f}(\bar{x},\bar{y})=f(c_{11}\bar{x}+c_{12}\bar{y},\ c_{21}\bar{x}+c_{22}\bar{y})=\sum_{k=0}^{n}\binom{n}{k}\bar{a}_k\bar{x}^k\bar{y}^{n-k},$$

where the new coefficients

$$\bar{a}_k=\sum_{i=0}^{n}\left(\sum_{j=\max(0,i-n+k)}^{\min(i,k)}\binom{k}{j}\binom{n-k}{i-j}c_{11}^{j}c_{12}^{i-j}c_{21}^{k-j}c_{22}^{n-k-i+j}\right)a_i.$$

are linear combinations of the original a_i with coefficients that are polynomials in the c_{ij}.

Example 3.2 (Quadratic Binary Forms ($n=2$))

For $f=a_2x^2+2a_1xy+a_0y^2$,

$$\bar{f}(\bar{x},\bar{y})=\bar{a}_2\bar{x}^2+2\bar{a}_1\bar{x}\bar{y}+\bar{a}_0\bar{y}^2,$$

where

$$\begin{pmatrix}\bar{a}_0\\ \bar{a}_1\\ \bar{a}_2\end{pmatrix}=\begin{pmatrix} c_{22}^2 & 2c_{22}c_{12} & c_{12}^2\\ c_{21}c_{22} & c_{11}c_{22}+c_{12}c_{21} & c_{11}c_{12}\\ c_{21}^2 & 2c_{11}c_{21} & c_{11}^2\end{pmatrix}\begin{pmatrix}a_0\\ a_1\\ a_2\end{pmatrix}.$$

We observe above that the action of $\mathrm{GL}_2(\mathbb{C})$ on binary quadratics induces a linear action

$$\begin{pmatrix}\bar{a}_0\\ \bar{a}_1\\ \bar{a}_2\\ \bar{x}\\ \bar{y}\end{pmatrix}=\begin{pmatrix}\begin{pmatrix} c_{22}^2 & 2c_{22}c_{12} & c_{12}^2\\ c_{21}c_{22} & c_{11}c_{22}+c_{12}c_{21} & c_{11}c_{12}\\ c_{21}^2 & 2c_{11}c_{21} & c_{11}^2\end{pmatrix} & 0\\ 0 & \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22}\end{pmatrix}^{-1}\end{pmatrix}\begin{pmatrix}a_0\\ a_1\\ a_2\\ x\\ y\end{pmatrix}.$$

on the polynomial ring $\mathbb{C}[a_0,a_1,a_2,x,y]$. In general, the action of $\mathrm{GL}_2(\mathbb{C})$ on binary forms induces a linear action on the polynomial ring $\mathbb{C}[a_0,\dots,a_n,x,y]$. From the invariant–theoretic point of view, we are interested in polynomial functions on $\mathbb{C}[a_0,\dots,a_n,x,y]$ that are constant (or scale in a controlled way) under this induced action.

§3.2 Invariants of Binary Forms

Recall our earlier definition: For $\Gamma \subseteq \mathrm{GL}_n(\mathbb{C})$ acting on $\mathbb{C}[x_1, \dots, x_n]$, an *invariant* satisfies

$$\pi \cdot f = f \quad \text{for all } \pi \in \Gamma.$$

For binary forms, we use a more general notion:

Definition 3.3. A polynomial $f \in \mathbb{C}[a_0, \dots, a_n, x, y]$ is a *covariant of index* g if, for all $\pi \in \mathrm{GL}_2(\mathbb{C})$,

$$\pi \cdot f = (\det \pi)^g f \quad \text{for all } g \in \mathrm{GL}_2(\mathbb{C}).$$

If f has index 0 then it is an *invariant.*

Let

$$f_2 = a_2x^2 + 2a_1xy + a_0y^2,$$
$$f_3 = b_3x^3 + 3b_2x^2y + 3b_1xy^2 + b_0y^3,$$
$$f_4 = c_4x^4 + 4c_3xsy + 6c_2x^2y^2 + 4c_1xy^3 + c_0y^4.$$

Example 3.4 (Discriminant of Quadratic)

The discriminant

$$D_2(a_0, a_1, a_2) := a_0a_2 - a_1^2$$

of a quadratic satisfies

$$D_2(\bar{a}_0, \bar{a}_1, \bar{a}_2) = (\det \pi)^2 D_2(a_0, a_1, a_2).$$

It is an invariant of index 2. It vanishes if and only if f has a double root.

Example 3.5 (Sylvester Resultant)

The Sylvester resultant

$$Res_{2,3}(a_0, a_1, a_2, b_0, b_1, b_2, b_3) := \det \mathrm{S}_{f_2, f_3}$$

is the determinant of the Sylvester matrix of f_2, f_3. It vanishes if and only if f_2, f_3 have a common root.

Example 3.6 (Catelecticant)

The catelecticant of a quartic is defined as

$$C_4 := \det \begin{pmatrix} c_0 & c_1 & c_2 \\ c_1 & c_2 & c_3 \\ c_2 & c_3 & c_4 \end{pmatrix}.$$

It is a fact that all binary quartics can be written as the sum of three fourth powers of linear forms:

$$f_4(x, y) = \ell_1(x, y)^4 + \ell_2(x, y)^4 + \ell_3(x, y)^4.$$

The catelecticant is 0 if and only if f_4 can be written as the sum of only two linear fourth powers.

Although the coefficients of a binary form change under a linear change of variables, the geometry of its roots does not. Invariant theory isolates precisely those polynomial expressions in the coefficients (and variables) that are unchanged, or change only by a power of $\det \pi$, under the action of $\mathrm{GL}_2(\mathbb{C})$. These invariants and covariants therefore encode coordinate-free information about the polynomial: whether roots collide (discriminants), whether two forms share a root (resultants), or whether a form decomposes into few sums of powers (catelecticants).