INVARIANT THEORY OF FINITE GROUPS

M. SKILLETER

Abstract. In this paper, we introduce invariants of finite matrix groups, developing the language and elementary results of invariant theory. We culminate with the statement and proof of the Chevalley-Todd-Shephard Theorem, which relates the geometry of the action on a finite-dimensional vector space $\mathbb{C}^n$ to the induced action on the polynomial ring in n variables. This gives an easy geometric condition for when the ring of invariants is a polynomial algebra. This condition is both necessary and sufficient, so gives a full characterisation of such actions.

Introduction

Polynomials and matrices abound in modern mathematics, and there is a natural action of a group of $n \times n$ matrices on the polynomial ring in n variables. We can then ask the question: what are the fixed points of this action? Invariant theory is the study of these fixed points.

Let $G \leq \mathrm{GL}_n(\mathbb{C})$ be a finite group of matrices. We can define an action of G on the polynomial ring $\mathbb{C}[\mathbf{x}] = \mathbb{C}[x_1, ..., x_n]$ as follows: for $g \in G$ and $f \in \mathbb{C}[\mathbf{x}]$, the polynomial $g \cdot f$ is given by

$$(g \cdot f)(\mathbf{x}) = f(g^{-1} \cdot \mathbf{x}).$$

A polynomial f is called an invariant if

$$g \cdot f = f$$

for every $g \in G$. The collection of such invariants forms a subring of $\mathbb{C}[\mathbf{x}]$, called the invariant ring and denoted $\mathbb{C}[\mathbf{x}]^G$. The following is the quintessential example in invariant theory.

Example 1. We can embed S_n into $\mathrm{GL}_n(\mathbb{C})$ as the set of $n \times n$ permutation matrices. Polynomials in $\mathbb{C}[\mathbf{x}]$ invariant under this action are called symmetric polynomials. The Fundamental Theorem of Symmetric Polynomials says that the elementary symmetric polynomials

$$e_j^n = \sum_{1 \leq r_1 < r_2 < ... < r_j \leq n} x_{r_1} x_{r_2} ... x_{r_j}$$

form a basis for the ring of symmetric polynomials. That is,

$$\mathbb{C}[\mathbf{x}]^{S_n} = \mathbb{C}[e_1^n, ..., e_n^n].$$

The calculation of $\mathbb{C}[\mathbf{x}]^G$ for various groups G is a classical problem in algebraic geometry. In this paper, we will develop the rudimentary tools in modern invariant theory, as well as give illustrative examples of their application.

It is worth noting that much of the theory generalizes to an arbitrary algebraically closed field of characteristic coprime to $|G|$, but for simplicity we will work over $\mathbb{C}$.

Generating Sets and Syzygies

Given the subring $\mathbb{C}[\mathbf{x}]^G$, we might ask if there is a finite generating set akin to a basis for a vector space. For G a finite group, the answer is always affirmative (this is Theorem 2), although Hilbert showed that there need not be a finite generating set for infinite G. We have already seen one example of a finite set of generators for the invariant ring, given by the elementary symmetric polynomials. In this case, the elementary symmetric polynomials form a basis for $\mathbb{C}[\mathbf{x}]^{S_n}$, in that expressions in terms of the generators are unique. However, uniqueness need not hold in general.

Example 2. Consider $\mathbb{Z}/2 = \left\{\pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\} \leq \mathrm{GL}_2(\mathbb{C})$. It is easy to see that a monomial in $\mathbb{C}[x, y]$ is fixed under $\mathbb{Z}/2$ if and only if its total degree is even, so

$$\mathbb{C}[x, y]^{\mathbb{Z}/2} = \mathbb{C}[x^2, xy, y^2].$$

□

Here, expressions in terms of the generators need not be unique. For example,

$$x^2y^2 = x^2 \cdot y^2 = (xy)^2. \qquad (\dagger)$$

This lack of uniqueness is caused by the existence of non-trivial algebraic relations between the generators (called *syzygies*). In the case of Example 2, it turns out that (†) generates all such relations; see the Appendix for Macaulay2 code to verify this. For this reason, syzygies can be thought of as a measure of the failure of uniqueness in $\mathbb{C}[\mathbf{x}]^G$.

In a parallel to linear algebra, invariant theory is concerned with answering the following questions:

(1) Can we find a finite generating set for $\mathbb{C}[\mathbf{x}]^G$? This is analagous to a spanning set in linear algebra.
(2) Given such a generating set, what are the syzygies? How linearly independent are our generators?

It is surprisingly difficult to give a uniform treatment of all groups, so much of invariant theory consists of ad hoc arguments to compute the invariant ring of specific groups. We will now see an example of the computation of an invariant ring, which will illustrate why it is difficult to address this problem in general.

Theorem 1. *Let $A_n \leq \mathrm{GL}_n(\mathbb{C})$ be the alternating group on n elements and let δ denote the Vandermonde polynomial*

$$\delta = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Then every $f \in \mathbb{C}[\mathbf{x}]^{A_n}$ can be written uniquely as

$$f = f_1 + f_2\delta$$

where f_1, f_2 are symmetric. Hence $\mathbb{C}[\mathbf{x}]^{A_n} = \mathbb{C}[\mathbf{x}]^{S_n} \oplus \delta\mathbb{C}[\mathbf{x}]^{S_n}$.

Furthermore, the defining syzygy between the generators $e_1^n, ..., e_n^n, \delta$ is given by expressing $\delta^2 \in \mathbb{C}[\mathbf{x}]^{S_n}$ in terms of the elementary symmetric polynomials.

Proof.

First some terminology. We say that a polynomial is *alternating* if it is fixed under A_n, but the action by any odd permutation multiplies it by -1. Over the course of this proof, we will show that the set of alternating polynomials is exactly $\delta\mathbb{C}[\mathbf{x}]^{S_n}$. With this in mind, we start by proving that any $f \in \mathbb{C}[\mathbf{x}]^{A_n}$ can be written as $f = f_1 + A$ for f_1 symmetric and A an alternating polynomial.

Define $f^* \in \mathbb{C}[\mathbf{x}]$ by

$$f^*(x_1, ..., x_n) = f(x_2, x_1, ..., x_n) = ((12) \cdot f)(x_1, ..., x_n).$$

Then $f + f^*$ is symmetric and $f - f^*$ is alternating, so

$$f = \frac{1}{2}(f + f^*) + \frac{1}{2}(f - f^*)$$

is the sum of a symmetric and an alternating polynomial, as claimed. We next show that δ divides any alternating polynomial A. This is because

$$A(x_2, x_1, x_3, ..., x_n) = -A(x_1, x_2, ..., x_n)$$

and so

$$A(x_1, x_1, x_3, ..., x_n) = 0.$$

Hence $(x_2 - x_1) \mid A$. Similar reasoning shows that $(x_j - x_i) \mid A$ for every $i < j$. Since the $x_j - x_i$ are coprime and $\mathbb{C}[\mathbf{x}]$ is a unique factorization domain, $\delta \mid A$. Finally, we will show that if $A = f_2\delta$ for A alternating then f_2 is symmetric. We need to show that f_2 is invariant under any permutation. For τ an odd permutation, we have

$$\begin{aligned} \tau(A) &= \tau(f_2\delta) \\ &= \tau(f_2)\tau(\delta) \\ -A &= \tau(f_2)(-\delta) \\ A &= \tau(f_2)\delta \end{aligned}$$

By unique factorization, we must have $\tau(f_2) = f_2$. An analagous argument for even permutations shows that f_2 is invariant under every permutation and hence is symmetric.

To show that such a decomposition is unique, suppose

$$f = f_1 + f_2\delta = f_1' + f_2'\delta$$

for $f_1, f_2, f_1', f_2' \in \mathbb{C}[\mathbf{x}]^{S_n}$. Then

$$f_1 - f_1' = f_2'\delta - f_2\delta = (f_2' - f_2)\delta.$$

The product of a symmetric polynomial with δ is alternating, so $f_1 - f_1'$ is both symmetric and alternating. The only polynomial with this property is the zero polynomial, so $f_1 = f_1'$. Similarly $f_2 = f_2'$.

Finally, we address the syzygies in $\mathbb{C}[\mathbf{x}]^{A_n}$. Note that δ^2 is symmetric, so we can express it as a polynomial in the elementary symmetric polynomials; write $\delta^2 = p(e_1^n, ..., e_n^n)$. Then given any other relation $q\delta = r_1e_1^n + ...r_ne_n^n$, the right-hand side is symmetric and hence the left-hand side must be too. This is only possible if an even power of δ divides $q\delta$, so our syzygy divides the relation.

□

We now turn to the problem of computing $\mathbb{C}[\mathbf{x}]^G$ for arbitrary finite G. Clearly our calculation of $\mathbb{C}[\mathbf{x}]^{A_n}$ does not generalize. However, by computing the syzygies in $\mathbb{C}[\mathbf{x}]^{A_n}$, we have also proven the weaker statement that there *exists some* syzygy, without reference to what that syzygy might be. It turns out that this new statement can easily be generalized.

Recall that for a $\mathbb{C}$-algebra K, the *transcendence degree* of K over $\mathbb{C}$, written $\mathrm{trdeg}_{\mathbb{C}}(K)$, is the maximum number of elements of K which satisfy no non-trivial $\mathbb{C}$-linear relations. Such elements are said to be algebraically independent over $\mathbb{C}$.

Proposition 1. *Let $G \leq \mathrm{GL}_n(\mathbb{C})$ be a finite matrix group. Then $\mathrm{trdeg}_{\mathbb{C}}(\mathbb{C}[\mathbf{x}]^G) = n$.*

Proof. It is immediate that $\mathrm{trdeg}_{\mathbb{C}}(\mathbb{C}[\mathbf{x}]) = n$. Observe that we have the following tower of extensions:

$$\begin{array}{c} \mathbb{C}[\mathbf{x}] \\ | \\ \mathbb{C}[\mathbf{x}]^G \\ | \\ \mathbb{C} \end{array}$$

In such towers, the transcendence degrees satisfies an additive formula. This gives

$$\mathrm{trdeg}_{\mathbb{C}}(\mathbb{C}[\mathbf{x}]) = \mathrm{trdeg}_{\mathbb{C}}(\mathbb{C}[\mathbf{x}]^G) + \mathrm{trdeg}_{\mathbb{C}[\mathbf{x}]^G}(\mathbb{C}[\mathbf{x}]).$$

Hence it suffices to show that $\mathrm{trdeg}_{\mathbb{C}[\mathbf{x}]^G}(\mathbb{C}[\mathbf{x}]) = 0$ i.e. that $\mathbb{C}[\mathbf{x}]$ is algebraic over $\mathbb{C}[\mathbf{x}]^G$. For each x_i, we will exhibit a polynomial $P_i \in \mathbb{C}[\mathbf{x}]^G[t]$ which has x_i as a root.

Consider the polynomial

$$P_i(t) := \prod_{g \in G} (t - g \cdot x_i).$$

Then x_i is a root of P_i because $g = I_n$ is always an element of the group, so $(t - x_i) \mid P_i$. Furthermore, if we expand P_i then we find that the coefficients "look like" the elementary symmetric polynomials, with the roles of the indeterminates replaced by the elements of G. Hence the coeffients lie in $\mathbb{C}[\mathbf{x}]^G$, so P_i is a polynomial in $\mathbb{C}[\mathbf{x}]^G[t]$ which has x_i as a root, showing that x_i is algebraic over $\mathbb{C}[\mathbf{x}]^G$.

□

Some caution is required when applying Proposition 1. While there are always n algebraically independent invariants, it is *not true* that these invariants need span. For example, it can be shown that the generating set for $\mathbb{C}[\mathbf{x}]^{A_n}$ given in Theorem 1 is minimal, in the sense that any generating set must be of size $n + 1$ or more.

However, Proposition 1 does have a useful application. If our generating set is of size greater than n, then there must be an algebraic relation between the generators; such a relation is precisely the definition of a syzygy. Moreover, any generating set must be of size at least n, because otherwise there is some invariant which is not in the span of the generators.

The Reynolds Operator

The proof of Proposition 1 suggests that "averaging" the action of G on $\mathbb{C}[\mathbf{x}]$ might provide a way to study invariants. With this in mind, we introduce the Reynolds operator. The definition and several immediate properties are as follows:

Definition. For $G \leq \mathrm{GL}_n(\mathbb{C})$ a finite matrix group, the Reynolds operator $\rho : \mathbb{C}[\mathbf{x}] \to \mathbb{C}[\mathbf{x}]$ is

$$\rho = \frac{1}{|G|} \sum_{g \in G} g$$

Proposition 2. *The map ρ has the following properties:*

(i) For $f \in \mathbb{C}[\mathbf{x}]^G$ and $h \in \mathbb{C}[\mathbf{x}]$, $\rho(fh) = f\rho(h)$. Also $\rho(1) = 1$.
(ii) For any $f \in \mathbb{C}[\mathbf{x}]$, $\rho(f) \in \mathbb{C}[\mathbf{x}]^G$. Hence we can think of ρ as a map $\mathbb{C}[\mathbf{x}] \to \mathbb{C}[\mathbf{x}]^G$.
(iii) $\rho \mid_{\mathbb{C}[\mathbf{x}]^G} = \mathrm{id}$. Hence ρ is surjective.
(iv) $\rho^2 = \rho$ i.e. ρ is idempotent.
(v) ρ is $\mathbb{C}$-linear.
(vi) $\mathrm{tr}(\rho) = \frac{1}{|G|} \sum_{g \in G} \mathrm{tr}(g)$

Proof. Properties (i) and (ii) can easily be verified via explicit computation, and property (iii) is a special case of (i) with $h = 1$. Property (iv) follows immediately from (ii) and (iii). Property (v) is because ρ is a sum of matrices and property (vi) is from linearity of the trace. □

This definition is deceptively simple, since ρ has many powerful applications. There is a more general definition of a Reynolds operator, which is a linear operator that satisfies property (i). As an example of the usefulness of the Reynolds operator, we can now prove the Hilbert Finiteness Theorem:

Theorem 2. *Let $G \leq \mathrm{GL}_n(\mathbb{C})$ be a finite matrix group. Then $\mathbb{C}[\mathbf{x}]^G$ is finitely generated.*

Proof. We first claim that it suffices to show that the homogeneous invariants are finitely generated. Given any $f \in \mathbb{C}[\mathbf{x}]$, we can write $f = \sum_\alpha c_\alpha x^\alpha$ as a finite sum of monomials. Then $\rho(f) = \sum_\alpha c_\alpha \rho(x^\alpha)$. Since $\rho(x^\alpha)$ is homogeneous, any invariant can be written as a finite sum of homogeneous invariants, proving our claim.

Let $I \subseteq \mathbb{C}[\mathbf{x}]$ be the **ideal** generated by all homogeneous invariants of positive degree. By Hilbert's Basis Theorem, $\mathbb{C}[\mathbf{x}]$ is Noetherian and so every ideal is finitely generated. Write

$$I = \langle f_1, ..., f_s \rangle.$$

By replacing each f_i by its (finitely many) homogeneous components if necessary, we may assume that the generators are themselves homogeneous invariants of positive degree. We claim that $f_1, ..., f_s$ generate the homogeneous invariants as a $\mathbb{C}$-algebra, which is stronger than generating the ideal I.

Let $f \in \mathbb{C}[\mathbf{x}]^G$ be a homogeneous invariant of degree d. We will prove that $f \in \mathbb{C}[f_1, ..., f_s]$ by induction on d. The base case is $d = 0$. Since $\mathbb{C}[f_1, ..., f_s]$ contains all constant functions, this is immediate.

Now suppose that $d > 0$ and that every homogeneous invariant of degree less than d is in $\mathbb{C}[f_1, ..., f_s]$. Since $f \in I$, we can write f as a linear combination

$$f = q_1 f_1 + ... + q_s f_s.$$

Applying the Reynolds operator and simplifying using Proposition 2 then gives

$$f = \rho(q_1)f_1 + ... + \rho(q_s)f_s.$$

The elements $\rho(q_1), ..., \rho(q_s)$ are invariants. Because the f_i and f are homogeneous of positive degree, each $\rho(q_i)$ must have degree strictly less than d. By the inductive hypothesis, each $\rho(q_i) \in \mathbb{C}[f_1, ..., f_s]$ (note that $\rho(q_i)$ might not be homogeneous, but the inductive hypothesis applies to its homogeneous components). Thus we have expressed f as a linear combination of $f_1, .., f_s$ with coefficients in $\mathbb{C}[f_1, ..., f_s]$. By induction, this finishes the proof.

□

This shows that there is always a finite generating set for $\mathbb{C}[\mathbf{x}]^G$ if G is finite. It is worth remarking that the only place we have used finiteness in this proof was to construct the Reynolds operator. More generally, Theorem 2 will hold so long as there is an operator to play the role of ρ.

In the case that G is finite, Emmy Noether proved a much stronger statement than Theorem 2 by giving a bound on the number of invariants needed to generate $\mathbb{C}[\mathbf{x}]^G$, as well as bounding the degree of each generator.

Theorem 3. *Let $G \leq \mathrm{GL}_n(\mathbb{C})$ be a finite matrix group. Then $\mathbb{C}[\mathbf{x}]^G$ is generated by at most $\binom{n+|G|}{n}$ homogeneous invariants, each of degree at most $|G|$.*

Proof.

From the proof of Theorem 2, $\mathbb{C}[\mathbf{x}]^G$ is generated by the $\rho(x^\alpha)$ ranging over all monomials x^α. We will show that we actually only need the $\rho(x^\alpha)$ for $|\alpha| \leq |G|$.

By the multinomial expansion theorem, we have

$$(x_1 + ... + x_n)^k = \sum_{|\alpha|=k} c_\alpha x^\alpha,$$

where $c_\alpha = \binom{k}{\alpha} = \frac{k!}{\alpha_1!...\alpha_n!}$ is the multinomial coefficient. Let $t_1, ..., t_n$ be a new set of indeterminates, with $\mathbf{t}$ the column vector of these variables. For sake of notation, let $\mathbf{tx}$ denote the polynomial $t_1x_1 + ... + t_nx_n$.

By the same reasoning as above, $(\mathbf{tx})^k = \sum_{|\alpha|=k} c_\alpha t^\alpha x^\alpha$. For $g \in G$, we then have

$$(\mathbf{t}(g \cdot \mathbf{x}))^k = \sum_{|\alpha|=k} c_\alpha t^\alpha (g \cdot x)^\alpha.$$

We now sum both sides of this expression over all $g \in G$ to get

$$\sum_{g \in G} (\mathbf{t}(g \cdot \mathbf{x}))^k = \sum_{|\alpha|=k} c_\alpha t^\alpha \left(\sum_{g \in G} (g \cdot x)^\alpha \right). \qquad (*)$$

The left-hand side of $(*)$ is the k^{th} power sum in the $|G|$ polynomials $\mathbf{t}(g \cdot \mathbf{x})$; write $p_k = p_k(\mathbf{t}(g \cdot \mathbf{x}) : g \in G)$ to denote this power sum. It is a fact that the power sums for $k > |G|$ are a polynomial in the first $|G|$ power sums (see Theorem 2.5 in [1]). Hence we have

$$p_k = F(p_1, ..., p_{|G|}) \qquad (\dagger)$$

for $k > |G|$ and $F \in \mathbb{C}[p_1, ..., p_{|G|}]$.

The right-hand side of $(*)$ simplifies to $\sum_{|\alpha|=k} |G| c_\alpha t^\alpha \rho(x^\alpha)$, which encodes $\rho(x^\alpha)$ as a coefficient of t^α (we introduced the variables $t_1, ..., t_n$ to prevent cancellation).

Substituting $(*)$ into $(\dagger)$ gives

$$\sum_{|\alpha|=k} |G|c_\alpha \rho(x^\alpha)t^\alpha = F\left(\sum_{|\beta|=1} |G|c_\beta\rho(x^\beta)t^\beta, ..., \sum_{|\beta|=|G|} |G|c_\beta\rho(x^\beta)t^\beta\right).$$

Expanding the right-hand side and equating coefficients in the t^α shows that $|G|c_\alpha\rho(x^\alpha)$ (for $|\alpha| > |G|$) is a polynomial in the $\rho(x^\beta)$ for $|\beta| \leq |G|$. Since $\mathbb{C}$ has characteristic zero, $|G|c_\alpha \neq 0$, and so $\rho(x^\alpha) \in \mathbb{C}[\rho(x^\beta) : |\beta| \leq |G|]$, as required.

□

In many cases, Noether's bound gives a much worse generating set than is necessary. For example, consider $\mathbb{C}[\mathbf{x}]^{A_n}$ when $n = 5$. Then Theorem 1 gives a generating set of size 6, where the generator of highest degree is δ with degree $\binom{5}{2} = 10$. In contrast, Theorem 3 says that we can find a generating set where each generator is of degree at most $|A_n| = \frac{5!}{2} = 60$ and there are at most $\binom{5+60}{5} = 8,259,888$ generators.

Interestingly, we have already seen an example showing that Noether's bound is best possible (at least the bound on the degree). Example 2 gave a generating set of $\mathbb{C}[x, y]^{\mathbb{Z}/2}$ which consists entirely of generators of degree 2. This example can be generalized considerably.

Example 2 (revisited). Let p be a prime number and $n \geq 2$. The cyclic group $\mathbb{Z}/p$ embeds into $\mathrm{GL}_n(\mathbb{C})$ as

$$\mathbb{Z}/p = \left\{ \begin{pmatrix} \zeta_p^k & & 0 \\ & \ddots & \\ 0 & & \zeta_p^k \end{pmatrix} : k = 0, 1, ..., p-1 \right\}$$

where $\zeta_p = e^{\frac{2\pi i}{p}}$ is the standard primitive p^{th} root of unity.

For any monomial x^α, we have

$$\rho(x^\alpha) = \frac{1}{p}\sum_{k=0}^{p-1} (\zeta_p^k x_1)^{\alpha_1} ... (\zeta_p^k x_n)^{\alpha_n} = \frac{1}{p}\sum_{k=0}^{p-1} \zeta_p^{k|\alpha|} x^\alpha = \frac{x^\alpha}{p}\sum_{k=0}^{p-1} \zeta_p^{k|\alpha|}.$$

If $\gcd(p, |\alpha|) = 1$ then $\zeta_p^{|\alpha|}$ is a primitive p^{th} root of unity, so $\sum_{k=0}^{p-1} \zeta_p^{k|\alpha|} = 0$. If $p \mid |\alpha|$ then $\zeta_p^{|\alpha|} = 1$, so $\rho(x^\alpha) = x^\alpha$. This shows that the monomials in $\mathbb{C}[\mathbf{x}]^{\mathbb{Z}/p}$ are exactly those which have total degree a multiple of p, so

$$\mathbb{C}[\mathbf{x}]^{\mathbb{Z}/p} = \mathbb{C}[x^\alpha : |\alpha| = p].$$

□

In fact, Barbara Schmid proved in [3] that these cyclic groups are exactly those for which Noether's bound is achieved. That is, a generator of degree $|G|$ is required if and only if G is one of the above groups. This discovery sparked an interest in sharpening Noether's bound, and this is still an active field. We direct the reader to [4] if they wish to learn more.

The Molien Series

We now know that $\mathbb{C}[\mathbf{x}]^G$ has finitely many generators, and the proof of Theorem 3 shows that we can find such generators by taking $\rho(x^\alpha)$ for x^α monomials. However, this method often gives many redundant generators. The purpose of this section is to use the graded structure on $\mathbb{C}[\mathbf{x}]^G$ to refine our search for generators.

Let $\mathbb{C}[\mathbf{x}]_i^G$ denote the vector space generated by the homogeneous invariants of degree i. We define the *Molien series* of G to be

$$M(G,t) := \sum_{i=0}^{\infty} \dim(\mathbb{C}[\mathbf{x}]_i^G) t^i.$$

The Molien series stores information about how many linearly independent invariants of G are in each degree, and hence can be a powerful tool for explicit computation. While this definition is not easy to use for calculations, the following theorem of Molien makes it more accessible:

Theorem 4. *Let $G \leq \mathrm{GL}_n(\mathbb{C})$ be a finite matrix group. Then*

$$M(G,t) = \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(I - tg)}.$$

Proof. Because the Reynolds operator $\rho : \mathbb{C}[\mathbf{x}] \to \mathbb{C}[\mathbf{x}]^G$ is $\mathbb{C}$-linear, it induces a map on each grade. Let $\rho_i : \mathbb{C}[\mathbf{x}]_i \to \mathbb{C}[\mathbf{x}]_i^G$ be the map on the i^{th} grade. One can verify that all of the properties from Proposition 2 hold for ρ_i; in particular, ρ_i is idempotent. A result from elementary linear algebra then says that $\mathrm{rank}(\rho_i) = \mathrm{tr}(\rho_i)$. Because ρ_i is surjective, $\dim(\mathbb{C}[\mathbf{x}]_i^G) = \mathrm{rank}(\rho_i) = \mathrm{tr}(\rho_i)$. Substituting this into the definition of the Molien series yields

$$\begin{aligned} M(G,t) &= \sum_{i=0}^{\infty} \dim(\mathbb{C}[\mathbf{x}]_i^G) t^i \\ &= \sum_{i=0}^{\infty} \mathrm{tr}(\rho_i) t^i \\ &= \sum_{i=0}^{\infty} \left(\frac{1}{|G|} \sum_{g \in G} \mathrm{tr}\left(g \mid_{\mathbb{C}[\mathbf{x}]_i}\right) \right) t^i \\ &= \frac{1}{|G|} \sum_{g \in G} \left(\sum_{i=0}^{\infty} \mathrm{tr}\left(g \mid_{\mathbb{C}[\mathbf{x}]_i}\right) t^i \right) \end{aligned}$$

To prove the theorem, it now suffices to show that

$$\sum_{i=0}^{\infty} \mathrm{tr}\left(g \mid_{\mathbb{C}[\mathbf{x}]_i}\right) t^i = \frac{1}{\det(I - tg)}$$

for each $g \in G$.

Since g has finite order and $\mathbb{C}$ is algebraically closed, g can be diagonalised. Let $f_1, ..., f_n \in \mathbb{C}[\mathbf{x}]_1$ be linearly independent eigenvectors, with eigenvalues $\lambda_1, ..., \lambda_n$.

A basis for $\mathbb{C}[\mathbf{x}]_i$ is given by all monomials (in the f_j) of degree i, so the eigenvalues of $g \mid_{\mathbb{C}[\mathbf{x}]_i}$ are precisely numbers of the form $\lambda_1^{\alpha_1} ... \lambda_n^{\alpha_n}$ for $\alpha_1 + ... + \alpha_n = i$.

The trace of a matrix is the sum of its eigenvalues, so

$$\operatorname{tr}\left(g\mid_{\mathbb{C}[\mathbf{x}]_i}\right) = \sum_{\alpha_1+\ldots+\alpha_n=i} \lambda_1^{\alpha_1}\ldots\lambda_n^{\alpha_n}.$$

We now recall the power series expansion for $\frac{1}{1-\lambda_j t}$, which is

$$\frac{1}{1-\lambda_j t} = \sum_{i=0}^{\infty} \lambda_j^i t^i.$$

Taking the product over all eigenvalues gives

$$\prod_{j=1}^{n} \frac{1}{1-\lambda_j t} = \sum_{i=0}^{\infty} \left(\sum_{\alpha_1+\ldots+\alpha_n=i} \lambda_1^{\alpha_1}\ldots\lambda_n^{\alpha_n} \right) t^i$$

Now observe that the $1-\lambda_j t$ are exactly the eigenvalues of $I-tg$. Since the determinant of a matrix is the product of its eigenvalues,

$$\prod_{j=1}^{n} \frac{1}{1-\lambda_j t} = \frac{1}{\det(I-tg)}.$$

Combining these results yields

$$\begin{aligned}
\sum_{i=0}^{\infty} \operatorname{tr}\left(g\mid_{\mathbb{C}[\mathbf{x}]_i}\right) t^i &= \sum_{i=0}^{\infty} \left(\sum_{\alpha_1+\ldots+\alpha_n=i} \lambda_1^{\alpha_1}\ldots\lambda_n^{\alpha_n} \right) t^i \\
&= \prod_{j=1}^{n} \frac{1}{1-\lambda_j t} \\
&= \frac{1}{\det(I-tg)}
\end{aligned}$$

as required. □

Remark. The value $\det(I-tg)$ is constant on conjugacy classes. This means that we need only pick representatives for each conjugacy class and weight the sum instead of computing $\det(I-tg)$ for every $g \in G$.

Example 3. Consider the quaternions $Q_8 \leq \mathrm{GL}_2(\mathbb{C})$ (the embedding is given in the table below). This has five conjugacy classes of sizes $1, 1, 2, 2$ and 2, represented by $I_2, -I_2, i, j$ and k respectively. We calculate the value $\det(I-tg)$ for each of these representatives:

g	$\det(I-tg)$
$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$(t-1)^2$
$-I_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$(t+1)^2$
$i = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$	t^2+1
$j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	t^2+1
$k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$	t^2+1

Thus

$$M(Q_8,t)=\frac{1}{8}\left(\frac{1}{(t-1)^2}+\frac{1}{(t+1)^2}+3\times\frac{2}{t^2+1}\right)=\frac{1+t^6}{(1-t^4)^2}=1+2t^4+t^6+...$$

which shows that $\mathbb{C}[x,y]^{Q_8}$ has 2 invariants of degree 4, 1 of degree 6 and so forth.

We will now use this to find the invariant ring $\mathbb{C}[x,y]^{Q_8}$. We are looking for 2 independent invariants of degree 4, so we compute the following values:

$$\rho(x^4)=\frac{1}{4}(x^4+y^4)=\rho(y^4)$$
$$\rho(x^2y^2)=x^2y^2$$

Hence $\alpha=x^4+y^4$ and $\beta=x^2y^2$ are two independent invariants of degree 4. For our degree 6 invariant, we calculate $\rho(x^3y^3)=x^5y-xy^5=:\gamma\in\mathbb{C}[x,y]^{Q_8}$. Using elimination ideals, we find that α,β and γ satisfy the relation $\gamma^2=\alpha^2\beta-4\beta^3$ and that this generates all relations between the invariants. Thus

$$\mathbb{C}[\alpha,\beta,\gamma]\cong\frac{\mathbb{C}[a,b,g]}{\langle g^2-a^2b+4b^3\rangle}$$

where $|a|=|b|=4$ and $|g|=6$. Using Macaulay2 (see Appendix), one can verify that the Hilbert series of $\mathbb{C}[\alpha,\beta,\gamma]$ is

$$HS(\mathbb{C}[\alpha,\beta,\gamma])=\frac{1+t^6}{(1-t^4)^2}=M(Q_8,t).$$

Therefore the rings $\mathbb{C}[x,y]^{Q_8}$ and $\mathbb{C}[\alpha,\beta,\gamma]$ have equal dimensions in each grade. Since $\mathbb{C}[\alpha,\beta,\gamma]\leq\mathbb{C}[x,y]^{Q_8}$, this shows that $\mathbb{C}[x,y]^{Q_8}=\mathbb{C}[\alpha,\beta,\gamma]$. □

The Chevalley-Todd-Shephard Theorem

We have now seen many examples of invariant rings, most of which have quite complicated structure. A natural question to ask is when $\mathbb{C}[\mathbf{x}]^G$ is particularly simple. To make this more precise, we introduce the following two definitions:

Definition. An element $s \in \mathrm{GL}_n(\mathbb{C})$ is called a **pseudoreflection** if it fixes pointwise a codimension 1 subspace of $\mathbb{C}^n$ and is not the identity transformation; equivalently, if $\ker(s - I)$ has codimension 1 in $\mathbb{C}^n$. A group generated by pseudoreflections is called a *reflection group.*

Definition. For $G \leq \mathrm{GL}_n(\mathbb{C})$, we say that $\mathbb{C}[\mathbf{x}]^G$ is a **polynomial algebra** if it is generated by n algebraically independent homogeneous elements, say $\mathbb{C}[\mathbf{x}]^G = \mathbb{C}[f_1, ..., f_n]$.

We have already seen an example of a polynomial algebra: the invariant ring $\mathbb{C}[\mathbf{x}]^{S_n}$, generated by the elementary symmetric polynomials. Interestingly, S_n is also generated by pseudoreflections (the transpositions) which suggests that there may be a relationship between these two concepts. This is the content of the Chevalley-Todd-Shephard (CTS) Theorem:

Theorem 5. *Let $G \leq \mathrm{GL}_n(\mathbb{C})$ be a finite matrix group. Then $\mathbb{C}[\mathbf{x}]^G$ is a polynomial algebra if and only if G is generated by pseudoreflections.*

Remark. Chevalley's original paper ([2]) only proved the forward implication. We will give a slightly modified proof, but the essential idea is the same.

Shephard and Todd's proof of the converse implication involved explicitly identifying all groups in $\mathrm{GL}_n(\mathbb{C})$ generated by pseudoreflections and computing each of their invariant rings. We will give a more elegant combinatorial argument (due to [5]).

Before we prove the CTS Theorem, we will need a few lemmas for use in the forward implication. The first of these is an elementary result about homogeneous polynomials, which we will state without proof. The second, however, relies on the fact that G is a reflection group.

Lemma 1. *Let $f_1, ..., f_s$ be homogeneous polynomials and let $H \in \mathbb{C}[f_1, ..., f_s]$ be a homogeneous relation. If $\frac{\partial H}{\partial f_1}$ is a linear combination of $\frac{\partial H}{\partial f_2}, ..., \frac{\partial H}{\partial f_s}$ then $\frac{\partial H}{\partial f_1} = 0$.*

Lemma 2. *If $f_1, f_2, ..., f_s$ are homogeneous invariants of a reflection group G such that $f_1 \notin \langle f_2, ..., f_s \rangle$, and there is some relation*

$$p_1 f_1 + ... + p_s f_s = 0$$

with the p_i homogeneous then p_1 is in the ideal generated by homogeneous invariants of positive degree.

Proof. For sake of notation, let I denote the ideal generated by homogeneous invariants of positive degree.

Let $g \in G$ be a pseudoreflection, so g fixes pointwise a codimension 1 subspace. Such a subspace is given by the vanishing locus of some polynomial f. For any polynomial $p \in \mathbb{C}[\mathbf{x}]$ and $\mathbf{x} \in \mathcal{V}(f)$, we then have:

$$(p - g \cdot p)(\mathbf{x}) = p(\mathbf{x}) - (g \cdot p)(\mathbf{x}) = p(\mathbf{x}) - p(g^{-1} \cdot \mathbf{x}) = p(\mathbf{x}) - p(\mathbf{x}) = 0.$$

Thus $p - g \cdot p \in \mathcal{I}(\mathcal{V}(f))$. Since $\mathbb{C}$ is algebraically closed, the Nullstellensatz tells us that $\mathcal{I}(\mathcal{V}(f)) = \sqrt{\langle f \rangle}$. In $\mathbb{C}[\mathbf{x}]$, the radical of a principal ideal is principal, so replacing f by the generator of the radical if necessary, we find that $f \mid p - g \cdot p$.

We thus get the relation

$$\frac{p_1 - g \cdot p_1}{f} f_1 + ... + \frac{p_1 - g \cdot p_1}{f} f_s = 0. \qquad (*)$$

We now argue by induction on the degree of p_1 that $p_1 - g \cdot p_1 \in I$. The base case is when p_1 is constant. If p_1 was a non-zero constant polynomial then we could rearrange our relation to write f_1 as a linear combination of $f_2, ..., f_s$. But by assumption $f_1 \notin \langle f_2, ..., f_s \rangle$, so it must be that $p_1 = 0$ and $p_1 - g \cdot p_1 = 0 \in I$.

For the inductive step, observe that $\frac{p_1 - g \cdot p_1}{f}$ has strictly smaller degree than $p_1 - g \cdot p_1$ and also satisfies the relation $(*)$, so by induction is in I. Thus $p_1 - g \cdot p_1 = \frac{p_1 - g \cdot p_1}{f} \cdot f \in I$.

Using this,

$$p_1 = \rho(p_1) + \frac{1}{|G|} \sum_{g \in G} (p_1 - g \cdot p_1)$$

is the sum of an invariant and an element of I. Again, p_1 is either 0 or homogeneous of positive degree, and in both cases $\rho(p_1) \in I$. We have now written p_1 as a sum of elements in I, so $p_1 \in I$.

□

With these lemmas in hand, we are now ready for the proof of the CTS Theorem.

Proof. Suppose first that G is a reflection group, and let $f_1, ..., f_s$ be a minimal set of generators for the ideal I of homogeneous invariants of positive degree. By the proof of Theorem 2, we know that $f_1, ..., f_s$ generate $\mathbb{C}[\mathbf{x}]^G$ as a $\mathbb{C}$-algebra. We will show that $f_1, ..., f_s$ are algebraically independent.

Suppose that $H(f_1, ..., f_s) = 0$ for some homogeneous relation $H \in \mathbb{C}[f_1, ..., f_s]$. Using the chain rule, we can relate the total derivative of H to the partial derivatives. For each $1 \leq i \leq n$, this gives us

$$\frac{\partial H}{\partial f_1} \frac{\partial f_1}{\partial x_i} + ... + \frac{\partial H}{\partial f_s} \frac{\partial f_s}{\partial x_i} = \frac{dH}{dx_i} = 0.$$

By reordering the f_j if necessary, we may assume that $\deg(f_1) \leq \deg(f_2) \leq ... \leq \deg(f_s)$. Then $\frac{\partial H}{\partial f_1}$ has degree at least the degrees of $\frac{\partial H}{\partial f_2}, ..., \frac{\partial H}{\partial f_s}$ so if $\frac{\partial H}{\partial f_1} = 0$ then all of the partials are zero, meaning H must be the zero polynomial and we are done.

Now suppose for sake of contradiction that $\frac{\partial H}{\partial f_1} \neq 0$. By the contrapositive of Lemma 1, $\frac{\partial H}{\partial f_1}$ is not a linear combination of $\frac{\partial H}{\partial f_2}, ..., \frac{\partial H}{\partial f_s}$ and hence $\frac{\partial H}{\partial f_1} \notin \langle \frac{\partial H}{\partial f_2}, ..., \frac{\partial H}{\partial f_s} \rangle$. By Lemma 2, the coefficient $\frac{\partial f_1}{\partial x_i}$ is in I. We now make use of Euler's Theorem, a formal identity expressing the homogeneous polynomial f_1 in terms of its partial derivatives. This says that

$$\deg(f_1) \cdot f_1 = \sum_{i=1}^{n} x_i \frac{\partial f_1}{\partial x_i}.$$

Each of the polynomials $\frac{\partial f_1}{\partial x_i}$ is of strictly smaller degree than f_1, so must be a linear combination of $f_2, ..., f_s$ without f_1. But this means that $f_1 \in \langle f_2, ..., f_s \rangle$, contradicting the fact that we chose $f_1, ..., f_s$ as a minimal generating set for I. Therefore it must be that $\frac{\partial H}{\partial f_1} = 0$ and hence $H = 0$ as before. Thus the $f_1, ..., f_s$ are algebraically independent, as claimed.

Because the $f_1, ..., f_s$ are generators for $\mathbb{C}[\mathbf{x}]^G$ and are algebraically independent, Proposition 1 says that $s = n$ and so $\mathbb{C}[\mathbf{x}]^G$ is a polynomial algebra.

Suppose now that $\mathbb{C}[\mathbf{x}]^G = \mathbb{C}[f_1, ..., f_n]$ is a polynomial algebra, with generators in degrees $d_1, ..., d_n$. Let $T \subseteq G$ be the set of all pseudoreflections in G, with $H \leq G$ the subgroup generated by T. We will show that $H = G$.

By Molien's Theorem, we have

$$M(G,t) = \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(I - tg)}.$$

We claim that the Laurent expansion about $t = 1$ is

$$M(G,t) = \frac{1}{|G|}(1-t)^{-n} + \frac{|T|}{2|G|}(1-t)^{-n+1} + \mathcal{O}((1-t)^{-n+2}).$$

First observe that the only element of G for which $\frac{1}{\det(I-tg)}$ has a pole of order n at 1 is $g = I_n$, so the coefficient of $(1-t)^{-n}$ is $\frac{1}{|G|}$. Next, the elements of G which give a pole of order $n-1$ at 1 are exactly the pseudoreflections. For $g \in T$, let $\lambda(g)$ be the eigenvalue of g which is not 1. The coefficient of $(1-t)^{-n+1}$ is then

$$\frac{1}{|G|} \sum_{g \in T} \frac{1}{1 - \lambda(g)}$$

because $\det(g)$ is the product of the eigenvalues of g, so $\lambda(g) = \det(g)$. An element of G is a pseudoreflection if and only if its inverse is, so we conclude that

$$2\sum_{g \in T} \frac{1}{1-\det(g)} = \sum_{g \in T} \frac{1}{1-\det(g)} + \sum_{g \in T} \frac{1}{1-\det(g^{-1})} = \sum_{g \in T} \left(\frac{1}{1-\det(g)} + \frac{1}{1-\det(g)^{-1}} \right).$$

For any invertible matrix, $\frac{1}{1-\det(g)} + \frac{1}{1-\det(g)^{-1}} = 1$ so the above reduces to $\sum_{g \in T} 1 = |T|$. Thus the coefficient of $(1-t)^{-n+1}$ is $\frac{|T|}{2|G|}$, as claimed.

We now find another expression for the Molien series. By the assumption that the f_i are homogeneous, we see that the homogeneous grades of $\mathbb{C}[\mathbf{x}]^G$ are precisely determined by the $d_1, ..., d_n$ i.e. $\dim(\mathbb{C}[\mathbf{x}]^G_i)$ is the number of possible ways to write i as a sum of $d_1, ..., d_n$. Thus

$$M(G,t) = \prod_{j=1}^{n} \frac{1}{1 - t^{d_j}}.$$

Expanding this in a Laurent expansion about $t = 1$, we find that

$$M(G,t) = \frac{1}{d_1 \cdot ... \cdot d_n}(1-t)^{-n} + \frac{d_1 + ... + d_n - n}{2d_1 \cdot ... \cdot d_n}(1-t)^{-n+1} + \mathcal{O}((1-t)^{-n+2}).$$

Equating coefficients of $(1-t)^{-n}$ in these two Laurent expansions shows that $|G| = d_1 \cdot ... \cdot d_n$. If we now equate coefficients on $(1-t)^{-n+1}$ and substitute in this formula for $|G|$, we find that $|T| = d_1 + ... + d_n - n = \sum_i (d_i - 1)$.

We are now almost finished. Since H is generated by pseudoreflections, the forward direction of the CTS Theorem says that $\mathbb{C}[\mathbf{x}]^H$ is a polynomial algebra: write $\mathbb{C}[\mathbf{x}]^H = \mathbb{C}[g_1, ..., g_n]$ where $\deg(g_i) = e_i$. Since $\mathbb{C}[\mathbf{x}]^G \leq \mathbb{C}[\mathbf{x}]^H$, the f_i can be written as

$$f_i = c_{i,1} g_1 + ... + c_{i,n} g_n.$$

Hence after reordering so that $d_1 \leq d_2 \leq ... \leq d_n$ and $e_1 \leq e_2 \leq ... \leq e_n$, we see that $d_i \geq e_i$ for all i. Repeating the same proof to compute the Laurent expansion for $M(H,t)$ shows that $|T| = \sum_i (e_i - 1)$ and so $\sum_i (e_i - 1) = \sum_i (d_i - 1)$. Thus $d_i = e_i$ and so

$$|G| = \prod_i d_i = \prod_i e_i = |H|.$$

Since G is finite, this shows that $G = H$.

□

Example 4. Consider the dihedral group $D_{2k} = \langle r, s \mid r^2 = s^k = (rs)^2 = 1 \rangle$. This embeds into $\mathrm{GL}_2(\mathbb{C})$ by

$$r \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad s \mapsto \begin{pmatrix} \zeta_k & 0 \\ 0 & \zeta_k^{-1} \end{pmatrix}$$

where $\zeta_k = e^{\frac{2pi}{k}}$ is the standard primitive k^{th} root of unity.

Written in this form, it is not obvious that D_{2k} is a reflection group. However, we can also take the elements r and rs as generators, which fix the hyperplanes $\mathcal{V}(y-x)$ and $\mathcal{V}(y - \zeta_k x)$ respectively. These elements are pseudoreflections and so D_{2k} is a reflection group. By the CTS Theorem, $\mathbb{C}[x,y]^{D_{2k}}$ is a polynomial algebra. More explicitly, one can show that $\mathbb{C}[x,y]^{D_{2k}} = \mathbb{C}[xy, x^k + y^k]$ by mimicking Example 3.

With the completion of the Chevalley-Todd-Shephard Theorem, we have classified all finite groups G such that $\mathbb{C}[\mathbf{x}]^G$ is a polynomial algebra. These are the "nicest" groups, the ones for which invariant theory is easiest to understand. The problem of computing invariants of non-reflection groups is still an open one, but the machinery developed in this paper should give the reader sufficient background to do their own further investigation into the subject.

References

[1] J. Kerma. *Rings of Invariants of Finite Groups*. Available at `www.math.iitb.ac.in/~jkv/invar.pdf`.

[2] C. Chevalley. *Invariants of Finite Groups Generated by Reflections*. Available at `https://pdfs.semanticscholar.org/b08d/5d8a14dd8ca8f5e3ed740b60c819f8d0f761.pdf`.

[3] Schmid B.J. (1991) *Finite groups and invariant theory*. In: Topics in Invariant Theory. Lecture Notes in Mathematics, vol 1478. Springer, Berlin, Heidelberg.

[4] M. Sezer. (2001) *Sharpening the generalized Noether bound in the invariant theory of finite groups*. Available at `https://www.sciencedirect.com/science/article/pii/S0021869302000182/pdf?md5=3f1649e1235ef3c453d2211db52b7ea6&pid=1-s2.0-S0021869302000182-main.pdf&_valck=1`.

[5] R. Stanley. (1979) *Invariants of finite groups and their applications to combinatorics*. Available at `https://projecteuclid.org/euclid.bams/1183544328`.

Appendix

Computing relations between generators for Example 2

```
i1 : R = CC[x, y, a, b, c]
i2 : I = ideal(a-x^2, b-x*y, c-y^2)
i3 : eliminate(I, {x,y})
o1= ideal(b^2-ac)
```

Computing relations between α, β, γ for Example 3

```
i1 : R = CC[x, y, a, b, g]
i2 : I = ideal(a-x^4-y^4, b-x^2*y^2, g-x^5*y+x*y^5)
i3 : eliminate(I, {x,y})
o1= ideal(g^2-a^2*b+4*b^3)
```

Computing Hilbert series for graded polynomial ring in Example 3

```
i1 : R = CC[a, b, g, Degrees => {4, 4, 6}] / ideal(g^2-a^2*b+4*b^3)
i2 : hilbertSeries R
o1= (1-T^12)/((1-T^6)(1-T^4)^2)
```


Australian National University, Mathematical Sciences Institute