An Introduction to Invariant Theory

Krishna Dhulipala
Euler Circle

July 13, 2020

"The theory of invariants came into existence about the middle of the nineteenth century somewhat like Minerva: a grown-up virgin, mailed in the shining armor of algebra, she sprang forth from Cayley's Jovian head."

– Hermann Weyl [1]

Abstract

The birth of Invariant Theory is credited to the 19th-century mathematicians Arthur Cayley and George Boole, who, in [2] and [1], respectively, began the study of describing polynomial forms under linear transformations. This work was carried on by mathematicians like Hilbert and Klein, and inspired important results like Hilbert's basis and finiteness theorems, as well as entire fields of study like moduli spaces, symmetric functions, and commutative algebra. This paper will introduce the theory of invariants by investigating the behaviour of polynomial forms over linear transformations, and will go on to discuss invariance under group action, Dickson's Lemma, and finally Hilbert's basis and finiteness theorems.

1 Preliminaries

Definition 1.1 (Forms)**.** Let $\mathcal{F}(x_1, x_2, \ldots, x_r) \in \mathbb{C}[x]$ be a polynomial of r variables. We can then rewrite this polynomial by grouping all the terms of degree i under the constituent function F_i, to yield:

$$\mathcal{F}(x_1, x_2, \ldots, x_r) = F_1 + F_2 + \cdots + F_n.$$

If all terms other than those contained in $[n]$ vanish, then we call $\mathcal{F}(x_1, x_2, \ldots, x_r)$ a *form*, or a *homogeneous polynomial*. In short, the nonzero terms for a form all share the same degree. For the form of r variables containing terms of some order n, we write $\mathcal{F}(x_1, x_2, \ldots, x_r)$. Forms will be heavily utilized throughout this paper, as we shall study how certain properties of forms behave under linear transformations.[2]

[1] This quote was taken from [5]

[2] This introduction to forms and transformations over forms has been largely inspired by Hilbert's own book on the theory of algebraic invariants. For a much deeper look into topics discussed here, visit [3].

Example. Suppose we were to work only in two variables, say, x and y. Then, let c_1, c_2, and c_3 be from $\mathbb{C}$. We can construct a form from these, namely the homogeneous polynomial $f(x, y) = c_1x^2 + c_2xy + c_3y^2$. This form, called the *binary form*, is very convenient, so we shall continually refer to it when dealing with other form-related concepts.

Definition 1.2 (Forms over Transformations). Let us consider the form given by $\mathcal{F}(x_1, x_2, \ldots, x_r)$. Then, we can attempt to generate a different form $\mathcal{F}'$ in which each variable x_i is substituted by x_i', where the relation between x_i and x_i' is given as follows: for forms $\psi_1, \psi_2, \ldots, \psi_n$ having the same order, write

$$x_i = \psi_i(x_1', x_2', \ldots, x_n').$$

We can then represent this new form $\mathcal{F}'$ as a variation of the original form $\mathcal{F}$, namely:

$$\mathcal{F}'(x_1', x_2', \ldots, x_n') = \mathcal{F}(\psi_1(x_1', \ldots, x_n'), \psi_2(x_1', \ldots, x_n'), \ldots, \psi_n(x_1', \ldots, x_n')).$$

We call this newly generated form $\mathcal{F}'$ the *transformed form*. The most important parts of Invariant Theory concern themselves with linear transformations, so we will mainly be examining linearly transformed forms.

Definition 1.3 (Linear Transformations). We define *linear transformations* to be the transformations for which the forms $\psi_1, \psi_2, \ldots, \psi_n$ operating on the x_i's are linear forms. This means that for each x_i, we have:

$$x_i = c_{i1}x_1' + c_{i2}x_2' + \cdots + c_{in}x_n'$$

. We will use the following three properties to categorize linear transformations:

1. Both the original and transformed forms have the same order.

2. The transformation is invertible; this makes sense, because linear transformations only require one to solve a system of equations in order to retrieve the original variables from the transformed form. However, this quality only holds if the linear transformation has a nonzero *transformation determinant*, which is the determinant of the c_{ij} coefficients, namely

$$\delta = \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ & & \cdots & \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix}$$

3. The composition of multiple linear transformations is equivalent to a single linear transformation. (This is sometimes referred to as the *group property of linear transformations*)

Example. Let $f(x, y)$ be a binary form with coefficients c_1, c_2, and c_3. Suppose that D were a 2×2 matrix with determinant 1, with

$$D = \begin{bmatrix} q & u \\ v & w \end{bmatrix}.$$

Then, D gives a linear change of coordinates in $\mathbb{C}^2$, since we can perform the substitution of variables $(x, y) \mapsto (qx + vy, ux + wy)$, and yield the new polynomial given by $f'(x, y) = f(qx + vy, ux + wy) = c_1'x^2 + c_2'xy + c_3'y^2$, where the coefficients c_1', c_2', and c_3' are generated by the matrix operation

$$\begin{bmatrix} c_1' \\ c_2' \\ c_3' \end{bmatrix} = \begin{bmatrix} q^2 & qu & u^2 \\ qv & qw + uv & uw \\ v^2 & vw & w^2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}.$$

Thus, we have generated a linearly transformed form $f'(x, y)$ from the transformation matrix D.

2 A First Look at Invariants and Covariants

It is natural when considering transformed forms to wonder whether there exist even simpler types of relationships between $\mathcal{F}$ and $\mathcal{F}'$. In the case where the transformation is linear, we are able find such a relationship. Certain transformed forms, for example, may be scalings of the original form by some constant factor. In other instances, transformed forms may be scalings of a slight variation of the original form by some constant factor. The notions of *invariance* and *covariance* are motivated by these types of linearly transformed forms.

Definition 2.1 (Invariance). An *invariant* of the form $\mathcal{F}$ is defined as a polynomial having the coefficients $c_0, c_1, \ldots, c_n$ that changes only by a factor equal to a power of the determinant δ of the transformation matrix if one replaces the coefficients $c_0, c_1, \ldots, c_n$ of the given base form by the corresponding coefficients $c_0', c_1', \ldots, c_n'$ of the linearly transformed form. Thus, we are able to describe invariants using the following relation:

$$\mathcal{I}(c_0', c_1', \ldots, c_n') = \delta^s \mathcal{I}(c_0, c_1, \ldots, c_n).$$

Definition 2.2 (Covariance). We define a *covariant* as a polynomial of the coefficients c_0 $c_1, \ldots, c_n$ and the variables x_1, x_2 that changes only by a factor equal to a power of the determinant δ of the transformation matrix under the following transformation: replacing the coefficients $a_0, a_1, \ldots, a_n$ with coefficients of the linearly transformed form $a_0', a_1', \ldots, a_n'$, and replacing the variables x_1, x_2 with the linearly transformed variables x_1', x_2'.

Then, we are able to describe covariants using the following relation:

$$\mathcal{C}\left(\{c_0', c_1', \ldots, c_n'\}, \{x_1', x_2'\}\right) = \delta^p \mathcal{C}\left(\{c_0, c_1, \ldots, c_n\}, \{x_1, x_2\}\right).$$

We can also think of covariance as a generalization of invariance; an invariant is just a covariant in which the degree p of the transformation determinant is simply 1.

Definition 2.3 (Group Action). Suppose that G were a group, and S some set. Then, we define a *group action* of G on S to be a map $\phi : G \times X \to X$ such that for all $s \in S$ and $g_1, g_2 \in G$, we have $\phi(1, s) = s$ as well as $\phi(g_2, \phi(g_1, s)) = \phi(g_2 g_1, s)$.[3]

[3]Though we use the set S here for general purposes, we will be working primarily with sets of functions such as $\mathbb{C}[x]$, so that we can observe functions and forms which are invariant under group actions.

Definition 2.4 (Invariance under Group Action). Let $f(x)$ be some function in $\mathbb{C}[x]$. We say that f is *G-invariant* if, for all $g \in G$, we have $(gf)(x) = f(x)$. We then define the set of all G-invariant polynomials $f \in \mathbb{C}[x]$ as $\mathbb{C}[x]^G$.

3 Hilbert's Theorems

Now that we have established the necessary preliminaries, we are able to discuss some interesting theorems that are rooted in Invariant Theory. We shall begin by stating Dickson's Lemma, which is used to justify Hilbert's Basis Theorem, an important result that we will make use of in the proof of the finiteness theorem.

Lemma 3.1 (Dickson'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 $m_i \mid m_j$.*

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 $n - 1 \geq 0$. Then, we can define the infinite sequences of exponents $e_1 \leq e_2 \leq \ldots$ and $i_1 < i_2 < \ldots$, such that 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. Then, 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. Thus, by induction, there must exist $j < l$ such that $m_{i_j}/x_n^{e_j} \mid m_{i_l}/x_n^{e_l}$. Since $e_j \leq e_l$, it must be that $m_{i_j} \mid m_{i_l}$, from which it follows that $i_j < i_l$, thus ending the proof. ■

Although there are multiple versions of the proof for Hilbert's Basis Theorem, the one used in this paper will build on Dickson's Lemma and use leading monomials.

Theorem 3.2 (Hilbert's Basis Theorem). The ring $\mathbb{C}[x_1, \ldots, x_n]$ is Noetherian. In other words, every ideal $I \in \mathbb{C}[x]$ is generated by a finite set.

Proof. Suppose that I were an ideal of $\mathbb{C}[x_1, x_2, \ldots, x_k]$. Note that we can order the monomials inside $\mathbb{C}[x_1, x_2, \ldots, x_k]$ lexicographically; we will be interested in the largest of these monomials. We then write $LM(f)$ (standing for leading monomial) for the largest monomial having non-zero coefficients in f, where f is any polynomial in $\mathbb{C}[x_1, x_2, \ldots, x_k]$ (lexicographically, this is the first term with the highest total degree). Since the leading monomial is a sequence of monomials in the variables of f, we are able to determine, using 3.1, that the set of minimal monomials in $\{\mathrm{LM}(f) : f \in I\}$ must be finite. It follows that there must exist finitely many polynomials $f_1, \ldots, f_k \in I$ such that for all $f \in I$, there exists some i for which $\mathrm{LM}(f_i) \mid \mathrm{LM}(f)$. Now, we shall show that the ideal $J = (f_1, \ldots, f_k)$ is generated by the f_i's is equal to I. If it is not, then we take some $f \in I \backslash J$ with the smallest leading monomial among all counter examples. Then, we pick i such that $\mathrm{LM}(f_i) \mid \mathrm{LM}(f)$, perhaps with $\mathrm{LM}(f) = m\,\mathrm{LM}(f_i)$. Then, we subtract the appropriate scalar multiple of mf_i contained in J from f. This then yields a polynomial with a smaller leading monomial which is still in $I \backslash J$. However, we have reached a contradiction, since we have disputed the minimality of $\mathrm{LM}(f)$, thus ending the proof.

■

Theorem 3.3 (Hilbert's Finiteness Theorem). The set $\mathbb{C}[x]^G := \{f \in \mathbb{C}[x] : gf = f\}$ is a finitely generated algebra. In other words, there exist functions $f_1, f_2, \ldots, f_k \in \mathbb{C}[x]^G$ such that every G-invariant polynomial in $\mathbb{C}[x]$ is a polynomial in the set of f_i's.

We may observe how the proof[4] to the above theorem may be approached: supposing that I were an ideal of $\mathbb{C}[x]$ generated by all the homogeneous non-constant functions from $\mathbb{C}[x]$, we may use the Hilbert Basis Theorem to show that I is generated by the finite set of homogeneous functions $f_1, f_2, \ldots, f_k$ in $\mathbb{C}[x]^G$. Then, it is plausible that we can induct to yield $\mathbb{C}[x]^G = \mathbb{C}[f_1, f_2, \ldots, f_k]$.

Hilbert's proofs of the basis and finiteness theorems using topics from invariant theory was a groundbreaking accomplishment. In response to a letter Hilbert sent to mathematician Paul Gordan containing these proofs, Gordan famously replied, "*Das ist nicht Mathematik. Das ist Theologie!*"

[4]We cannot construct a rigorous proof of Hilbert's finiteness theorem using only the materials provided in this paper, as the proof requires advanced topics from representation theory. The study of isotypic decompositions and Schur's lemma used the rigorous proof can be pursued using [4].

References

[1] G. Bool. Exposition of a general theory of linear transformations. *Cambridge Mathematical Journal*, 3, 1841.

[2] A. Cayley. On the theory of linear transformations. *Cambridge Mathematical Journal*, 4, 1845.

[3] R. C. L. David Hilbert. *Theory of Algebraic Invariants*. Cambridge Mathematical Library. Cambridge University Press, 1993.

[4] B. Steinberg. *Representation Theory of Finite Groups: An Introductory Approach*. Universitext. Springer, 1 edition, 2012.

[5] H. Weyl. Invariants. *Duke Mathematical Journal*, 5, 1939.