7

Invariant Theory of Finite Groups

Invariant theory has had a profound effect on the development of algebraic geometry. For example, the Hilbert Basis Theorem and Hilbert Nullstellensatz, which play a central role in the earlier chapters in this book, were proved by Hilbert in the course of his investigations of invariant theory.

In this chapter, we will study the invariants of finite groups. The basic goal is to describe *all* polynomials which are unchanged when we change variables according to a given finite group of matrices. Our treatment will be elementary and by no means complete. In particular, we do not presume a prior knowledge of group theory.

§1 Symmetric Polynomials

Symmetric polynomials arise naturally when studying the roots of a polynomial. For example, consider the cubic $f = x^3 + bx^2 + cx + d$ and let its roots be $\alpha_1, \alpha_2, \alpha_3$. Then

$$x^3 + bx^2 + cx + d = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3).$$

If we expand the right-hand side, we obtain

$$x^3 + bx^2 + cx + d = x^3 - (\alpha_1 + \alpha_2 + \alpha_3)x^2 + (\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3)x - \alpha_1\alpha_2\alpha_3,$$

and thus,

$$\begin{aligned} b &= -(\alpha_1 + \alpha_2 + \alpha_3), \\ c &= \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3, \\ d &= -\alpha_1\alpha_2\alpha_3. \end{aligned} \tag{1}$$

This shows that the coefficients of f are polynomials in its roots. Further, since changing the order of the roots does not affect f, it follows that the polynomials expressing b, c, d in terms of $\alpha_1, \alpha_2, \alpha_3$ are unchanged if we permute $\alpha_1, \alpha_2, \alpha_3$. Such polynomials are said to be *symmetric*. The general concept is defined as follows.

Definition 1. *A polynomial* $f \in k[x_1, \ldots, x_n]$ *is* **symmetric** *if*

$$f(x_{i_1}, \ldots, x_{i_n}) = f(x_1, \ldots, x_n)$$

for every possible permutation $x_{i_1}, \ldots, x_{i_n}$ *of the variables* $x_1, \ldots, x_n$.

For example, if the variables are x, y, and z, then $x^2+y^2+z^2$ and xyz are obviously symmetric. The following symmetric polynomials will play an important role in our discussion.

Definition 2. *Given variables* $x_1, \ldots, x_n$, *we define the* **elementary symmetric functions** $\sigma_1, \ldots, \sigma_n \in k[x_1, \ldots, x_n]$ *by the formulas*

$$\begin{aligned}
\sigma_1 &= x_1 + \cdots + x_n, \\
&\vdots \\
\sigma_r &= \sum_{i_1 < i_2 < \cdots < i_r} x_{i_1} x_{i_2} \cdots x_{i_r}, \\
&\vdots \\
\sigma_n &= x_1 x_2 \cdots x_n.
\end{aligned}$$

Thus, σ_r is the sum of all monomials that are products of r distinct variables. In particular, every term of σ_r has total degree r. To see that these polynomials are indeed symmetric, we will generalize observation (1). Namely, introduce a new variable X and consider the polynomial

$$f(X) = (X - x_1)(X - x_2) \cdots (X - x_n) \tag{2}$$

with roots $x_1, \ldots, x_n$. If we expand the right-hand side, it is straightforward to show that

$$f(X) = X^n - \sigma_1 X^{n-1} + \sigma_2 X^{n-2} + \cdots + (-1)^{n-1}\sigma_{n-1} X + (-1)^n \sigma_n$$

(we leave the details of the proof as an exercise). Now suppose that we rearrange $x_1, \ldots, x_n$. This changes the order of the factors on the right-hand side of (2), but f itself will be unchanged. Thus, the coefficients $(-1)^r \sigma_r$ of f are symmetric functions.

One corollary is that for any polynomial with leading coefficient 1, the other coefficients are the elementary symmetric functions of its roots (up to a factor of ± 1). The exercises will explore some interesting consequences of this fact.

From the elementary symmetric functions, we can construct other symmetric functions by taking polynomials in $\sigma_1, \ldots, \sigma_n$. Thus, for example,

$$\sigma_2^2 - \sigma_1 \sigma_3 = x^2y^2 + x^2yz + x^2z^2 + xy^2z + xyz^2 + y^2z^2$$

is a symmetric polynomial. What is more surprising is that *all* symmetric polynomials can be represented in this way.

Theorem 3 (The Fundamental Theorem of Symmetric Polynomials). *Every symmetric polynomial in $k[x_1, \ldots, x_n]$ can be written uniquely as a polynomial in the elementary symmetric functions $\sigma_1, \ldots, \sigma_n$.*

Proof. We will use lex order with $x_1 > x_2 > \cdots > x_n$. Given a nonzero symmetric polynomial $f \in k[x_1, \ldots, x_n]$, let $\mathrm{LT}(f) = ax^\alpha$. If $\alpha = (\alpha_1, \ldots, \alpha_n)$, we first claim that $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$. To prove this, suppose that $\alpha_i < \alpha_{i+1}$ for some i. Let β be the exponent vector obtained from α by switching α_i and α_{i+1}. We will write this as $\beta = (\ldots, \alpha_{i+1}, \alpha_i, \ldots)$. Since ax^α is a term of f, it follows that ax^β is a term of $f(\ldots, x_{i+1}, x_i, \ldots)$. But f is symmetric, so that $f(\ldots, x_{i+1}, x_i, \ldots) = f$, and thus, ax^β is a term of f. This is impossible since $\beta > \alpha$ under lex order, and our claim is proved.

Now let

$$h = \sigma_1^{\alpha_1-\alpha_2}\sigma_2^{\alpha_2-\alpha_3}\cdots\sigma_{n-1}^{\alpha_{n-1}-\alpha_n}\sigma_n^{\alpha_n}.$$

To compute the leading term of h, first note that $\mathrm{LT}(\sigma_r) = x_1x_2\cdots x_r$ for $1 \leq r \leq n$. Hence,

$$\begin{aligned}\mathrm{LT}(h) &= \mathrm{LT}(\sigma_1^{\alpha_1-\alpha_2}\sigma_2^{\alpha_2-\alpha_3}\cdots\sigma_n^{\alpha_n})\\ &= \mathrm{LT}(\sigma_1)^{\alpha_1-\alpha_2}\mathrm{LT}(\sigma_2)^{\alpha_2-\alpha_3}\cdots\mathrm{LT}(\sigma_n)^{\alpha_n}\\ &= x_1^{\alpha_1-\alpha_2}(x_1x_2)^{\alpha_2-\alpha_3}\cdots(x_1\cdots x_n)^{\alpha_n}\\ &= x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n} = x^\alpha.\end{aligned} \tag{3}$$

It follows that f and ah have the same leading term, and thus,

$$\mathrm{multideg}(f - ah) < \mathrm{multideg}(f)$$

whenever $f - ah \neq 0$.

Now set $f_1 = f - ah$ and note that f_1 is symmetric since f and ah are. Hence, if $f_1 \neq 0$, we can repeat the above process to form $f_2 = f_1 - a_1h_1$, where a_1 is a constant and h_1 is a product of $\sigma_1, \ldots, \sigma_n$ to various powers. Further, we know that $\mathrm{LT}(f_2) < \mathrm{LT}(f_1)$ when $f_2 \neq 0$. Continuing in this way, we get a sequence of polynomials $f, f_1, f_2, \ldots$ with

$$\mathrm{multideg}(f) > \mathrm{multideg}(f_1) > \mathrm{multideg}(f_2) > \cdots.$$

Since lex order is a well-ordering, the sequence must be finite. But the only way the process terminates is when $f_{t+1} = 0$ for some t. Then it follows easily that

$$f = ah + a_1h_1 + \cdots + a_th_t,$$

which shows that f is a polynomial in the elementary symmetric functions. Finally, we need to prove uniqueness. Suppose that we have a symmetric polynomial f which can be written

$$f = g_1(\sigma_1, \ldots, \sigma_n) = g_2(\sigma_1, \ldots, \sigma_n).$$

Here, g_1 and g_2 are polynomials in n variables, say $y_1, \ldots, y_n$. We need to prove that $g_1 = g_2$ in $k[y_1, \ldots, y_n]$.

If we set $g = g_1 - g_2$, then $g(\sigma_1, \ldots, \sigma_n) = 0$ in $k[x_1, \ldots, x_n]$. Uniqueness will be proved if we can show that $g = 0$ in $k[y_1, \ldots, y_n]$. So suppose that $g \neq 0$. If we write $g = \sum_\beta a_\beta y^\beta$, then $g(\sigma_1, \ldots, \sigma_n)$ is a sum of the polynomials $g_\beta = a_\beta \sigma_1^{\beta_1} \sigma_2^{\beta_2} \cdots \sigma_n^{\beta_n}$, where $\beta = (\beta_1, \ldots, \beta_n)$. Furthermore, the argument used in (3) above shows that

$$\text{LT}(g_\beta) = a_\beta x_1^{\beta_1 + \cdots + \beta_n} x_2^{\beta_2 + \cdots + \beta_n} \cdots x_n^{\beta_n}.$$

It is an easy exercise to show that the map

$$(\beta_1, \ldots \beta_n) \mapsto (\beta_1 + \cdots + \beta_n, \beta_2 + \cdots + \beta_n, \ldots, \beta_n)$$

is injective. Thus, the g_β's have distinct leading terms. In particular, if we pick β such that $\text{LT}(g_\beta) > \text{LT}(g_\gamma)$ for all $\gamma \neq \beta$, then $\text{LT}(g_\beta)$ will be greater than *all* terms of the g_γ's. It follows that there is nothing to cancel $\text{LT}(g_\beta)$ and, thus, $g(\sigma_1, \ldots, \sigma_n)$ cannot be zero in $k[x_1, \ldots, x_n]$. This contradiction completes the proof of the theorem. □

The proof just given is due to Gauss, who needed the properties of symmetric polynomials for his second proof (dated 1816) of the fundamental theorem of algebra. Here is how Gauss states lex order: "Then among the two terms

$$Ma^\alpha b^\beta c^\gamma \cdots \quad \text{and} \quad Ma^{\alpha'} b^{\beta'} c^{\gamma'} \cdots$$

superior order is attributed to the first rather than the second, if

$$\text{either } \alpha > \alpha', \text{ or } \alpha = \alpha' \text{ and } \beta > \beta', \text{ or } \alpha = \alpha', \beta = \beta' \text{ and } \gamma > \gamma', \text{ or etc."}$$

[see p. 36 of GAUSS (1876)]. This is the earliest known explicit statement of lex order.

Note that the proof of Theorem 3 gives an algorithm for writing a symmetric polynomial in terms of $\sigma_1, \ldots, \sigma_n$. For an example of how this works, consider

$$f = x^3y + x^3z + xy^3 + xz^3 + y^3z + yz^3 \in k[x, y, z].$$

The leading term of f is $x^3y = \text{LT}(\sigma_1^2\sigma_2)$, which gives

$$f_1 = f - \sigma_1^2\sigma_2 = -2x^2y^2 - 5x^2yz - 2x^2z^2 - 5xy^2z - 5xyz^2 - 2y^2z^2.$$

The leading term is now $-2x^2y^2 = -2\text{LT}(\sigma_2^2)$, and thus,

$$f_2 = f - \sigma_1^2\sigma_2 + 2\sigma_2^2 = -x^2yz - xy^2z - xyz^2.$$

Then one easily sees that

$$f_3 = f - \sigma_1^2\sigma_2 + 2\sigma_2^2 + \sigma_1\sigma_3 = 0$$

and hence,

$$f = \sigma_1^2\sigma_2 - 2\sigma_2^2 - \sigma_1\sigma_3$$

is the unique expression of f in terms of the elementary symmetric polynomials.

Surprisingly, we do not need to write a general algorithm for expressing a symmetric polynomial in $\sigma_1, \ldots, \sigma_n$, for we can do this process using the division algorithm from Chapter 2. We can even use the division algorithm to check for symmetry. The precise method is as follows.

Proposition 4. *In the ring $k[x_1, \dots, x_n, y_1, \dots, y_n]$, fix a monomial order where any monomial involving one of $x_1, \dots, x_n$ is greater than all monomials in $k[y_1, \dots, y_n]$. Let G be a Groebner basis of the ideal $\langle \sigma_1 - y_1, \dots, \sigma_n - y_n \rangle \subset k[x_1, \dots, x_n, y_1, \dots, y_n]$. Given $f \in k[x_1, \dots, x_n]$, let $g = \overline{f}^G$ be the remainder of f on division by G. Then:*

(i) *f is symmetric if and only if $g \in k[y_1, \dots, y_n]$.*

(ii) *If f is symmetric, then $f = g(\sigma_1, \dots, \sigma_n)$ is the unique expression of f as a polynomial in the elementary symmetric polynomials $\sigma_1, \dots, \sigma_n$.*

Proof. As above, we have $f \in k[x_1, \dots, x_n]$, and $g \in k[x_1, \dots, x_n, y_1, \dots, y_n]$ is its remainder on division by $G = \{g_1, \dots, g_t\}$. This means that

$$f = A_1 g_1 + \cdots + A_t g_t + g,$$

where $A_1, \dots, A_t \in k[x_1, \dots, x_n, y_1, \dots, y_n]$. We can assume that $g_i \neq 0$ for all i.

To prove (i), first suppose that $g \in k[y_1, \dots, y_n]$. Then for each i, substitute σ_i for y_i in the above formula for f. This will not affect f since it involves only $x_1, \dots, x_n$. The crucial observation is that under this substitution, every polynomial in $\langle \sigma_1 - y_1, \dots, \sigma_n - y_n \rangle$ goes to zero. Since $g_1, \dots, g_t$ lie in this ideal, it follows that

$$f = g(\sigma_1, \dots, \sigma_n).$$

Hence, f is symmetric.

Conversely, suppose that $f \in k[x_1, \dots, x_n]$ is symmetric. Then $f = g(\sigma_1, \dots, \sigma_n)$ for some $g \in k[y_1, \dots, y_n]$. We want to show that g is the remainder of f on division by G. To prove this, first note that in $k[x_1, \dots, x_n, y_1, \dots, y_n]$, a monomial in $\sigma_1, \dots, \sigma_n$ can be written as follows:

$$\begin{aligned} \sigma_1^{\alpha_1} \cdots \sigma_n^{\alpha_n} &= (y_1 + (\sigma_1 - y_1))^{\alpha_1} \cdots (y_n + (\sigma_n - y_n))^{\alpha_n} \\ &= y_1^{\alpha_1} \cdots y_n^{\alpha_n} + B_1 \cdot (\sigma_1 - y_1) + \cdots + B_n \cdot (\sigma_n - y_n) \end{aligned}$$

for some $B_1, \dots, B_n \in k[x_1, \dots, x_n, y_1, \dots, y_n]$. Multiplying by an appropriate constant and adding over the exponents appearing in g, it follows that

$$g(\sigma_1, \dots, \sigma_n) = g(y_1, \dots, y_n) + C_1 \cdot (\sigma_1 - y_1) + \cdots + C_n \cdot (\sigma_n - y_n),$$

where $C_1, \dots, C_n \in k[x_1, \dots, x_n, y_1, \dots, y_n]$. Since $f = g(\sigma_1, \dots, \sigma_n)$, we can write this as

$$f = C_1 \cdot (\sigma_1 - y_1) + \cdots + C_n \cdot (\sigma_n - y_n) + g(y_1, \dots, y_n). \tag{4}$$

We want to show that g is the remainder of f on division by G.

The first step is to show that no term of g is divisible by an element of $\mathrm{LT}(G)$. If this were false, then there would be $g_i \in G$, where $\mathrm{LT}(g_i)$ divides some term of g. Hence, $\mathrm{LT}(g_i)$ would involve only $y_1, \dots, y_n$ since $g \in k[y_1, \dots, y_n]$. By our hypothesis on the ordering, it would follow that $g_i \in k[y_1, \dots, y_n]$. Now replace every y_i with the corresponding σ_i. Since $g_i \in \langle \sigma_1 - y_1, \dots, \sigma_n - y_n \rangle$, we have already observed that g_i goes to zero under the substitution $y_i \mapsto \sigma_i$. Then $g_i \in k[y_1, \dots, y_n]$ would mean $g_i(\sigma_1, \dots, \sigma_n) = 0$. By the uniqueness part of Theorem 3, this would imply $g_i = 0$, which is impossible since $g_i \neq 0$. This proves our claim. It follows that in (4), no term

of g is divisible by an element of $\mathrm{LT}(G)$, and since G is a Groebner basis, Proposition 1 of Chapter 2, §6 tells us that g is the remainder of f on division by G. This proves that the remainder lies in $k[y_1, \ldots, y_n]$ when f is symmetric.

Part (ii) of the proposition follows immediately from the above arguments and we are done. □

A seeming drawback to the above proposition is the necessity to compute a Groebner basis for $\langle \sigma_1 - y_1, \ldots, \sigma_n - y_n \rangle$. However, when we use lex order, it is quite simple to write down a Groebner basis for this ideal. We first need some notation. Given variables $u_1, \ldots, u_s$, let

$$h_i(u_1, \ldots, u_s) = \sum_{|\alpha|=i} u^\alpha$$

be the sum of *all* monomials of total degree i in $u_1, \ldots, u_s$. Then we get the following Groebner basis.

Proposition 5. *Fix lex order on $k[x_1, \ldots, x_n, y_1, \ldots, y_n]$ with $x_1 > \cdots > x_n > y_1 > \cdots > y_n$. Then the polynomials*

$$g_k = h_k(x_k, \ldots, x_n) + \sum_{i=1}^{k} (-1)^i h_{k-i}(x_k, \ldots, x_n) y_i, \quad k = 1, \ldots, n,$$

form a Groebner basis for the ideal $\langle \sigma_1 - y_1, \ldots, \sigma_n - y_n \rangle$.

Proof. We will sketch the proof, leaving most of the details for the exercises. The first step is to note the polynomial identity

$$0 = h_k(x_k, \ldots, x_n) + \sum_{i=1}^{k} (-1)^i h_{k-i}(x_k, \ldots, x_n) \sigma_i. \tag{5}$$

The proof will be covered in Exercises 10 and 11.

The next step is to show that $g_1, \ldots, g_n$ form a basis of $\langle \sigma_1 - y_1, \ldots, \sigma_n - y_n \rangle$. If we subtract the identity (5) from the definition of g_k, we obtain

$$g_k = \sum_{i=1}^{k} (-1)^i h_{k-i}(x_k, \ldots, x_n)(y_i - \sigma_i), \tag{6}$$

which proves that $\langle g_1, \ldots, g_n \rangle \subset \langle \sigma_1 - y_1, \ldots, \sigma_n - y_n \rangle$. To prove the opposite inclusion, note that since $h_0 = 1$, we can write (6) as

$$g_k = (-1)^k (y_k - \sigma_k) + \sum_{i=1}^{k-1} (-1)^i h_{k-i}(x_k, \ldots, x_n)(y_i - \sigma_i). \tag{7}$$

Then induction on k shows that $\langle \sigma_1 - y_1, \ldots, \sigma_n - y_n \rangle \subset \langle g_1, \ldots, g_n \rangle$ (see Exercise 12).

Finally, we need to show that we have a Groebner basis. In Exercise 12, we will ask you to prove that

$$\mathrm{LT}(g_k) = x_k^k.$$

This is where we use lex order with $x_1 > \cdots > x_n > y_1 > \cdots > y_n$. Thus the leading terms of $g_1, \ldots, g_k$ are relatively prime, and using the theory developed in §9 of Chapter 2, it is easy to show that we have a Groebner basis (see Exercise 12 for the details). This completes the proof. □

In dealing with symmetric polynomials, it is often convenient to work with ones that are *homogeneous.* Here is the definition.

Definition 6. *A polynomial $f \in k[x_1, \ldots, x_n]$ is* **homogeneous of total degree** k *provided that every term appearing in f has total degree k.*

As an example, note that the i-th elementary symmetric function σ_i is homogeneous of total degree i. An important fact is that every polynomial can be written uniquely as a sum of homogeneous polynomials. Namely, given $f \in k[x_1, \ldots, x_n]$, let f_k be the sum of all terms of f of total degree k. Then each f_k is homogeneous and $f = \sum_k f_k$. We call f_k the k-th *homogeneous component of* f.

We can understand symmetric polynomials in terms of their homogeneous components as follows.

Proposition 7. *A polynomial $f \in k[x_1, \ldots, x_n]$ is symmetric if and only if all of its homogeneous components are symmetric.*

Proof. Given a symmetric polynomial f, let $x_{i_1}, \ldots, x_{i_n}$ be a permutation of $x_1, \ldots, x_n$. This permutation takes a term of f of total degree k to one of the same total degree. Since $f(x_{i_1}, \ldots, x_{i_n}) = f(x_1, \ldots, x_n)$, it follows that the k-th homogeneous component must also be symmetric. The converse is trivial and the proposition is proved. □

Proposition 7 tells us that when working with a symmetric polynomial, we can assume that it is homogeneous. In the exercises, we will explore what this implies about how the polynomial is expressed in terms of $\sigma_1, \ldots, \sigma_n$.

The final topic we will explore is a different way of writing symmetric polynomials. Specifically, we will consider the *power sums*

$$s_k = x_1^k + x_2^k + \cdots + x_n^k.$$

Note that s_k is symmetric. Then we can write an arbitrary symmetric polynomial in terms of $s_1, \ldots, s_n$ as follows.

Theorem 8. *If k is a field containing the rational numbers $\mathbb{Q}$, then every symmetric polynomial in $k[x_1, \ldots, x_n]$ can be written as a polynomial in the power sums $s_1, \ldots, s_n$.*

Proof. Since every symmetric polynomial is a polynomial in the elementary symmetric functions (by Theorem 3), it suffices to prove that $\sigma_1, \ldots, \sigma_n$ are polynomials in $s_1, \ldots, s_n$. For this purpose, we will use the *Newton identities*, which state that

$$\begin{aligned} &s_k - \sigma_1 s_{k-1} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1 + (-1)^k k\sigma_k = 0, \quad 1 \le k \le n, \\ &s_k - \sigma_1 s_{k-1} + \cdots + (-1)^{n-1}\sigma_{n-1}s_{k-n+1} + (-1)^n \sigma_n s_{k-n} = 0, \quad k > n. \end{aligned}$$

The proof of these identities will be given in the exercises.

We now prove by induction on k that σ_k is a polynomial in $s_1, \ldots, s_n$. This is true for $k = 1$ since $\sigma_1 = s_1$. If the claim is true for $1, 2, \ldots, k-1$, then the Newton identities imply that

$$\sigma_k = (-1)^{k-1}\frac{1}{k}(s_k - \sigma_1 s_{k-1} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1).$$

We can divide by the integer k because $\mathbb{Q}$ is contained in the coefficient field (see Exercise 16 for an example of what can go wrong when $\mathbb{Q} \not\subset k$). Then our inductive assumption and the above equation show that σ_k is a polynomial in $s_1, \ldots, s_n$. □

As a consequence of Theorems 3 and 8, every elementary symmetric function can be written in terms of power sums, and vice versa. For example,

$$\begin{aligned} s_2 = \sigma_1^2 - 2\sigma_2 &\longleftrightarrow \sigma_2 = \frac{1}{2}(s_1^2 - s_2), \\ s_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3 &\longleftrightarrow \sigma_3 = \frac{1}{6}(s_1^3 - 3s_1s_2 + 2s_3). \end{aligned}$$

Power sums will be unexpectedly useful in §3 when we give an algorithm for finding the invariant polynomials for a finite group.

EXERCISES FOR §1

1. Prove that $f \in k[x, y, z]$ is symmetric if and only if $f(x, y, z) = f(y, x, z) = f(y, z, x)$.
2. (Requires abstract algebra) Prove that $f \in k[x_1, \ldots, x_n]$ is symmetric if and only if $f(x_1, x_2, x_3, \ldots, x_n) = f(x_2, x_1, x_3, \ldots, x_n) = f(x_2, x_3, \ldots x_n, x_1)$. Hint: Show that the cyclic permutations (1, 2) and $(1, 2, \ldots, n)$ generate the symmetric group S_n. See Exercise 11 in §2.10 of HERSTEIN (1975).
3. Let σ_i^n be the i-th elementary symmetric function in variables $x_1, \ldots, x_n$. The superscript n denotes the number of variables and is not an exponent. We also set $\sigma_0^n = 1$ and $\sigma_i^n = 0$ if $i < 0$ or $i > n$. Prove that $\sigma_i^n = \sigma_i^{n-1} + x_n\sigma_{i-1}^{n-1}$ for all $n > 1$ and all i. This identity is useful in induction arguments involving elementary symmetric functions.
4. As in (2), let $f(X) = (X - x_1)(X - x_2)\cdots(X - x_n)$. Prove that $f = X^n - \sigma_1 X^{n-1} + \sigma_2 X^{n-2} + \cdots + (-1)^{n-1}\sigma_{n-1}X + (-1)^n\sigma_n$. Hint: You can give an induction proof using the identities of Exercise 3.
5. Consider the polynomial
$$f = (x^2 + y^2)(x^2 + z^2)(y^2 + z^2) \in k[x, y, z].$$
   a. Use the method given in the proof of Theorem 3 to write f as a polynomial in the elementary symmetric functions $\sigma_1, \sigma_2, \sigma_3$.
   b. Use the method described in Proposition 4 to write f in terms of $\sigma_1, \sigma_2, \sigma_3$.

You can use a computer algebra system for both parts of the exercise. Note that by stripping off the coefficients of powers of X in the polynomial $(X-x)(X-y)(X-z)$, you can get the computer to generate the elementary symmetric functions.

6. If the variables are $x_1, \ldots, x_n$, show that $\sum_{i \neq j} x_i^2 x_j = \sigma_1\sigma_2 - 3\sigma_3$. Hint: If you get stuck, see Exercise 13. Note that a computer algebra system cannot help here!
7. Let $f = x^n + a_1 x^{n-1} + \cdots + a_n \in k[x]$ have roots $\alpha_1, \ldots, \alpha_n$, which lie in some bigger field K containing k.
   a. Prove that any symmetric polynomial $g(\alpha_1, \ldots, \alpha_n)$ in the roots of f can be expressed as a polynomial in the coefficients $a_1, \ldots, a_n$ of f.
   b. In particular, if the symmetric polynomial g has coefficients in k, conclude that $g(\alpha_1, \ldots, \alpha_n) \in k$.
8. As in Exercise 7, let $f = x^n + a_1 x^{n-1} + \cdots + a_n \in k[x]$ have roots $\alpha_1, \ldots, \alpha_n$, which lie in some bigger field K containing k. The *discriminant* of f is defined to be
$$D(f) = \prod_{i \neq j} (\alpha_i - \alpha_j)$$
   a. Use Exercise 7 to show that $D(f)$ is a polynomial in $a_1, \ldots, a_n$.
   b. When $n = 2$, express $D(f)$ in terms of a_1 and a_2. Does your result look familiar?
   c. When $n = 3$, express $D(f)$ in terms of a_1, a_2, a_3.
   d. Explain why a cubic polynomial $x^3 + a_1x^2 + a_2x + a_3$ has a multiple root if and only if $-4a_1^3a_3 + a_1^2a_2^2 + 18a_1a_2a_3 - 4a_2^3 - 27a_3^2 = 0$.
9. Given a cubic polynomial $f = x^3 + a_1x^2 + a_2x + a_3$, what condition must the coefficients of f satisfy in order for one of its roots to be the average of the other two? Hint: If α_1 is the average of the other two, then $2\alpha_1 - \alpha_2 - \alpha_3 = 0$. But it could happen that α_2 or α_3 is the average of the other two. Hence, you get a condition stating that the product of three expressions similar to $2\alpha_1 - \alpha_2 - \alpha_3$ is equal to zero. Now use Exercise 7.
10. As in Proposition 5, let $h_i(x_1, \ldots, x_n)$ be the sum of all monomials of total degree i in $x_1, \ldots, x_n$. Also, let $\sigma_0 = 1$ and $\sigma_i = 0$ if $i > n$. The goal of this exercise is to show that
$$0 = \sum_{i=0}^{k} (-1)^i h_{k-i}(x_1, \ldots, x_n)\sigma_i(x_1, \ldots, x_n).$$
    In Exercise 11, we will use this to prove the closely related identity (5) that appears in the text. To prove the above identity, we will compute the coefficients of the monomials x^α that appear in $h_{k-i}\sigma_i$. Since every term in $h_{k-i}\sigma_i$ has total degree k, we can assume that x^α has total degree k. We will let a denote the number of variables that actually appear in x^α.
    a. If x^α appears in $h_{k-i}\sigma_i$, show that $i \leq a$. Hint: How many variables appear in each term of σ_i?
    b. If $i \leq a$, show that exactly $\binom{a}{i}$ terms of σ_i involve only variables that appear in x^α. Note that all of these terms have total degree i.
    c. If $i \leq a$, show that x^α appears in $h_{k-i}\sigma_i$ with coefficient $\binom{a}{i}$. Hint: This follows from part b because h_{k-i} is the sum of all monomials of total degree $k-i$, and each monomial has coefficient 1.
    d. Conclude that the coefficient of x^α in $\sum_{i=0}^{k}(-1)^i h_{k-i}\sigma_i$ is $\sum_{i=0}^{a}(-1)^i\binom{a}{i}$. Then use the binomial theorem to show that the coefficient of x^α is zero. This will complete the proof of our identity.
11. In this exercise, we will prove the identity
$$0 = h_k(x_k, \ldots, x_n) + \sum_{i=1}^{k} (-1)^i h_{k-i}(x_k, \ldots, x_n)\sigma_i(x_1, \ldots, x_n)$$

used in the proof of Proposition 5. As in Exercise 10, let $\sigma_0 = 1$, so that the identity can be written more compactly as

$$0 = \sum_{i=0}^{k} (-1)^i h_{k-i}(x_k, \ldots, x_n)\sigma_i(x_1, \ldots, x_n)$$

The idea is to separate out the variables $x_1, \ldots, x_{k-1}$. To this end, if $S \subset \{1, \ldots, k-1\}$, let x^S be the product of the corresponding variables and let $|S|$ denote the number of elements in S.

a. Prove that

$$\sigma_i(x_1, \ldots, x_n) = \sum_{S \subset \{1,\ldots,k-1\}} x^S \sigma_{i-|S|}(x_k, \ldots, x_n),$$

where we set $\sigma_j = 0$ if $j < 0$.

b. Prove that

$$\sum_{i=0}^{k} (-1)^i h_{k-i}(x_k, \ldots, x_n)\sigma_i(x_1, \ldots, x_n)$$
$$= \sum_{S \subset \{1,\ldots,k-1\}} x^S \left(\sum_{i=|S|}^{k} (-1)^i h_{k-i}(x_k, \ldots, x_n)\sigma_{i-|S|}(x_k, \ldots, x_n) \right).$$

c. Use Exercise 10 to conclude that the sum inside the parentheses is zero for every S. This proves the desired identity. Hint: Let $j = i - |S|$.

12. This exercise is concerned with the proof of Proposition 5. Let g_k be as defined in the statement of the proposition.
    a. Use equation (7) to prove that $\langle \sigma_1 - y_1, \ldots, \sigma_n - y_n \rangle \subset \langle g_1, \ldots, g_n \rangle$.
    b. Prove that $\mathrm{LT}(g_k) = x_k^k$.
    c. Combine part (b) with the results from §9 of Chapter 2 (especially Theorem 3 and Proposition 4 of that section) to prove that $g_1, \ldots, g_n$ form a Groebner basis.
13. Let f be a homogeneous symmetric polynomial of total degree k.
    a. Show that f can be written as a linear combination (with coefficients in k) of polynomials of the form $\sigma_1^{i_1}\sigma_2^{i_2}\cdots\sigma_n^{i_n}$ where $k = i_1 + 2i_2 + \cdots + ni_n$.
    b. Let m be the maximum degree of x_1 that appears in f. By symmetry, m is the maximum degree in f of any variable. If $\sigma_1^{i_1}\sigma_2^{i_2}\cdots\sigma_n^{i_n}$ appears in the expression of f from part (a), then prove that $i_1 + i_2 + \cdots + i_n \le m$.
    c. Show that the symmetric polynomial $\sum_{i \ne j} x_i^2 x_j$ can be written as $a\sigma_1\sigma_2 + b\sigma_3$ for some constants a and b. Then determine a and b. Compare this to what you did in Exercise 6.
14. In this exercise, you will prove the Newton identities used in the proof of Theorem 8. Let the variables be $x_1, \ldots, x_n$.
    a. As in Exercise 3, set $\sigma_0 = 1$ and $\sigma_i = 0$ if either $i < 0$ or $i > n$. Then show that the Newton identities are equivalent to

$$s_k - \sigma_1 s_{k-1} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1 + (-1)^k k\sigma_k = 0$$

for *all* $k \ge 1$.

    b. Prove the identity of part (a) by induction on n. Hint: Write σ_i as σ_i^n, where the exponent denotes the number of variables, and similarly write s_k as s_k^n. Use Exercise 3, and note that $s_k^n = s_k^{n-1} + x_n^k$.

15. This exercise will use the identity (5) to prove the following *nonsymmetric Newton identities:*

$$x_i^k - \sigma_1 x_i^{k-1} + \cdots + (-1)^{k-1}\sigma_{k-1}x_i + (-1)^k\sigma_k = (-1)^k\hat{\sigma}_k^i, \quad 1 \le k < n,$$
$$x_i^k - \sigma_1 x_i^{k-1} + \cdots + (-1)^{n-1}\sigma_{n-1}x_i^{k-n+1} + (-1)^n\sigma_n x_i^{k-n} = 0, \quad k \ge n,$$

where $\hat{\sigma}_k^i = \sigma_k(x_i, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ is the k-th elementary symmetric function of all variables except x_i. We will then give a second proof of the Newton identities.

a. Show that the nonsymmetric Newton identity for $k = n$ follows from (5). Then prove that this implies the nonsymmetric Newton identities for $k \ge n$. Hint: Treat the case $i = n$ first.
b. Show that the nonsymmetric Newton identity for $k = n - 1$ follows from the one for $k = n$. Hint: $\sigma_n = x_i\hat{\sigma}_{k-1}^i$.
c. Prove the nonsymmetric Newton identity for $k < n$ by decreasing induction on k. Hint: By Exercise 3, $\sigma_k = \hat{\sigma}_k^i + x_i\hat{\sigma}_{k-1}^i$.
d. Prove that $\sum_{i=1}^n \hat{\sigma}_k^i = (n-k)\hat{\sigma}_k$. Hint: A term $x_{i_1}\cdots x_{i_k}$, where $1 \le i_1 < \cdots < i_k \le n$, appears in how many of the $\hat{\sigma}_k^i$'s?
e. Prove the Newton identities.

16. Consider the field $\mathbb{F}_2 = \{0, 1\}$ consisting of two elements. Show that it is impossible to express the symmetric polynomial $xy \in \mathbb{F}_2[x, y]$ as a polynomial in s_1 and s_2 with coefficients $\mathbb{F}_2$. Hint: Show that $s_2 = s_1^2$!
17. Express s_4 as a polynomial in $\sigma_1, \ldots, \sigma_4$ and express σ_4 as a polynomial in $s_1 \ldots, s_4$.
18. We can use the division algorithm to automate the process of writing a polynomial $g(\sigma_1, \ldots, \sigma_n)$ in terms of $s_1, \ldots, s_n$. Namely, regard $\sigma_1, \ldots, \sigma_n, s_1, \ldots, s_n$ as variables and consider the polynomials

$$g_k = s_k = \sigma_1 s_{k-1} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1 + (-1)^k k\sigma_k, \quad 1 \le k \le n.$$

Show that if we use the correct lex order, the remainder of $g(\sigma_1, \ldots, \sigma_n)$ on division by $g_1, \ldots, g_n$ will be a polynomial $(s_1, \ldots, s_n)$ such that $g(\sigma_1, \ldots, \sigma_n) = h(s_1, \ldots, s_n)$. Hint: The lex order you need is *not* $\sigma_1 > \sigma_2 > \cdots > \sigma_n > s_1 > \cdots > s_n$.

§2 Finite Matrix Groups and Rings of Invariants

In this section, we will give some basic definitions for invariants of finite matrix groups and we will compute some examples to illustrate what questions the general theory should address. For the rest of this chapter, we will always assume that our field k contains the rational numbers $\mathbb{Q}$. Such fields are said to be of *characteristic zero.*

Definition 1. *Let* $\mathrm{GL}(n, k)$ *be the set of all invertible* $n \times n$ *matrices with entries in the field* k.

If A and B are invertible $n \times n$ matrices, then linear algebra implies that the product AB and inverse A^{-1} are also invertible (see Exercise 1). Also, recall that the $n \times n$ identity matrix I_n has the properties that $A \cdot I_n = I_n \cdot A = A$ and $A \cdot A^{-1} = I_n$ for all $A \in \mathrm{GL}(n, k)$. In the terminology of Appendix A, we say that $\mathrm{GL}(n, k)$ is a *group.*

Note that $A \in \mathrm{GL}(n, k)$ gives an invertible linear map $A : k^n \to k^n$ via matrix multiplication. Since every linear map from k^n to itself arises in this way, it is customary to call $\mathrm{GL}(n, k)$ the *general linear group.*

We will be most interested in the following subsets of $\mathrm{GL}(n,k)$.

Definition 2. *A finite subset $G \subset \mathrm{GL}(n,k)$ is called a* **finite matrix group** *provided it is nonempty and closed under matrix multiplication. The number of elements of G is called the* **order** *of G and is denoted $|G|$.*

Let us look at some examples of finite matrix groups.

Example 3. Suppose that $A \in \mathrm{GL}(n,k)$ is a matrix such that $A^m = I_n$ for some positive integer m. If m is the smallest such integer, then it is easy to show that

$$C_m = \{I_n, A, \ldots, A^{m-1}\} \subset \mathrm{GL}(n,k)$$

is closed under multiplication (see Exercise 2) and, hence, is a finite matrix group. We call C_m a *cyclic group* of order m. An example is given by

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{GL}(2,k).$$

One can check that $A^4 = I_2$, so that $C_4 = \{I_2, A, A^2, A^3\}$ is a cyclic matrix group of order 4 in $\mathrm{GL}(2,k)$.

Example 4. An important example of a finite matrix group comes from the permutations of variables discussed in §1. Let τ denote a permutation $x_{i_1}, \ldots, x_{i_n}$ of $x_1, \ldots, x_n$. Since τ is determined by what it does to the subscripts, we will set $i_1 = \tau(1), i_2 = \tau(2), \ldots, i_n = \tau(n)$. Then the corresponding permutation of variables is $x_{\tau(1)}, \ldots, x_{\tau(n)}$.

We can create a matrix from τ as follows. Consider the linear map that takes $(x_1, \ldots, x_n)$ to $(x_{\tau(1)}, \ldots, x_{\tau(n)})$. The matrix representing this linear map is denoted M_τ and is called a *permutation matrix*. Thus, M_τ has the property that under matrix multiplication, it permutes the variables according to τ:

$$M_\tau \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_{\tau(1)} \\ \vdots \\ x_{\tau(n)} \end{pmatrix}.$$

We leave it as an exercise to show that M_τ is obtained from the identity matrix by permuting its columns according to τ. More precisely, the $\tau(i)$-th column of M_τ is the i-th column of I_n. As an example, consider the permutation τ that takes (x, y, z) to (y, z, x). Here, $\tau(1) = 2$, $\tau(2) = 3$, and $\tau(3) = 1$, and one can check that

$$M_\tau \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ z \\ x \end{pmatrix}.$$

Since there are $n!$ ways to permute the variables, we get $n!$ permutation matrices. Furthermore, this set is closed under matrix multiplication, for it is easy to show that

$$M_\tau \cdot M_\nu = M_{\nu\tau},$$

where $\upsilon\tau$ is the permutation that takes i to $\upsilon(\tau(i))$ (see Exercise 4). Thus, the permutation matrices form a finite matrix group in $\mathrm{GL}(n, k)$. We will denote this matrix group by S_n. (Strictly speaking, the group of permutation matrices is only isomorphic to S_n in the sense of group theory. We will ignore this distinction.)

Example 5. Another important class of finite matrix groups comes from the symmetries of regular polyhedra. For example, consider a cube in $\mathbb{R}^3$ centered at the origin. The set of rotations of $\mathbb{R}^3$ that take the cube to itself is clearly finite and closed under multiplication. Thus, we get a finite matrix group in $\mathrm{GL}(3, \mathbb{R})$. In general, all finite matrix groups in $\mathrm{GL}(3, \mathbb{R})$ have been classified, and there is a rich geometry associated with such groups (see Exercises 5–9 for some examples). To pursue this topic further, the reader should consult BENSON and GROVE (1985), KLEIN (1884), or COXETER (1973).

Finite matrix groups have the following useful properties.

Proposition 6. *Let $G \subset \mathrm{GL}(n, k)$ be a finite matrix group. Then:*
(i) $I_n \in G$.
(ii) *If $A \in G$, then $A^m = I_n$ for some positive integer m.*
(iii) *If $A \in G$, then $A^{-1} \in G$.*

Proof. Take $A \in G$. Then $\{A, A^2, A^3, \ldots\} \in G$ since G is closed under multiplication. The finiteness of G then implies that $A^i = A^j$ for some $i > j$, and since A is invertible, we can multiply each side by A^{-j} to conclude that $A^m = I_n$, where $m = i - j > 0$. This proves (ii).

To prove (iii), note that (ii) implies $I_n = A^m = A \cdot A^{m-1} = A^{m-1} \cdot A$. Thus, $A^{-1} = A^{m-1} \in G$ since G is closed under multiplication. As for (i), since $G \neq \phi$, we can pick $A \in G$, and then, by (ii), $I_n = A^m \in G$. □

We next observe that elements of $\mathrm{GL}(n, k)$ act on polynomials in $k[x_1, \ldots, x_n]$. To see how this works, let $A = (a_{ij}) \in \mathrm{GL}(n, k)$ and $f \in k[x_1, \ldots, x_n]$. Then

$$g(x_1, \ldots, x_n) = f(a_{11}x_1 + \cdots + a_{1n}x_n, \ldots, a_{n1}x_1 + \cdots + a_{nn}x_n) \tag{1}$$

is again a polynomial in $k[x_1, \ldots, x_n]$. To express this more compactly, let $\mathbf{x}$ denote the column vector of the variables $x_1, \ldots, x_n$. Thus,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Then we can use matrix multiplication to express equation (1) as

$$g(\mathbf{x}) = f(A \cdot \mathbf{x}).$$

If we think of A as a change of basis matrix, then g is simply f viewed using the new coordinates.

For an example of how this works, let $f(x, y) = x^2 + xy + y^2 \in \mathbb{R}[x, y]$ and

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in \mathrm{GL}(2, \mathbb{R}).$$

Then

$$\begin{aligned} g(x, y) = f(A \cdot \mathbf{x}) &= f\left(\frac{x-y}{\sqrt{2}}, \frac{x+y}{\sqrt{2}}\right) \\ &= \left(\frac{x-y}{\sqrt{2}}\right)^2 + \frac{x-y}{\sqrt{2}} \cdot \frac{x+y}{\sqrt{2}} + \left(\frac{x+y}{\sqrt{2}}\right)^2 \\ &= \frac{3}{2}x^2 + \frac{1}{2}y^2. \end{aligned}$$

Geometrically, this says that we can eliminate the xy term of f by rotating the axes 45°.

A remarkable fact is that sometimes this process gives back the same polynomial we started with. For example, if we let $h(x, y) = x^2 + y^2$ and use the above matrix A, then one can check that

$$h(\mathbf{x}) = h(A \cdot \mathbf{x}).$$

In this case, we say that h is *invariant* under A.

This leads to the following fundamental definition.

Definition 7. *Let $G \subset \mathrm{GL}(n, k)$ be a finite matrix group. Then a polynomial $f(\mathbf{x}) \in k[x_1, \ldots, x_n]$ is* **invariant under** *G if*

$$f(\mathbf{x}) = f(A \cdot \mathbf{x})$$

for all $A \in G$. The set of all invariant polynomials is denoted $k[x_1, \ldots, x_n]^G$.

The most basic example of invariants of a finite matrix group is given by the symmetric polynomials.

Example 8. If we consider the group $S_n \subset \mathrm{GL}(n, k)$ of permutation matrices, then it is obvious that

$$k[x_1, \ldots, x_n]^{S_n} = \{\text{all symmetric polynomials in } k[x_1, \ldots, x_n]\}.$$

By Theorem 3 of §1, we know that symmetric polynomials are polynomials in the elementary symmetric functions with coefficients in k. We can write this as

$$k[x_1, \ldots, x_n]^{S_n} = k[\sigma_1, \ldots, \sigma_n].$$

Thus, every invariant can be written as a polynomial in finitely many invariants (the elementary symmetric functions). In addition, we know that the representation in terms of the elementary symmetric functions is unique. Hence, we have a very explicit knowledge of the invariants of S_n.

One of the goals of invariant theory is to examine whether all invariants $k[x_1, \ldots, x_n]^G$ are as nice as Example 8. To begin our study of this question, we first show that the set of invariants $k[x_1, \ldots, x_n]^G$ has the following algebraic structure.

Proposition 9. *Let $G \subset \mathrm{GL}(n, k)$ be a finite matrix group. Then the set $k[x_1, \ldots, x_n]^G$ is closed under addition and multiplication and contains the constant polynomials.*

Proof. We leave the easy proof as an exercise. □

Multiplication and addition in $k[x_1, \ldots, x_n]^G$ automatically satisfy the distributive, associative, etc., properties since these properties are true in $k[x_1, \ldots, x_n]$. In the terminology of Chapter 5, we say that $k[x_1, \ldots, x_n]^G$ is a *ring*. Furthermore, we say that $k[x_1, \ldots, x_n]^G$ is a *subring* of $k[x_1, \ldots, x_n]$.

So far in this book, we have learned three ways to create new rings. In Chapter 5, we saw how to make the quotient ring $k[x_1, \ldots, x_n]/I$ of an ideal $I \subset k[x_1, \ldots, x_n]$ and the coordinate ring $k[V]$ of an affine variety $V \subset k^n$. Now we can make the ring of invariants $k[x_1, \ldots, x_n]^G$ of a finite matrix group $G \subset \mathrm{GL}(n, k)$. In §4, we will see how these constructions are related.

In §1, we saw that the homogeneous components of a symmetric polynomial were also symmetric. We next observe that this holds for the invariants of any finite matrix group.

Proposition 10. *Let $G \subset \mathrm{GL}(n, k)$ be a finite matrix group. Then a polynomial $f \in k[x_1, \ldots, x_n]$ is invariant under G if and only if its homogeneous components are.*

Proof. See Exercise 11. □

In many situations, Proposition 10 will allow us to reduce to the case of homogeneous invariants. This will be especially useful in some of the proofs given in §3.

The following lemma will prove useful in determining whether a given polynomial is invariant under a finite matrix group.

Lemma 11. *Let $G \subset \mathrm{GL}(n, k)$ be a finite matrix group and suppose that we have $A_1, \ldots, A_m \in G$ such that every $A \in G$ can be written in the form*

$$A = B_1 B_2 \cdots B_t,$$

*where $B_i \in \{A_1, \ldots, A_m\}$ for every i (we say that $A_1, \ldots, A_m$ **generate** G). Then $f \in k[x_1, \ldots, x_n]$ is in $k[x_1, \ldots, x_n]^G$ if and only if*

$$f(\mathbf{x}) = f(A_1 \cdot \mathbf{x}) = \cdots = f(A_m \cdot \mathbf{x}).$$

Proof. We first show that if f is invariant under matrices $B_1, \ldots, B_t$, then it is also invariant under their product $B_1 \cdots B_t$. This is clearly true for $t = 1$. If we assume it is

true for $t-1$, then

$$\begin{aligned} f((B_1 \cdots B_t)\cdot \mathbf{x}) &= f((B_1 \cdots B_{t-1})\cdot B_t\mathbf{x}) \\ &= f(B_t\mathbf{x}) && \text{(by our inductive assumption)} \\ &= f(\mathbf{x}) && \text{(by the invariance under } B_t\text{)}. \end{aligned}$$

Now suppose that f is invariant under $A_1, \ldots, A_m$. Since elements $A \in G$ can be written $A = B_1 \cdots B_t$, where every B_i is one of $A_1, \ldots, A_m$, it follows immediately that $f \in k[x_1, \ldots, x_n]^G$. The converse is trivial and the lemma is proved. □

We can now compute some interesting examples of rings of invariants.

Example 12. Consider the finite matrix group

$$V_4 = \left\{ \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix} \right\} \subset \mathrm{GL}(2, k).$$

This is sometimes called the *Klein four-group*. We use the letter V_4 because "four" in German is "vier." You should check that the two matrices

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

generate V_4. Then Lemma 11 implies that a polynomial $f \in k[x, y]$ is invariant under V_4 if and only if

$$f(x, y) = f(-x, y) = f(x, -y)$$

Writing $f = \sum_{ij} a_{ij} x^i y^j$, we can understand the first of these conditions as follows:

$$\begin{aligned} f(x, y) = f(-x, y) &\iff \sum_{ij} a_{ij} x^i y^j = \sum_{ij} a_{ij} (-x)^i y^j \\ &\iff \sum_{ij} a_{ij} x^i y^j = \sum_{ij} (-1)^i a_{ij} x^i y^i \\ &\iff a_{ij} = (-1)^i a_{ij} \quad \text{for all } i, j \\ &\iff a_{ij} = 0 \quad \text{for } i \text{ odd} \end{aligned}$$

It follows that x always appears to an even power. Similarly, the condition $f(x, y) = f(x, -y)$ implies that y appears to even powers. Thus, we can write

$$f(x, y) = g(x^2, y^2)$$

for a unique polynomial $g(x, y) \in k[x, y]$. Conversely, every polynomial f of this form is clearly invariant under V_4. This proves that

$$k[x, y]^{V_4} = k[x^2, y^2].$$

Hence, every invariant of V_4 can be uniquely written as a polynomial in the two

homogeneous invariants x^2 and y^2. In particular, the invariants of the Klein four-group behave very much like the symmetric polynomials.

Example 13. For a finite matrix group that is less well-behaved, consider the cyclic group $C_2 = \{\pm I_2\} \subset \mathrm{GL}(2, k)$ of order 2. In this case, the invariants consist of the polynomials $f \in k[x, y]$ for which $f(x, y) = f(-x, -y)$. We leave it as an exercise to show that this is equivalent to the condition

$$f(x, y) = \sum_{ij} a_{ij} x^i y^j, \quad \text{where } a_{ij} = 0 \text{ whenever } i + j \text{ is odd.}$$

This means that f is invariant under C_2 if and only if the exponents of x and y always have the same parity (i.e., both even or both odd). Hence, we can write a monomial $x^i y^j$ appearing in f in the form

$$x^i y^j = \begin{cases} x^{2k} y^{2l} = (x^2)^k (y^2)^l & \text{if } i,\ j \text{ are even} \\ x^{2k+1} y^{2l+1} = (x^2)^k (y^2)^l xy & \text{if } i,\ j \text{ are odd.} \end{cases}$$

This means that every monomial in f, and hence f itself, is a polynomial in the homogeneous invariants x^2, y^2 and xy. We will write this as

$$k[x, y]^{C_2} = k[x^2, y^2, xy].$$

Note also that we need all three invariants to generate $k[x, y]^{C_2}$.

The ring $k[x^2, y^2, xy]$ is fundamentally different from the previous examples because uniqueness breaks down: a given invariant can be written in terms of x^2, y^2, xy in more than one way. For example, $x^4 y^2$ is clearly invariant under C_2, but

$$x^4 y^2 = (x^2)^2 \cdot y^2 = x^2 \cdot (xy)^2.$$

In §4, we will see that the crux of the matter is the algebraic relation $x^2 \cdot y^2 = (xy)^2$ between the basic invariants. In general, a key part of the theory is determining all algebraic relations between invariants. Given this information, one can describe precisely how uniqueness fails.

From these examples, we see that given a finite matrix group G, invariant theory has two basic questions to answer about the ring of invariants $k[x_1, \ldots, x_n]^G$:

- (Finite Generation) Can we find finitely many homogeneous invariants $f_1, \ldots, f_m$ such that every invariant is a polynomial in $f_1, \ldots, f_m$?
- (Uniqueness) In how many ways can an invariant be written in terms of $f_1, \ldots, f_m$? In §4, we will see that this asks for the algebraic relations among $f_1, \ldots, f_m$.

In §§3 and 4, we will give complete answers to both questions. We will also describe algorithms for finding the invariants and the relations between them.

EXERCISES FOR §2

1. If $A, B \in \mathrm{GL}(n, k)$ are invertible matrices, show that AB and A^{-1} are also invertible.
2. Suppose that $A \in \mathrm{GL}(n, k)$ satisfies $A^m = I_n$ for some positive integer. If m is the smallest such integer, then prove that the set $C_m = \{I_n, A, A^2, \ldots, A^{m-1}\}$ has exactly m elements and is closed under matrix multiplication.

3. Write down the six permutation matrices in $\mathrm{GL}(3, k)$.
4. Let M_τ be the matrix of the linear transformation taking $x_1, \dots, x_n$ to $x_{\tau(1)}, \dots, x_{\tau(n)}$. This means that if $e_1, \dots, e_n$ is the standard basis of k^n, then $M_\tau \cdot (\sum_j x_j e_j) = \sum_j x_{\tau(j)} e_j$.
   a. Show that $M_\tau \cdot e_{\tau(i)} = e_i$. Hint: Observe that $\sum_j x_j e_j = \sum_j x_{\tau(j)} e_{\tau(j)}$.
   b. Prove that the $\tau(i)$-th column of M_τ is the i-th column of the identity matrix.
   c. Prove that $M_\tau \cdot M_\nu = M_{\nu\tau}$, where $\nu\tau$ is the permutation taking i to $\nu(\tau(i))$.
5. Consider a cube in $\mathbb{R}^3$ centered at the origin whose edges have length 2 and are parallel to the coordinate axes.
   a. Show that there are finitely many rotations of $\mathbb{R}^3$ about the origin which take the cube to itself and show that these rotations are closed under composition. Taking the matrices representing these rotations, we get a finite matrix group $G \subset \mathrm{GL}(3, \mathbb{R})$.
   b. Show that G has 24 elements. Hint: Every rotation is a rotation about a line through the origin. So you first need to identify the "lines of symmetry" of the cube.
   c. Write down the matrix of the element of G corresponding to the 120° counterclockwise rotation of the cube about the diagonal connecting the vertices $(-1, -1, -1)$ and $(1, 1, 1)$.
   d. Write down the matrix of the element of G corresponding to the 90° counterclockwise rotation about the z-axis.
   e. Argue geometrically that G is generated by the two matrices from parts (c) and (d).
6. In this exercise, we will use geometric methods to find some invariants of the rotation group G of the cube (from Exercise 5).
   a. Explain why $x^2+y^2+z^2 \in \mathbb{R}[x, y, z]^G$. Hint: Think geometrically in terms of distance to the origin.
   b. Argue geometrically that the union of the three coordinate planes $\mathbf{V}(xyz)$ is invariant under G.
   c. Show that $\mathbf{I}(\mathbf{V}(xyz)) = (xyz)$ and conclude that if $f = xyz$, then for each $A \in G$, we have $f(A \cdot \mathbf{x}) = axyz$ for some real number a.
   d. Show that $f = xyz$ satisfies $f(A \cdot \mathbf{x}) = \pm xyz$ for all $A \in G$ and conclude that $x^2y^2z^2 \in k[x, y, z]^G$. Hint: Use part (c) and the fact that $A^m = I_3$ for some positive integer m.
   e. Use similar methods to show that the polynomials
   $$\Big((x+y+z)(x+y-z)(x-y+z)(x-y-z)\Big)^2, \Big((x^2-y^2)(x^2-z^2)(y^2-z^2)\Big)^2$$
   are in $k[x, y, z]^G$. Hint: The plane $x+y+z=0$ is perpendicular to one of the diagonals of the cube.
7. This exercise will continue our study of the invariants of the rotation group G of the cube begun in Exercise 6.
   a. Show that a polynomial f is in $k[x, y, z]^G$ if and only if $f(x, y, z) = f(y, z, x) = f(-y, x, z)$. Hint: Use parts (c), (d), and (e) of Exercise 5.
   b. Let
   $$\begin{aligned} f &= xyz, \\ g &= (x+y+z)(x+y-z)(z-y+z)(x-y-z), \\ h &= (x^2-y^2)(x^2-z^2)(y^2-z^2). \end{aligned}$$
   In Exercise 6, we showed that $f^2, g^2, h^2 \in k[x, y, z]^G$. Show that $f, h \notin k[x, y, z]^G$, but g, $fh \in k[x, y, z]^G$. Combining this with the previous exercise, we have invariants $x^2+y^2+z^2$, g, f^2, fh, and h^2 of degrees 2, 4, 6, 9, and 12, respectively, in $k[x, y, z]^G$. In §3, we will see that h^2 can be expressed in terms of the others.
8. In this exercise, we will consider an interesting "duality" that occurs among the regular polyhedra.

a. Consider a cube and an octahedron in $\mathbb{R}^3$, both centered at the origin. Suppose the edges of the cube are parallel to the coordinate axes and the vertices of the octahedron are on the axes. Show that they have the same group of rotations. Hint: Put the vertices of the octahedron at the centers of the faces of the cube.
b. Show that the dodecahedron and the icosahedron behave the same way. Hint: What do you get if you link up the centers of the 12 faces of the dodecahedron?
c. Parts (a) and (b) show that in a certain sense, the "dual" of the cube is the octahedron and the "dual" of the dodecahedron is the icosahedron. What is the "dual" of the tetrahedron?

9. (Requires abstract algebra) In this problem, we will consider a tetrahedron centered at the origin of $\mathbb{R}^3$.
   a. Show that the rotations of $\mathbb{R}^3$ about the origin which take the tetrahedron to itself give us a finite matrix group G of order 12 in $\mathrm{GL}(3, \mathbb{R})$.
   b. Since every rotation of the tetrahedron induces a permutation of the four vertices, show that we get a group homomorphism $\rho : G \to S_4$.
   c. Show that ρ is injective and that its image is the alternating group A_4. This shows that the rotation group of the tetrahedron is isomorphic in A_4.
10. Prove Proposition 9.
11. Prove Proposition 10. Hint: If $A = (a_{ij}) \in \mathrm{GL}(n, k)$ and $x_1^{i_1} \cdots x_n^{i_n}$ is a monomial of total degree $k = i_1 + \cdots + i_n$ appearing in f, then show that

$$(a_{11}x_1 + \cdots + a_{1n}x_n)^{i_1} \cdots (a_{n1}x_1 + \cdots + a_{nn}x_n)^{i_n}$$

    is homogeneous of total degree k.
12. In Example 13, we studied polynomials $f \in k[x, y]$ with the property that $f(x, y) = f(-x, -y)$. If $f = \sum_{ij} a_{ij} x^i y^j$, show that the above condition is equivalent to $a_{ij} = 0$ whenever $i + j$ is odd.
13. In Example 13, we discovered the algebraic relation $x^2 \cdot y^2 = (xy)^2$ between the invariants x^2, y^2, and xy. We want to show that this is essentially the only relation. More precisely, suppose that we have a polynomial $g(u, v, w) \in k[u, v, w]$ such that $g(x^2, y^2, xy) = 0$. We want to prove that $g(u, v, w)$ is a multiple (in $k[u, v, w]$) of $uv - w^2$ (which is the polynomial corresponding to the above relation).
   a. If we divide g by $uv - w^2$ using lex order with $u > v > w$, show that the remainder can be written in the form $uA(u, w) + vB(v, w) + C(w)$.
   b. Show that a polynomial $r = uA(u, w) + vB(v, w) + C(w)$ satisfies $r(x^2, y^2, xy) = 0$ if and only if $r = 0$.
14. Consider the finite matrix group $C_4 \subset \mathrm{GL}(2, \mathbb{C})$ generated by

$$A = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \in \mathrm{GL}(2, \mathbb{C})$$

    a. Prove that C_4 is cyclic of order 4.
    b. Use the method of Example 13 to determine $\mathbb{C}[x, y]^{C_4}$.
    c. Is there an algebraic relation between the invariants you found in part (b)? Can you give an example to show how uniqueness fails?
    d. Use the method of Exercise 13 to show that the relation found in part (c) is the only relation between the invariants.
15. Consider

$$V_4 = \left\{ \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \pm \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\} \subset \mathrm{GL}(2, k)$$

a. Show that V_4 is a finite matrix group of order 4.
b. Determine $k[x, y]^{V_4}$.
c. Show that any invariant can be written uniquely in terms of the generating invariants you found in part (b).

16. In Example 3, we introduced the finite matrix group C_4 in $\mathrm{GL}(2, k)$ generated by

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{GL}(2, k).$$

Try to apply the methods of Examples 12 and 13 to determine $k[x, y]^{C_4}$. Even if you cannot find all of the invariants, you should be able to find some invariants of low total degree. In §3, we will determine $k[x, y]^{C_4}$ completely.

§3 Generators for the Ring of Invariants

The goal of this section is to determine, in an algorithmic fashion, the ring of invariants $k[x_1, \ldots, x_n]^G$ of a finite matrix group $G \subset \mathrm{GL}(n, k)$. As in §2, we assume that our field k has characteristic zero. We begin by introducing some terminology used implicitly in §2.

Definition 1. *Given $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$, we let $k[f_1, \ldots, f_m]$ denote the subset of $k[x_1, \ldots, x_n]$ consisting of all polynomial expressions in $f_1, \ldots, f_m$ with coefficients in k.*

This means that the elements $f \in k[f_1, \ldots, f_m]$ are those polynomials which can be written in the form

$$f = g(f_1, \ldots, f_m),$$

where g is a polynomial in m variables with coefficients in k.

Since $k[f_1, \ldots, f_m]$ is closed under multiplication and addition and contains the constants, it is a subring of $k[x_1, \ldots, x_n]$. We say that $k[f_1, \ldots, f_m]$ is *generated by $f_1, \ldots, f_m$ over k*. One has to be slightly careful about the terminology: the subring $k[f_1, \ldots, f_m]$ and the ideal $\langle f_1, \ldots, f_m \rangle$ are both "generated" by $f_1, \ldots, f_m$, but in each case, we mean something slightly different. In the exercises, we will give some examples to help explain the distinction.

An important tool we will use in our study of $k[x_1, \ldots, x_n]^G$ is the *Reynolds operator*, which is defined as follows.

Definition 2. *Given a finite matrix group $G \subset \mathrm{GL}(n, k)$, the* **Reynolds operator** *of G is the map $R_G : k[x_1, \ldots, x_n] \to k[x_1, \ldots, x_n]$ defined by the formula*

$$R_G(f)(\mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(A \cdot \mathbf{x})$$

for $f(\mathbf{x}) \in k[x_1, \ldots, x_n]$.

One can think of $R_G(f)$ as "averaging" the effect of G on f. Note that division by $|G|$ is allowed since k has characteristic zero. The Reynolds operator has the following crucial properties.

Proposition 3. *Let R_G be the Reynolds operator of the finite matrix group G.*
(i) *R_G is k-linear in f.*
(ii) *If $f \in k[x_1, \dots, x_n]$, then $R_G(f) \in k[x_1, \dots, x_n]^G$.*
(iii) *If $f \in k[x_1, \dots, x_n]^G$, then $R_G(f) = f$.*

Proof. We will leave the proof of (i) as an exercise. To prove (ii), let $B \in G$. Then

$$R_G(f)(B\mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(A \cdot B\mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(AB \cdot \mathbf{x}). \tag{1}$$

Writing $G = \{A_1, \dots, A_{|G|}\}$, note that $A_iB \neq A_jB$ when $i \neq j$ (otherwise, we could multiply each side by B^{-1} to conclude that $A_i = A_j$). Thus the subset $\{A_1B, \dots, A_{|G|}B\} \subset G$ consists of $|G|$ distinct elements of G and hence must equal G. This shows that

$$G = \{AB : A \in G\}.$$

Consequently, in the last sum of (1), the polynomials $f(AB \cdot \mathbf{x})$ are just the $f(A \cdot \mathbf{x})$, possibly in a different order. Hence,

$$\frac{1}{|G|} \sum_{A \in G} f(AB \cdot \mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(A \cdot \mathbf{x}) = R_G(f)(\mathbf{x}),$$

and it follows that $R_G(f)(B \cdot \mathbf{x}) = R_G(f)(\mathbf{x})$ for all $B \in G$. This implies $R_G(f) \in k[x_1, \dots, x_n]^G$.

Finally, to prove (iii), note that if $f \in k[x_1, \dots, x_n]^G$, then

$$R_G(f)(\mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(A \cdot \mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(\mathbf{x}) = f(\mathbf{x})$$

since f invariant. This completes the proof. □

One nice aspect of this proposition is that it gives us a way of creating invariants. Let us look at an example.

Example 4. Consider the cyclic matrix group $C_4 \subset \mathrm{GL}(2, k)$ of order 4 generated by

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

By Lemma 11 of §2, we know that

$$k[x, y]^{C_4} = \{f \in k[x, y] : f(x, y) = f(-y, x)\}.$$

One can easily check that the Reynolds operator is given by

$$R_{C_4}(f)(x, y) = \frac{1}{4}(f(x, y) + f(-y, x) + f(-x, -y) + f(y, -x))$$

(see Exercise 3). Using Proposition 3, we can compute some invariants as follows:

$$\begin{aligned}
R_{C_4}(x^2) &= \tfrac{1}{4}(x^2 + (-y)^2 + (-x)^2 + y^2) = \tfrac{1}{2}(x^2 - y^2),\\
R_{C_4}(xy) &= \tfrac{1}{4}(xy + (-y)x + (-x)(-y) + y(-x)) = 0,\\
R_{C_4}(x^3y) &= \tfrac{1}{4}(x^3y + (-y)^3x + (-x)^3(-y) + y^3(-x)) = \tfrac{1}{2}(x^3y - xy^3),\\
R_{C_4}(x^2y^2) &= \tfrac{1}{4}(x^2y^2 + (-y)^2x^2 + (-x)^2(-y)^2 + y^2(-x)^2) = x^2y^2.
\end{aligned}$$

Thus, x^2+y^2, x^3y-xy^3, $x^2y^2 \in k[x, y]^{C_4}$. We will soon see that these three invariants generate $k[x, y]^{C_4}$.

It is easy to prove that for any monomial x^α, the Reynolds operator gives us a homogeneous invariant $R_G(x^\alpha)$ of total degree $|\alpha|$ whenever it is nonzero. The following wonderful theorem of Emmy Noether shows that we can always find finitely many of these invariants that generate $k[x_1, \dots, x_n]^G$.

Theorem 5. *Given a finite matrix group $G \subset \mathrm{GL}(n, k)$, we have*

$$k[x_1, \dots, x_n]^G = k[R_G(x^\beta) : |\beta| \le |G|].$$

In particular, $k[x_1, \dots, x_n]^G$ is generated by finitely many homogeneous invariants.

Proof. If $f = \sum_\alpha c_\alpha x^\alpha \in k[x_1, \dots, x_n]^G$, then Proposition 3 implies that

$$f = R_G(f) = R_G\left(\sum_\alpha c_\alpha x^\alpha\right) = \sum_\alpha c_\alpha R_G(x^\alpha).$$

Hence every invariant is a linear combination (over k) of the $R_G(x^\alpha)$. Consequently, it suffices to prove that for all α, $R_G(x^\alpha)$ is a polynomial in the $R_G(x^\beta)$, $|\beta| \le |G|$.

Noether's clever idea was to fix an integer k and combine *all* $R_G(x^\beta)$ of total degree k into a power sum of the type considered in §1. Using the theory of symmetric functions, this can be expressed in terms of finitely many power sums, and the theorem will follow.

The first step in implementing this strategy is to expand $(x_1 + \cdots + x_n)^k$ into a sum of monomials x^α with $|\alpha| = k$:

$$(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} a_\alpha x^\alpha. \tag{2}$$

In Exercise 4, you will prove that a_α is a positive integer for all $|\alpha| = k$.

To exploit this identity, we need some notation. Given $A = (a_{ij}) \in G$, let A_i denote the i-th row of A. Thus, $A_i \cdot \mathbf{x} = a_{i1}x_1 + \cdots + a_{in}x_n$. Then, if $\alpha_1 = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n_{\ge 0}$, let

$$(A \cdot \mathbf{x})^\alpha = (A_1 \cdot \mathbf{x})^{\alpha_1} \cdots (A_n \cdot \mathbf{x})^{\alpha_n}.$$

In this notation, we have

$$R_G(x^\alpha) = \frac{1}{|G|}\sum_{A\in G}(A\cdot \mathbf{x})^\alpha.$$

Now introduce new variables $u_1, \dots, u_n$ and substitute $u_i A_i \cdot \mathbf{x}$ for x_i in (2). This gives the identity

$$(u_1 A_1\cdot \mathbf{x} + \cdots + u_n A_n \cdot \mathbf{x})^k = \sum_{|\alpha|=k} a_\alpha (A\cdot \mathbf{x})^\alpha u^\alpha.$$

If we sum over all $A \in G$, then we obtain

$$\begin{aligned} S_k = \sum_{A\in G}(u_1 A_1\cdot \mathbf{x} + \cdots + u_n A_n\cdot \mathbf{x})^k &= \sum_{|\alpha|=k} a_\alpha \left(\sum_{A\in G}(A\cdot \mathbf{x})^\alpha\right) u^\alpha \\ &= \sum_{|\alpha|=k} b_\alpha R_G(x^\alpha) u^\alpha, \end{aligned} \tag{3}$$

where $b_\alpha = |G|a_\alpha$. Note how the sum on the right encodes *all* $R_G(x^\alpha)$ with $|\alpha| = k$. This is why we use the variables $u_1, \dots, u_n$: they prevent any cancellation from occurring.

The left side of (3) is the k-th power sum S_k of the $|G|$ quantities

$$U_A = u_1 A_1\cdot \mathbf{x} + \cdots + u_n A_n \cdot \mathbf{x}$$

indexed by $A \in G$. We write this as $S_k = S_k(U_A : A \in G)$. By Theorem 8 of §1, every symmetric function in the $|G|$ quantities U_A is a polynomial in $S_1, \dots, S_{|G|}$. Since S_k is symmetric in the U_A, it follows that

$$S_k = F(S_1, \dots, S_{|G|})$$

for some polynomial F with coefficients in k. Substituting in (3), we obtain

$$\sum_{|\alpha|=k} b_\alpha R_G(x^\alpha)u^\alpha = F\left(\sum_{|\beta|=1} b_\beta R_G(x^\beta)u^\beta, \dots, \sum_{|\beta|=|G|} R_G(x^\beta)u^\beta\right).$$

Expanding the right side and equating the coefficients of u^α, it follows that

$$b_\alpha R_G(x^\alpha) = \text{a polynomial in the } R_G(x^\beta), \quad |\beta| \le |G|.$$

Since k has characteristic zero, the coefficient $b_\alpha = |G|a_\alpha$ is nonzero in k, and hence $R_G(x^\alpha)$ has the desired form. This completes the proof of the theorem. □

This theorem solves the *finite generation problem* stated at the end of §2. In the exercises, you will give a second proof of the theorem using the Hilbert Basis Theorem.

To see the power of what we have just proved, let us compute some invariants.

Example 6. We will return to the cyclic group $C_4 \subset \mathrm{GL}(2,k)$ of order 4 from Example 4. To find the ring of invariants, we need to compute $R_{C_4}(x^i y^j)$ for all $i+j \le 4$. The following table records the results:

$x^i y^j$	$R_{C_4}(x^i y^j)$	$x^i y^j$	$R_{C_4}(x^i y^j)$
x	0	xy^2	0
y	0	y^3	0
x^2	$\frac{1}{2}(x^2+y^2)$	x^4	$\frac{1}{2}(x^4+y^4)$
xy	0	x^3y	$\frac{1}{2}(x^3y-xy^3)$
y^2	$\frac{1}{2}(x^2+y^2)$	x^2y^2	x^2y^2
x^3	0	xy^3	$-\frac{1}{2}(x^3y-xy^3)$
x^2y	0	y^4	$\frac{1}{2}(x^4+y^4)$

By Theorem 5, it follows that $k[x,y]^{C_4}$ is generated by the four invariants x^2+y^2, x^4+y^4, x^3y-xy^3 and x^2y^2. However, we do not need x^4+y^4 since

$$x^4+y^4 = (x^2+y^2)^2 - 2x^2y^2.$$

Thus, we have proved that

$$k[x,y]^{C_4} = k[x^2+y^2, x^3y-xy^3, x^2y^2].$$

The main drawback of Theorem 5 is that when $|G|$ is large, we need to compute the Reynolds operator for *lots* of monomials. For example, consider the cyclic group $C_8 \subset \mathrm{GL}(2,\mathbb{R})$ of order 8 generated by the 45° rotation

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in \mathrm{GL}(2,\mathbb{R}).$$

In this case, Theorem 5 says that $k[x,y]^{C_8}$ is generated by the 44 invariants $R_{C_8}(x^i y^j)$, $i+j \le 8$. In reality, only 3 are needed. For larger groups, things are even worse, especially if more variables are involved. See Exercise 10 for an example.

Fortunately, there are more efficient methods for finding a generating set of invariants. The main tool is *Molien's Theorem*, which enables one to predict in advance the number of linearly independent homogeneous invariants of given total degree. This theorem can be found in Chapter 7 of BENSON and GROVE (1985) and Chapter 2 of STURMFELS (1993). The latter also gives an efficient algorithm, based on Molien's Theorem, for finding invariants that generate $k[x_1,\dots,x_n]^G$.

Once we know $k[x_1,\dots,x_n]^G = k[f_1,\dots,f_m]$, we can ask if there is an algorithm for writing a given invariant $f \in k[x_1,\dots,x_n]^G$ in terms of $f_1,\dots,f_m$. For example, it is easy to check that the polynomial

$$f(x,y) = x^8 + 2x^6y^2 - x^5y^3 + 2x^4y^4 + x^3y^5 + 2x^2y^6 + y^8 \tag{4}$$

satisfies $f(x, y) = f(-y, x)$, and hence is invariant under the group C_4 from Example 4. Then Example 6 implies that $f \in k[x, y]^{C_4} = k[x^2 + y^2, x^3y - xy^3, x^2y^2]$. But how do we write f in terms of these three invariants? To answer this question, we will use a method similar to what we did in Proposition 4 of §1.

We will actually prove a bit more, for we will allow $f_1, \ldots, f_m$ to be *arbitrary* elements of $k[x_1, \ldots, x_n]$. The following proposition shows how to test whether a polynomial lies in $k[f_1, \ldots, f_m]$ and, if so, to write it in terms of $f_1, \ldots, f_m$.

Proposition 7. *Suppose that $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$ are given. Fix a monomial order in $k[x_1, \ldots, x_n, y_1, \ldots, y_m]$ where any monomial involving one of $x_1, \ldots, x_n$ is greater than all monomials in $k[y_1, \ldots, y_m]$. Let G be a Groebner basis of the ideal $\langle f_1 - y_1, \ldots, f_m - y_m \rangle \subset k[x_1, \ldots x_n, y_1, \ldots, y_m]$. Given $f \in k[x_1, \ldots, x_n]$, let $g = \overline{f}^G$ be the remainder of f on division by G. Then:*

(i) *$f \in k[f_1, \ldots, f_m]$ if and only if $g \in k[y_1, \ldots, y_m]$.*

(ii) *If $f \in k[f_1, \ldots, f_m]$, then $f = g(f_1, \ldots, f_m)$ is an expression of f as a polynomial in $f_1, \ldots, f_m$.*

Proof. The proof will be similar to the argument given in Proposition 4 of §1 (with one interesting difference). When we divide $f \in k[x_1, \ldots, x_n]$ by $G = \{g_1, \ldots, g_t\}$, we get an expression of the form

$$f = A_1 g_1 + \cdots + A_t g_t + g.$$

with $A_1, \ldots, A_t, g \in k[x_1, \ldots, x_n, y_1, \ldots, y_m]$.

To prove (i), first suppose that $g \in k[y_1, \ldots, y_m]$. Then for each i, substitute f_i for y_i in the above formula for f. This substitution will not affect f since it involves only $x_1, \ldots, x_n$, but it sends every polynomial in $\langle f_1 - y_1, \ldots, f_m - y_m \rangle$ to zero. Since $g_1, \ldots, g_t$ lie in this ideal, it follows that $f = g(f_1, \ldots, f_m)$. Hence, $f \in k[f_1, \ldots, f_m]$.

Conversely, suppose that $f = g(f_1, \ldots, f_m)$ for some $g \in k[y_1, \ldots, y_m]$. Arguing as in §1, one sees that

$$f = C_1 \cdot (f_1 - y_1) + \cdots + C_m \cdot (f_m - y_m) + g(y_1, \ldots, y_m) \tag{5}$$

[see equation (4) of §1]. Unlike the case of symmetric polynomials, g need not be the remainder of f on division by G—we still need to reduce some more.

Let $G' = G \cap k[y_1, \ldots, y_m]$ consist of those elements of G involving only $y_1, \ldots, y_m$. Renumbering if necessary, we can assume $G' = \{g_1, \ldots, g_s\}$, where $s \leq t$. If we divide g by G', we get an expression of the form

$$g = B_1 g_1 + \cdots + B_s g_s + g', \tag{6}$$

where $B_1, \ldots, B_s, g' \in k[y_1, \ldots, y_m]$. If we combine equations (5) and (6), we can

write f in the form

$$f = C_1' \cdot (f_1 - y_1) + \cdots + C_m' \cdot (f_m - y_m) + g'(y_1, \ldots\ y_m).$$

This follows because, in (6), each g_i lies in $\langle f_1 - y_1, \ldots, f_m - y_m \rangle$. We claim that g' is the remainder of f on division by G. This will prove that the remainder lies in $k[y_1, \ldots, y_m]$.

Since G a Groebner basis, Proposition 1 of Chapter 2, §6 tells us that g' is the remainder of f on division by G provided that no term of g' is divisible by an element of $\mathrm{LT}(G)$. To prove that g' has this property, suppose that there is $g_i \in G$ where $\mathrm{LT}(g_i)$ divides some term of g'. Then $\mathrm{LT}(g_i)$ involves only $y_1, \ldots, y_m$ since $g' \in k[y_1, \ldots, y_m]$. By our hypothesis on the ordering, it follows that $g_i \in k[y_1, \ldots, y_m]$ and hence, $g_i \in G'$. Since g' is a remainder on division by G', $\mathrm{LT}(g_i)$ cannot divide any term of g'. This contradiction shows that g' is the desired remainder.

Part (ii) of the proposition follows immediately from the above arguments, and we are done. □

In the exercises, you will use this proposition to write the polynomial

$$f(x, y) = x^8 + 2x^6y^2 - x^5y^3 + 2x^4y^4 + x^3y^5 + 2x^2y^6 + y^8$$

from (4) in terms of the generating invariants $x^2 + y^2$, $x^3y - xy^3$, x^2y^2 of $k[x, y]^{C_4}$.

The problem of finding generators for the ring of invariants (and the associated problem of finding the relations between them—see §4) played an important role in the development of invariant theory. Originally, the group involved was the group of all invertible matrices over a field. A classic introduction can be found in HILBERT (1993), and STURMFELS (1993) also discusses this case. For more on the invariant theory of finite groups, we recommend BENSON (1993), BENSON and GROVE (1985), SMITH (1995) and STURMFELS (1993).

EXERCISES FOR §3

1. Given $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$, we can "generate" the following two objects:
   - The ideal $\langle f_1, \ldots, f_m \rangle \subset k[x_1, \ldots, x_n]$ generated by $f_1, \ldots, f_m$. This consists of all expressions $\sum_{i=1}^m h_i f_i$, where $h_1, \ldots, h_m \in k[x_1, \ldots, x_n]$.
   - The subring $k[f_1, \ldots, f_m] \subset k[x_1, \ldots, x_n]$ generated by $f_1, \ldots, f_m$ over k. This consists of all expressions $g(f_1, \ldots, f_m)$ where g is a polynomial in m variables with coefficients in k.

   To illustrate the differences between these, we will consider the simple case where $f_1 = x^2 \in k[x]$.
   a. Explain why $1 \in k[x^2]$ but $1 \notin \langle x^2 \rangle$.
   b. Explain why $x^3 \notin k[x^2]$ but $x^3 \in \langle x^2 \rangle$.
2. Let G be a finite matrix group in $\mathrm{GL}(n, k)$. Prove that the Reynolds operator R_G has the following properties:
   a. If $a, b \in k$ and $f, g \in k[x_1, \ldots., x_n]$, then $R_G(af + bg) = aR_G(f) + bR_G(g)$.
   b. R_G maps $k[x_1, \ldots, x_n]$ to $k[x_1, \ldots, x_n]^G$ and is onto.
   c. $R_G \circ R_G = R_G$.
   d. If $f \in k[x_1, \ldots, x_n]^G$ and $g \in k[x_1, \ldots, x_n]$, then $R_G(fg) = f \cdot R_G(g)$.

3. In this exercise, we will work with the cyclic group $C_4 \subset \mathrm{GL}(2,k)$ from Example 4 in the text.
   a. Prove that the Reynolds operator of C_4 is given by
   $$R_{C_4}(f)(x,y) = \tfrac{1}{4}(f(x,y) + f(-y,x) + f(-x,-y) + f(y,-x)).$$
   b. Compute $R_{C_4}(x^i y^j)$ for all $i + j \leq 4$. Note that some of the computations are done in Example 4. You can check your answers against the table in Example 6.
4. In this exercise, we will study the identity (2) used in the proof of Theorem 5. We will use the *multinomial coefficients*, which are defined as follows. For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$, let $|\alpha| = k$ and define
   $$\binom{k}{\alpha} = \frac{k!}{\alpha_1!\alpha_2!\cdots\alpha_n!}.$$
   a. Prove that $\binom{k}{\alpha}$ is an integer. Hint: Use induction on n and note that when $n = 2$, $\binom{k}{\alpha}$ is a binomial coefficient.
   b. Prove that
   $$(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} \binom{k}{\alpha} x^\alpha.$$
   In particular, the coefficient a_α in equation (2) is the positive integer $\binom{k}{\alpha}$. Hint: Use induction on n and note that the case $n = 2$ is the binomial theorem.
5. Let $G \subset \mathrm{GL}(n,k)$ be a finite matrix group. In this exercise, we will give Hilbert's proof that $k[x_1, \ldots, x_n]^G$ is generated by finitely many homogeneous invariants. To begin the argument, let $I \subset k[x_1, \ldots, x_n]$ be the *ideal* generated by all homogeneous invariants of positive total degree.
   a. Explain why there are finitely many homogeneous invariants $f_1, \ldots, f_m$ such that $I = \langle f_1, \ldots, f_m \rangle$. The strategy of Hilbert's proof is to show that $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m]$. Since the inclusion $k[f_1, \ldots, x_n] \subset k[x_1, \ldots, x_n]^G$ is obvious, we must show that $k[x_1, \ldots, x_n]^G \not\subset k[f_1, \ldots, f_m]$ leads to a contradiction.
   b. Prove that $k[x_1, \ldots, x_n]^G \not\subset k[f_1, \ldots, f_m]$ implies there is a homogeneous invariant f of positive degree which is not in $k[f_1, \ldots, f_m]$.
   c. For the rest of the proof, pick f as in part (b) with *minimal* total degree k. By definition, $f \in I$, so that $f = \sum_{i=1}^m h_i f_i$ for $h_1, \ldots, h_m, \in k[x_1, \ldots, x_n]$. Prove that for each i, we can assume $h_i f_i$ is 0 or homogeneous of total degree k.
   d. Use the Reynolds operator to show that $f = \sum_{i=1}^m R_G(h_i) f_i$. Hint: Use Proposition 3 and Exercise 2. Also show that for each i, $R_G(h_i) f_i$ is 0 or homogeneous of total degree k.
   e. Since f_i has positive total degree, conclude that $R_G(h_i)$ is a homogeneous invariant of total degree $< k$. By the minimality of k, $R_G(h_i) \in k[f_1, \ldots, f_m]$ for all i. Prove that this contradicts $f \notin k[f_1, \ldots, f_m]$.

   This proof is a lovely application of the Hilbert Basis Theorem. The one drawback is that it does not tell us how to find the generators—the proof is purely nonconstructive. Thus, for our purposes, Noether's theorem is much more useful.
6. If we have two finite matrix groups G and H such that $G \subset H \subset \mathrm{GL}(n,k)$, prove that $k[x_1, \ldots, x_n]^H \subset k[x_1, \ldots, x_n]^G$.
7. Consider the matrix
   $$A = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \in \mathrm{GL}(2,k).$$

a. Show that A generates a cyclic matrix group C_3 of order 3.
b. Use Theorem 5 to find finitely many homogeneous invariants which generate $k[x, y]^{C_3}$.
c. Can you find fewer invariants that generate $k[x, y]^{C_3}$? Hint: If you have invariants $f_1, \ldots, f_m$, you can use Proposition 7 to determine whether $f_1 \in k[f_2, \ldots, f_m]$.

8. Let A be the matrix of Exercise 7.
a. Show that $-A$ generates a cyclic matrix group C_6, of order 6.
b. Show that $-I_2 \in C_6$. Then use Exercise 6 and §2 to show that $k[x, y]^{C_6} \subset k[x^2, y^2, xy]$. Conclude that all nonzero homogeneous invariants of C_6 have even total degree.
c. Use part (b) and Theorem 5 to find $k[x, y]^{C_6}$. Hint: There are still a lot of Reynolds operators to compute. You should use a computer algebra program to design a procedure that has i, j as input and $R_{C_6}(x^i y^j)$ as output.

9. Let A be the matrix

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in \mathrm{GL}(2, k).$$

a. Show that A generates a cyclic matrix group $C_8 \subset \mathrm{GL}(2, k)$.
b. Give a geometric argument to explain why $x^2 + y^2 \in k[x, y]^{C_8}$. Hint: A is a rotation matrix.
c. As in Exercise 8, explain why all homogeneous invariants of C_8 have even total degree.
d. Find $k[x, y]^{C_8}$. Hint: Do not do this problem unless you know how to design a procedure (on some computer algebra program) that has i, j as input and $R_{C_8}(x^i y^j)$ as output.

10. Consider the finite matrix group

$$G = \left\{ \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm 1 \end{pmatrix} \right\} \subset \mathrm{GL}(3, k).$$

Note that G has order 8.
a. If we were to use Theorem 5 to determine $k[x, y, z]^G$, for how many monomials would we have to compute the Reynolds operator?
b. Use the method of Example 12 in §2 to determine $k[x, y, z]^G$.

11. Let f be the polynomial (4) in the text.
a. Verify that $f \in k[x, y]^{C_4} = k[x^2 + y^2, x^3 y - x y^3, x^2 y^2]$.
b. Use Proposition 7 to express f as a polynomial in $x^2 + y^2, x^2 y - x y^3, x^2 y^2$.

12. In Exercises 5, 6, and 7 of §2, we studied the rotation group $G \subset \mathrm{GL}(3, \mathbb{R})$ of the cube in $\mathbb{R}^3$ and we found that $k[x, y, z]^G$ contained the polynomials

$$\begin{aligned} f_1 &= x^2 + y^2 + z^2, \\ f_2 &= (x + y + z)(x + y - z)(x - y + z)(x - y - z), \\ f_3 &= x^2 y^2 z^2, \\ f_4 &= xyz(x^2 - y^2)(x^2 - z^2)(y^2 - z^2). \end{aligned}$$

a. Give an elementary argument using degrees to show that $f_4 \notin k[f_1, f_2, f_3]$.
b. Use Proposition 7 to show that $f_3 \notin k[f_1, f_2]$.
c. In Exercise 6 of §2, we showed that

$$\Big((x^2 - y^2)(x^2 - z^2)(y^2 - z^2)\Big)^2 \in k[x, y, z]^G.$$

Prove that this polynomial lies in $k[f_1, f_2, f_3]$. Why can we ignore f_4?
Using Molien's Theorem and the methods of STURMFELS (1993), one can prove that $k[x, y, z]^G = k[f_1, f_2, f_3, f_4]$.

§4 Relations Among Generators and the Geometry of Orbits

Given a finite matrix group $G \subset \mathrm{GL}(n, k)$, Theorem 5 of §3 guarantees that there are finitely many homogeneous invariants $f_1, \ldots, f_m$ such that

$$k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m].$$

In this section, we will learn how to describe the *algebraic relations* among $f_1, \ldots, f_m$, and we will see that these relations have some fascinating algebraic and geometric implications.

We begin by recalling the *uniqueness problem* stated at the end of §2. For a symmetric polynomial $f \in k[x_1, \ldots, x_n]^{S_n} = k[\sigma_1, \ldots, \sigma_n]$, we proved that f could be written uniquely as a polynomial in $\sigma_1, \ldots \sigma_n$. For a general finite matrix group $G \subset \mathrm{GL}(n, k)$, if we know that $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m]$, then one could similarly ask if $f \in k[x_1, \ldots, x_n]^G$ can be uniquely written in terms of $f_1, \ldots, f_m$.

To study this question, note that if g_1 and g_2 are polynomials in $k[y_1, \ldots, y_m]$, then

$$g_1(f_1, \ldots, f_m) = g_2(f_1, \ldots, f_m) \Longleftrightarrow h(f_1, \ldots, f_m) = 0,$$

where $h = g_1 - g_2$. It follows that uniqueness fails if and only if there is a nonzero polynomial $h \in k[y_1, \ldots, y_m]$ such that $h(f_1, \ldots, f_m) = 0$. Such a polynomial is a *nontrivial algebraic relation* among $f_1, \ldots, f_m$.

If we let $F = (f_1, \ldots, f_m)$, then the set

$$I_F = \{h \in k[y_1, \ldots, y_m] : h(f_1, \ldots, f_m) = 0 \text{ in } k[x_1, \ldots, x_n]\} \tag{1}$$

records *all* algebraic relations among $f_1, \ldots, f_m$. This set has the following properties.

Proposition 1. *If $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m]$, let $I_F \subset k[y_1, \ldots, y_m]$ be as in* (1). *Then:*

(i) *I_F is a prime ideal of $k[y_1, \ldots, y_m]$.*
(ii) *Suppose that $f \in k[x_1, \ldots, x_n]^G$ and that $f = g(f_1, \ldots, f_m)$ is one representation of f in terms of $f_1, \ldots, f_m$. Then all such representations are given by*

$$f = g(f_1, \ldots, f_m) + h(f_1, \ldots, f_m),$$

as h varies over I_F.

Proof. For (i), it is an easy exercise to prove that I_F is an ideal. To show that it is prime, we need to show that $fg \in I_F$ implies that $f \in I_F$ or $g \in I_F$ (see Definition 2

of Chapter 4, §5). But $fg \in I_F$ means that $f(f_1, \ldots, f_m)g(f_1, \ldots, f_m) = 0$. This is a product of polynomials in $k[x_1, \ldots, x_n]$, and hence, $f(f_1, \ldots, f_m)$ or $g(f_1, \ldots, f_m)$ must be zero. Thus f or g is in I_F.

We leave the proof of (ii) as an exercise. □

We will call I_F the *ideal of relations* for $F = (f_1, \ldots, f_m)$. Another name for I_F used in the literature is the *syzygy ideal.* To see what Proposition 1 tells us about the uniqueness problem, consider $C_2 = \{\pm I_2\} \subset \mathrm{GL}(2, k)$. We know from §2 that $k[x, y]^{C_2} = k[x^2, y^2, xy]$, and, in Example 4, we will see that $I_F = \langle uv - w^2 \rangle \subset k[u, v, w]$. Now consider $x^6 + x^3y^3 \in k[x, y]^{C_2}$. Then Proposition 1 implies that *all* possible ways of writing $x^6 + x^3y^3$ in terms of x^2, y^2, xy are given by

$$(x^2)^3 + (xy)^3 + (x^2 \cdot y^2 - (xy)^2) \cdot b(x^2, y^2, xy)$$

since elements of $\langle uv - w^2 \rangle$ are of the form $(uv - w^2) \cdot b(u, v, w)$.

As an example of what the ideal of relations I_F can tell us, let us show how it can be used to reconstruct the ring of invariants.

Proposition 2. *If $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m]$, let $I_F \subset k[y_1, \ldots, y_m]$ be the ideal of relations. Then there is a ring isomorphism*

$$k[y_1, \ldots, y_m]/I_F \cong k[x_1, \ldots, x_n]^G$$

between the quotient ring of I_F (as defined in Chapter 5, §2) and the ring of invariants.

Proof. Recall from §2 of Chapter 5 that elements of the quotient ring $k[y_1, \ldots, y_m]/I_F$ are written $[g]$ for $g \in k[y_1, \ldots, y_m]$, where $[g_1] = [g_2]$ if and only if $g_1 - g_2 \in I_F$.

Now define $\phi : k[y_1, \ldots, y_m]/I_F \to k[x_1, \ldots, x_n]^G$ by

$$\phi([g]) = g(f_1, \ldots, f_m).$$

We leave it as an exercise to check that ϕ is well-defined and is a ring homomorphism. We need to show that ϕ is one-to-one and onto.

Since $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m]$, it follows immediately that ϕ is onto. To prove that ϕ is one-to-one, suppose that $\phi([g_1]) = \phi([g_2])$. Then $g_1(f_1, \ldots, f_m) = g_2(f_1, \ldots, f_m)$, which implies that $g_1 - g_2 \in I_F$. Thus, $[g_1] = [g_2]$, and hence, ϕ is one-to-one.

It is a general fact that if a ring homomorphism is one-to-one and onto, then its inverse function is a ring homomorphism. This proves that ϕ is a ring isomorphism. □

A more succinct proof of this proposition can be given using the Isomorphism Theorem of Exercise 16 in Chapter 5, §2.

For our purposes, another extremely important property of I_F is that we can compute it explicitly using elimination theory. Namely, consider the system of equations

$$\begin{aligned} y_1 &= f_1(x_1, \dots, x_n), \\ &\vdots \\ y_m &= f_m(x_1, \dots, x_n). \end{aligned}$$

Then I_F can be obtained by eliminating $x_1, \dots, x_n$ from these equations.

Proposition 3. *If* $k[x_1, \dots, x_n]^G = k[f_1, \dots, f_m]$, *consider the ideal*

$$J_F = \langle f_1 - y_1, \dots, f_m - y_m \rangle \subset k[x_1, \dots, x_n, y_1, \dots, y_m].$$

(i) *I_F is the n-th elimination ideal of J_F. Thus, $I_F = J_F \cap k[y_1, \dots, y_m]$.*
(ii) *Fix a monomial order in $k[x_1, \dots, x_n, y_1, \dots, y_m]$ where any monomial involving one of $x_1, \dots, x_n$ is greater than all monomials in $k[y_1, \dots, y_m]$ and let G be a Groebner basis of J_F. Then $G \cap k[y_1, \dots, y_m]$ is a Groebner basis for I_F in the monomial order induced on $k[y_1, \dots, y_m]$.*

Proof. Note that the ideal J_F appeared earlier in Proposition 7 of §3. To relate J_F to the ideal of relations I_F, we will need the following characterization of J_F: if $p \in k[x_1, \dots, x_n, y_1, \dots, y_m]$, then we claim that

$$p \in J_F \iff p(x_1, \dots, x_n, f_1, \dots, f_m) = 0 \text{ in } k[x_1, \dots, x_n]. \tag{2}$$

One implication is obvious since the substitution $y_i \mapsto f_i$ takes all elements of $J_F = \langle f_1 - y_1, \dots, f_m - y_m \rangle$ to zero. On the other hand, given $p \in k[x_1, \dots, x_n, y_1, \dots, y_m]$, if we replace each y_i in p by $f_i - (f_i - y_i)$ and expand, we obtain

$$\begin{aligned} p(x_1, \dots, x_n, y_1, \dots, y_m) = p(x_1, \dots, x_n, f_1, \dots, f_m) \\ + B_1 \cdot (f_1 - y_1) + \cdots + B_m \cdot (f_m - y_m) \end{aligned}$$

for some $B_1, \dots, B_m \in k[x_1, \dots, x_n, y_1, \dots, y_m]$ (see Exercise 4 for the details). In particular, if $p(x_1, \dots, x_n, f_1, \dots, f_m) = 0$, then

$$p(x_1, \dots, x_n, y_1, \dots, y_m) = B_1 \cdot (f_1 - y_1) + \cdots + B_m \cdot (f_m - y_m) \in J_F.$$

This completes the proof of (2).

Now intersect each side of (2) with $k[y_1, \dots, y_m]$. For $p \in k[y_1, \dots, y_m]$, this proves

$$p \in J_F \cap k[y_1, \dots, y_m] \iff p(f_1, \dots, f_m) = 0 \text{ in } k[x_1, \dots, x_n],$$

so that $J_F \cap k[y_1, \dots, y_m] = I_F$ by the definition of I_F Thus, (i) is proved and (ii) is then an immediate consequence of the elimination theory of Chapter 3 (see Theorem 2 and Exercise 5 of Chapter 3, §1). □

We can use this proposition to compute the relations between generators.

Example 4. In §2 we saw that the invariants of $C_2 = \{\pm I_2\} \subset \mathrm{GL}(2, k)$ are given by $k[x, y]^{C_2} = k[x^2, y^2, xy]$. Let $F = (x^2, y^2, xy)$ and let the new variables be u, v, w. Then the ideal of relations is obtained by eliminating x, y from the equations

$$\begin{aligned} u &= x^2, \\ v &= y^2, \\ w &= xy. \end{aligned}$$

If we use lex order with $x > y > u > v > w$, then a Groebner basis for the ideal $J_F = \langle u - x^2, v - y^2, w - xy \rangle$ consists of the polynomials

$$x^2 - u, xy - w, xv - yw, xw - yu, y^2 - v, uv - w^2.$$

It follows from Proposition 3 that

$$I_F = \langle uv - w^2 \rangle.$$

This says that all relations between x^2, y^2, and xy are generated by the obvious relation $x^2 \cdot y^2 = (xy)^2$. Then Proposition 2 shows that the ring of invariants can be written as

$$k[x, y]^{C_2} \cong k[u, v, w]/\langle uv - w^2 \rangle.$$

Example 5. In §3, we studied the cyclic matrix group $C_4 \subset \mathrm{GL}(2, k)$ generated by

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

and we saw that

$$k[x, y]^{C_4} = k[x^2 + y^2, x^3y - xy^3, x^2y^2].$$

Putting $F = (x^2 + y^2, x^3y - xy^3, x^2y^2)$, we leave it as an exercise to show that $I_F \subset k[u, v, w]$ is given by $I_F = \langle u^2w - v^2 - 4w^2 \rangle$. So the one nontrivial relation between the invariants is

$$(x^2 + y^2)^2 \cdot x^2y^2 = (x^3y - xy^3)^2 + 4(x^2y^2)^2.$$

By Proposition 2, we conclude that the ring of invariants can be written as

$$k[x, y]^{C_4} \cong k[u, v, w]/\langle u^2w - v^2 - 4w^2 \rangle.$$

By combining Propositions 1, 2, and 3 with the theory developed in §3 of Chapter 5, we can solve the *uniqueness problem* stated at the end of §2. Suppose that $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m]$ and let $I_F \subset k[y_1, \ldots, y_m]$ be the ideal of relations. If $I_F \neq \{0\}$, we know that a given element $f \in k[x_1, \ldots, x_n]^G$ can be written in more than one way in terms of $f_1, \ldots, f_m$. Is there a consistent choice for how to write f?

To solve this problem, pick a monomial order on $k[y_1, \ldots, y_m]$ and use Proposition 3 to find a Groebner basis G of I_F. Given $g \in k[y_1, \ldots, y_m]$, let $\overline{g}^G$ be the remainder of g on division by G. In Chapter 5, we showed that the remainders $\overline{g}^G$ uniquely represent

elements of the quotient ring $k[y_1, \ldots, y_m]/I_F$ (see Proposition 1 of Chapter 5, §3). Using this together with the isomorphism

$$k[y_1, \ldots, y_m]/I_F \cong k[x_1, \ldots, x_n]^G$$

of Proposition 2, we get a consistent method for writing elements of $k[x_1, \ldots, x_n]^G$ in terms of $f_1, \ldots, f_m$. Thus, Groebner basis methods help restore the uniqueness lost when $I_F \neq \{0\}$.

So far in this section, we have explored the algebra associated with the ideal of relations I_F. It is now time to turn to the geometry. The basic geometric object associated with an ideal is its variety. Hence, we get the following definition.

Definition 6. *If $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m]$, let $I_F \subset k[y_1, \ldots, y_m]$ be the ideal of relations for $F = (f_1, \ldots, f_m)$. Then we have the affine variety*

$$V_F = \mathbf{V}(I_F) \subset k^m.$$

The variety V_F has the following properties.

Proposition 7. *Let I_F and V_F be as in Definition 6.*

(i) *V_F is the smallest variety in k^m containing the parametrization*

$$\begin{aligned} y_1 &= f_1(x_1, \ldots, x_n), \\ &\vdots \\ y_m &= f_m(x_1, \ldots, x_n). \end{aligned}$$

(ii) *$I_F = \mathbf{I}(V_F)$, so that I_F is the ideal of all polynomial functions vanishing on V_F.*

(iii) *V_F is an irreducible variety.*

(iv) *Let $k[V_F]$ be the coordinate ring of V_F as defined in §4 of Chapter 5. Then there is a ring isomorphism*

$$k[V_F] \cong k[x_1, \ldots, x_n]^G.$$

Proof. Let $J_F = \langle f_1 - y_1, \ldots, f_m - y_m \rangle$. By Proposition 3, I_F is the n-th elimination ideal of J_F. Then part (i) follows immediately from the Polynomial Implicitization Theorem of Chapter 3 (see Theorem 1 of Chapter 3, §3).

Turning to (ii), note that we always have $I_F \subset \mathbf{I}(\mathbf{V}(I_F)) = \mathbf{I}(V_F)$. To prove the opposite inclusion, suppose that $h \in \mathbf{I}(V_F)$. Given any point $(a_1, \ldots, a_n) \in k^n$, part (i) implies that

$$(f_1(a_1, \ldots, a_n), \ldots, f_m(a_1, \ldots, a_n)) \in V_F.$$

Since h vanishes on V_F, it follows that

$$h(f_1(a_1, \ldots, a_n), \ldots, f_m(a_1, \ldots, a_n)) = 0$$

for all $(a_1, \ldots, a_n) \in k^n$. By assumption, k has characteristic zero and, hence, is infinite. Then Proposition 5 of Chapter 1, §1 implies that $h(f_1, \ldots, f_m) = 0$ and, hence, $h \in I_F$.

By (ii) and Proposition 1, $\mathbf{I}(V_F) = I_F$ is a prime ideal, so that V_F is irreducible by Proposition 4 of Chapter 5, §1. (We can also use the parametrization and Proposition 5 of Chapter 4, §5 to give a second proof that V_F is irreducible.)

Finally, in Chapter 5, we saw that the coordinate ring $k[V_F]$ could be written as

$$k[V_F] \cong k[y_1, \ldots, y_m]/\mathbf{I}(V_F)$$

(see Theorem 7 of Chapter 5, §2). Since $\mathbf{I}(V_F) = I_F$ by part (ii), we can use the isomorphism of Proposition 2 to obtain

$$k[V_F] \cong k[y_1, \ldots, y_m]/I_F \cong k[x_1, \ldots, x_n]^G. \tag{3}$$

This completes the proof of the proposition. □

Note how the isomorphisms in (3) link together the three methods (coordinate rings, quotient rings and rings of invariants) that we have learned for creating new rings.

When we write $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m]$, note that $f_1, \ldots, f_m$ are not uniquely determined. So one might ask how changing to a different set of generators affects the variety V_F. The answer is as follows.

Corollary 8. *Suppose that* $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m] = k[f'_1, \ldots, f'_{m'}]$. *If we set* $F = (f_1, \ldots, f_m)$ *and* $F' = (f'_1, \ldots, f'_{m'})$, *then the varieties* $V_F \subset k^m$ *and* $V_{F'} \subset k^{m'}$ *are isomorphic (as defined in Chapter 5, §4).*

Proof. Applying Proposition 7 twice, we then have isomorphisms $k[V_F] \cong k[x_1, \ldots, x_n]^G \cong k[V_{F'}]$, and it is easy to see that these isomorphisms are the identity on constants. But in Theorem 9 of Chapter 5, §4, we learned that two varieties are isomorphic if and only if there is an isomorphism of their coordinate rings which is the identity on constants. The corollary follows immediately. □

One of the lessons we learned in Chapter 4 was that the algebra–geometry correspondence works best over an algebraically closed field k. So for the rest of this section we will assume that k is algebraically closed.

To uncover the geometry of V_F, we need to think about the matrix group $G \subset \mathrm{GL}(n, k)$ more geometrically. So far, we have used G to act on polynomials: if $f(\mathbf{x}) \in k[x_1, \ldots, x_n]$, then a matrix $A \in G$ gives us the new polynomial $g(\mathbf{x}) = f(A \cdot \mathbf{x})$. But we can also let G act on the underlying affine space k^n. We will write a point $(a_1, \ldots, a_n) \in k^n$ as a column vector $\mathbf{a}$. Thus,

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Then a matrix $A \in G$ gives us the new point $A \cdot \mathbf{a}$ by matrix multiplication.

We can then use G to describe an equivalence relation on k^n: given $\mathbf{a}, \mathbf{b} \in k^n$, we say that $\mathbf{a} \sim_G \mathbf{b}$ if $\mathbf{b} = A \cdot \mathbf{a}$ for some $A \in G$. We leave it as an exercise to verify that $\sim_G$ is indeed an equivalence relation. It is also straightforward to check that the equivalence class of $\mathbf{a} \in k^n$ is given by

$$\{\mathbf{b} \in k^n : \mathbf{b} \sim_G \mathbf{a}\} = \{A \cdot \mathbf{a} : A \in G\}.$$

These equivalence classes have a special name.

Definition 9. *Given a finite matrix group $G \subset \mathrm{GL}(n, k)$ and $\mathbf{a} \in k^n$, the G-***orbit** *of $\mathbf{a}$ is the set*

$$G \cdot \mathbf{a} = \{A \cdot \mathbf{a} : A \in G\}.$$

The set of all G-orbits in k^n is denoted k^n/G and is called the **orbit space**.

Note that an orbit $G \cdot \mathbf{a}$ has at most $|G|$ elements. In the exercises, you will show that the number of elements in an orbit is always a divisor of $|G|$.

Since orbits are equivalence classes, it follows that the orbit space k^n/G is the set of equivalence classes of $\sim_G$. Thus, we have constructed k^n/G as a set. But for us, the objects of greatest interest are affine varieties. So it is natural to ask if k^n/G has the structure of a variety in some affine space. The answer is as follows.

Theorem 10. *Let $G \subset \mathrm{GL}(n, k)$ be a finite matrix group, where k is algebraically closed. Suppose that $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m]$. Then:*

(i) *The polynomial mapping $F : k^n \to V_F$ defined by $F(\mathbf{a}) = (f_1(\mathbf{a}), \ldots, f_m(\mathbf{a}))$ is surjective. Geometrically, this means that the parametrization $y_i = f_i(x_1, \ldots, x_n)$ covers all of V_F.*

(ii) *The map sending the G-orbit $G \cdot \mathbf{a} \subset k^n$ to the point $F(\mathbf{a}) \in V_F$ induces a one-to-one correspondence*

$$k^n/G \cong V_F.$$

Proof. We prove part (i) using elimination theory. Let $J_F = \langle f_1 - y_1, \ldots, f_m - y_m \rangle$ be the ideal defined in Proposition 3. Since $I_F = J_F \cap k[y_1, \ldots, y_m]$ is an elimination ideal of J_F, it follows that a point $(b_1, \ldots, b_m) \in V_F = \mathbf{V}(I_F)$ is a partial solution of the system of equations

$$\begin{aligned} y_1 &= f_1(x_1, \ldots, x_n), \\ &\vdots \\ y_m &= f_m(x_1, \ldots, x_n). \end{aligned}$$

If we can prove that $(b_1, \ldots, b_m) \in \mathbf{V}(I_F)$ extends to $(a_1, \ldots, a_n, b_1, \ldots, b_m) \in \mathbf{V}(J_F)$, then $F(a_1, \ldots, a_n) = (b_1, \ldots, b_m)$ and the surjectivity of $F : k^n \to V_F$ will follow.

We claim that for each i, there is an element $p_i \in J_F \cap k[x_i, \ldots, x_n, y_1, \ldots, y_m]$ such that

$$p_i = x_i^N + \text{terms in which } x_i \text{ has degree } < N, \tag{4}$$

where $N = |G|$. For now, we will assume that the claim is true.

Suppose that inductively we have extended $(b_1, \ldots, b_m)$ to a partial solution

$$(a_{i+1}, \ldots, a_n, b_1, \ldots, b_m) \in \mathbf{V}(J_F \cap k[x_{i+1}, \ldots, x_n, y_1, \ldots, y_m]).$$

Since k is algebraically closed, the Extension Theorem of Chapter 3, §1 asserts that we can extend to $(a_i, a_{i+1}, \ldots, a_n, b_1, \ldots, b_m)$, provided the leading coefficient in x_i of one of the generators of $J_F \cap k(x_i, \ldots, x_n, y_1, \ldots, y_m]$ does not vanish at the partial solution. Because of our claim, this ideal contains the above polynomial p_i and we can assume that p_i is a generator (just add it to the generating set). By (4), the leading coefficient is 1, which never vanishes, so that the required a_i exists (see Corollary 4 of Chapter 3, §1).

It remains to prove the existence of p_i. We will need the following lemma.

Lemma 11. *Suppose that $G \subset \mathrm{GL}(n, k)$ is a finite matrix group and $f \in k[x_1, \ldots, x_n]$. Let $N = |G|$. Then there are invariants $g_1, \ldots, g_N \in k[x_1, \ldots, x_n]^G$ such that*

$$f^N + g_1 f^{N-1} + \cdots + g_N = 0.$$

Proof. Consider the polynomial $\prod_{A \in G}(X - f(A \cdot \mathbf{x}))$. If we multiply it out, we get

$$\prod_{A \in G}(X - f(A \cdot \mathbf{x})) = X^N + g_1(\mathbf{x})X^{N-1} + \cdots + g_N(\mathbf{x}),$$

where the coefficients $g_1, \ldots, g_N$ are in $k[x_1, \ldots, x_n]$. We claim that $g_1, \ldots, g_N$ are invariant under G. To prove this, suppose that $B \in G$. In the proof of Proposition 3 of §3, we saw that the $f(AB \cdot \mathbf{x})$ are just the $f(A \cdot \mathbf{x})$, possibly in a different order. Thus

$$\prod_{A \in G}(X - f(AB \cdot \mathbf{x})) = \prod_{A \in G}(X - f(A \cdot \mathbf{x})),$$

and then multiplying out each side implies that

$$X^N + g_1(B \cdot \mathbf{x})X^{N-1} + \cdots + g_N(B \cdot \mathbf{x}) = X^N + g_1(\mathbf{x})X^{N-1} + \cdots + g_N(\mathbf{x})$$

for each $B \in G$. This proves that $g_1, \ldots, g_N \in k[x_1, \ldots, x_n]^G$.

Since one of the factors is $X - f(I_n \cdot \mathbf{x}) = X - f(\mathbf{x})$, the polynomial vanishes when $X = f$, and the lemma is proved. □

We can now prove our claim about the polynomial p_i. If we let $f = x_i$ in Lemma 11, then we get

$$x_i^N + g_1 x_i^{N-1} + \cdots + g_N = 0, \tag{5}$$

where $N = |G|$ and $g_1, \ldots, g_N \in k[x_1, \ldots, x_n]^G$. Using $k[x_1, \ldots, x_n]^G = k[f_1, \ldots, f_m]$, we can write $g_j = h_j(f_1, \ldots, f_m)$ for $j = 1, \ldots, N$. Then let

$$p_i(x_i, y_1, \ldots, y_m) = x_i^N + h_1(y_1, \ldots, y_m)x_i^{N-1} + \cdots + h_N(y_1, \ldots, y_m)$$

in $k[x_i, y_1, \ldots, y_m]$. From (5), it follows that $p(x_i, f_1, \ldots, f_m) = 0$ and, hence, by (2), we see that $p_i \in J_F$. Then $p_i \in J_F \cap k[x_i, \ldots, x_n, y_1, \ldots, y_m]$, and our claim is proved.

To prove (ii), first note that the map

$$\tilde{F} : k^n/G \to V_F$$

defined by sending $G \cdot \mathbf{a}$ to $F(\mathbf{a}) = (f_1(\mathbf{a}), \ldots, f_m(\mathbf{a}))$ is well-defined since each f_i is invariant and, hence, takes the same value on all points of a G-orbit $G \cdot \mathbf{a}$. Furthermore, F is onto by part (i) and it follows that $\tilde{F}$ is also onto.

It remains to show that $\tilde{F}$ is one-to-one. Suppose that $G \cdot \mathbf{a}$ and $G \cdot \mathbf{b}$ are distinct orbits. Since $\sim_G$ is an equivalence relation, it follows that the orbits are disjoint. We will construct an invariant $g \in k[x_1, \ldots, x_n]^G$ such that $g(\mathbf{a}) \neq g(\mathbf{b})$. To do this, note that $S = G \cdot \mathbf{b} \cup G \cdot \mathbf{a} - \{\mathbf{a}\}$ is a finite set of points in k^n and, hence, is an affine variety. Since $\mathbf{a} \notin S$, there must be some defining equation f of S which does not vanish at $\mathbf{a}$. Thus, for $A \in G$, we have

$$f(A \cdot \mathbf{b}) = 0 \quad \text{and} \quad f(A \cdot \mathbf{a}) = \begin{cases} 0 & \text{if} A \cdot \mathbf{a} \neq \mathbf{a} \\ f(\mathbf{a}) \neq 0 & \text{if} A \cdot \mathbf{a} = \mathbf{a}. \end{cases}$$

Then let $g = R_G(f)$. We leave it as an exercise to check that

$$g(\mathbf{b}) = 0 \text{ and } g(\mathbf{a}) = \frac{M}{|G|} f(\mathbf{a}) \neq 0,$$

where M is the number of elements $A \in G$ such that $A \cdot \mathbf{a} = \mathbf{a}$. We have thus found an element $g \in k[x_1, \ldots, x_n]^G$ such that $g(\mathbf{a}) \neq g(\mathbf{b})$.

Now write g as a polynomial $g = h(f_1, \ldots, f_m)$ in our generators. Then $g(\mathbf{a}) \neq g(\mathbf{b})$ implies that $f_i(\mathbf{a}) \neq f_i(\mathbf{b})$ for some i, and it follows that $\tilde{F}$ takes different values on $G \cdot \mathbf{a}$ and $G \cdot \mathbf{b}$. The theorem is now proved. □

Theorem 10 shows that there is a bijection between the *set* k^n/G and the *variety* V_F. This is what we mean by saying that k^n/G has the structure of an affine variety. Further, whereas I_F depends on the generators chosen for $k[x_1, \ldots, x_n]^G$, we noted in Corollary 8 that V_F is unique up to isomorphism. This implies that the variety structure on k^n/G is unique up to isomorphism.

One nice consequence of Theorem 10 and Proposition 7 is that the "polynomial functions" on the orbit space k^n/G are given by

$$k[V_F] \cong k[x_1, \ldots, x_n]^G.$$

Note how natural this is: an invariant polynomial takes the same value on all points of the G-orbit and, hence, defines a function on the orbit space. Thus, it is reasonable to expect that $k[x_1, \ldots, x_n]^G$ should be the "coordinate ring" of whatever variety structure we put on k^n/G.

Still, the bijection $k^n/G \cong V_F$ is rather remarkable if we look at it slightly differently. Suppose that we start with the geometric action of G on k^n which sends $\mathbf{a}$ to $A \cdot \mathbf{a}$ for $A \in G$. From this, we construct the orbit space k^n/G as the set of orbits. To give this set the structure of an affine variety, look at what we had to do:

- we made the action algebraic by letting G act on polynomials;
- we considered the invariant polynomials and found finitely many generators; and
- we formed the ideal of relations among the generators.

The equations coming from this ideal define the desired variety structure V_F on k^n/G.

In general, an important problem in algebraic geometry is to take a set of interesting objects (G-orbits, lines tangent to a curve, etc.) and give it the structure of an affine (or projective—see Chapter 8) variety. Some simple examples will be given in the exercises.

EXERCISES FOR §4

1. Given $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$, let $I = \{g \in k[y_1, \ldots, y_m] : g(f_1, \ldots, f_m) = 0\}$.
   a. Prove that I is an ideal of $k[y_1, \ldots, y_m]$.
   b. If $f \in k[f_1, \ldots, f_m]$ and $f = g(f_1, \ldots, f_m)$ is one representation of f in terms of $f_1, \ldots, f_m$, prove that all such representations are given by $f = g(f_1, \ldots, f_m) + h(f_1, \ldots, f_m)$ as h varies over I.
2. Let $f_1, \ldots, f_m, \in k[x_1, \ldots, x_n]$ and let $I \subset k[y_1, \ldots, y_m]$ be the ideal of relations defined in Exercise 1.
   a. Prove that the map sending a coset $[g]$ to $g(f_1, \ldots, f_m)$ defines a well-defined ring homomorphism
   $$\phi : k[y_1, \ldots, y_m]/I \longrightarrow k[f_1, \ldots, f_m].$$
   b. Prove that the map ϕ of part (a) is one-to-one and onto. Thus ϕ is a ring isomorphism.
   c. Use Exercise 13 in Chapter 5, §2 to give an alternate proof that $k[y_1, \ldots, y_m]/I$ and $k[f_1, \ldots, f_m]$ are isomorphic. Hint: Consider the ring homomorphism $\Phi\colon k[y_1, \ldots, y_m] \to k[f_1, \ldots, f_m]$ which sends y_i to f_i.
3. Although Propositions 1 and 2 were stated for $k[x_1, \ldots, x_n]^G$, we saw in Exercises 1 and 2 that these results held for any subring of $k[x_1, \ldots, x_n]$ of the form $k[f_1, \ldots, f_m]$. Give a similar generalization of Proposition 3. Does the proof given in the text need any changes?
4. Given $p \in k[x_1, \ldots, x_n, y_1, \ldots, y_m]$, prove that
   $$\begin{aligned} p(x_1, \ldots, x_n, y_1, \ldots, y_m) = {} & p(x_1, \ldots, x_n, f_1, \ldots, f_m) \\ & + B_1 \cdot (f_1 - y_1) + \cdots + B_m \cdot (f_m - y_m) \end{aligned}$$
   for some $B_1, \ldots, B_m \in k[x_1, \ldots, x_n, y_1, \ldots, y_m]$. Hint: In p, replace each occurrence of y_i by $f_i - (f_i - y_i)$. The proof is similar to the argument given to prove (4) in §1.
5. Complete Example 5 by showing that $I_F \subset k[u, v, w]$ is given by $I_F = \langle u^2 w - v^2 - 4w^2 \rangle$ when $F = (x^2 + y^2, x^3 y - xy^3, x^2 y^2)$.
6. In Exercise 7 of §3, you were asked to compute the invariants of a certain cyclic group $C_3 \subset \mathrm{GL}(2, k)$ of order 3. Take the generators you found for $k[x, y]^{C_3}$ and find the relations between them.
7. Repeat Exercise 6, this time using the cyclic group $C_6 \subset \mathrm{GL}(2, k)$ of order 6 from Exercise 8 of §3.

8. In Exercise 12 of §3, we listed four invariants f_1, f_2, f_3, f_4 of the group of rotations of the cube in $\mathbb{R}^3$.
   a. Using $(f_4/xyz)^2$ and part (c) of Exercise 12 of §3, find an algebraic relation between f_1, f_2, f_3, f_4.
   b. Show that there are no nontrivial algebraic relations between f_1, f_2, f_3.
   c. Show that the relation you found in part (a) generates the ideal of all relations between f_1, f_2, f_3, f_4. Hint: If $p(f_1, f_2, f_3, f_4) = 0$ is a relation, use part (a) to reduce to a relation of the form $p_1(f_1, f_2, f_3) + p_2(f_1, f_2, f_3)f_4 = 0$. Then explain how degree arguments imply $p_1(f_1, f_2, f_3) = 0$.
9. Given a finite matrix group $G \subset \mathrm{GL}(n, k)$, we defined the relation $\sim_G$ on k^n by $\mathbf{a} \sim_G \mathbf{b}$ if $b = A \cdot \mathbf{a}$ for some $A \in G$.
   a. Verify that $\sim_G$ is an equivalence relation.
   b. Prove that the equivalence class of $\mathbf{a}$ is the set $G \cdot \mathbf{a}$ defined in the text.
10. Consider the group of rotations of the cube in $\mathbb{R}^3$. We studied this group in Exercise 5 of §2, and we know that it has 24 elements.
    a. Draw a picture of the cube which shows orbits consisting of 1, 6, 8, 12 and 24 elements.
    b. Argue geometrically that there is no orbit consisting of four elements.
11. (Requires abstract algebra) Let $G \subset \mathrm{GL}(n, k)$ be a finite matrix group. In this problem, we will prove that the number of elements in an orbit $G \cdot \mathbf{a}$ divides $|G|$.
    a. Fix $\mathbf{a} \in k^n$ and let $H = \{A \in G : A \cdot \mathbf{a} = \mathbf{a}\}$. Prove that H is a subgroup of G. We call H the *isotropy subgroup* or *stabilizer* of $\mathbf{a}$.
    b. Given $A \in G$, we get the *left coset* $AH = \{AB : B \in H\}$ of H in G and we let G/H denote the set of all left cosets (note that G/H will not be a group unless H is normal). Prove that the map sending AH to $A \cdot \mathbf{a}$ induces a bijective map $G/H \cong G \cdot \mathbf{a}$. Hint: You will need to prove that the map is well-defined. Recall that two cosets AH and BH are equal if and only if $B^{-1}A \in H$.
    c. Use part (b) to prove that the number of elements in $G \cdot \mathbf{a}$ divides $|G|$.
12. As in the proof of Theorem 10, suppose that we have disjoint orbits $G \cdot \mathbf{a}$ and $G \cdot \mathbf{b}$. Set $S = G \cdot \mathbf{b} \cup G \cdot \mathbf{a} - \{\mathbf{a}\}$, and pick $f \in k[x_1, \ldots, x_n]$ such that $f = 0$ on all points of S but $f(\mathbf{a}) \neq 0$. Let $g = R_G(f)$, where R_G is the Reynolds operator of G.
    a. Explain why $g(\mathbf{b}) = 0$.
    b. Explain why $g(\mathbf{a}) = \frac{M}{|G|} f(\mathbf{a}) \neq 0$, where M is the number of elements $A \in G$ such that $A \cdot \mathbf{a} = \mathbf{a}$.
13. In this exercise, we will see how Theorem 10 can fail when we work over a field that is not algebraically closed. Consider the group of permutation matrices $S_2 \subset \mathrm{GL}(2, \mathbb{R})$.
    a. We know that $\mathbb{R}[x, y]^{S_2} = \mathbb{R}[\sigma_1, \sigma_2]$. Show that $I_F = \{0\}$ when $F = (\sigma_1, \sigma_2)$, so that $V_F = \mathbb{R}^2$. Thus, Theorem 10 is concerned with the map $\tilde{F} : \mathbb{R}^2/S_2 \to \mathbb{R}^2$ defined by sending $S_2 \cdot (x, y)$ to $(y_1, y_2) = (x + y, xy)$.
    b. Show that the image of $\tilde{F}$ is the set $\{(y_1, y_2) \in \mathbb{R}^2 : y_1^2 \geq 4y_2\} \subset \mathbb{R}^2$. This is the region lying below the parabola $y_1^2 = 4y_2$. Hint: Interpret y_1 and y_2 as coefficients of the quadratic $X^2 - y_1 X + y_2$. When does the quadratic have real roots?
14. There are many places in mathematics where one takes a set of equivalences classes and puts an algebraic structure on them. Show that the construction of a quotient ring $k[x_1, \ldots, x_n]/I$ is an example. Hint: See §2 of Chapter 5.
15. In this exercise, we will give some examples of how something initially defined as a set can turn out to be a variety in disguise. The key observation is that the set of nonvertical lines in the plane k^2 has a natural geometric structure. Namely, such a line L has a unique equation of the form $y = mx + b$, so that L can be identified with the point (m, b) in another

2-dimensional affine space, denoted $k^{2\vee}$. (If we use projective space—to be studied in the next chapter—then we can also include vertical lines.)

Now suppose that we have a curve C in the plane. Then consider all lines which are tangent to C somewhere on the curve. This gives us a subset $C^\vee \subset k^{2\vee}$. Let us compute this subset in some simple cases and show that it is an affine variety.

a. Suppose our curve C is the parabola $y = x^2$. Given a point (x_0, y_0) on the parabola, show that the tangent line is given by $y = 2x_0x - x_0^2$ and conclude that $C^\vee$ is the parabola $m^2 + 4b = 0$ in $k^{2\vee}$.
b. Show that $C^\vee$ is an affine variety when C is the cubic curve $y = x^3$.

In general, more work is needed to study $C^\vee$. In particular, the method used in the above examples breaks down when there are vertical tangents or singular points. Nevertheless, one can develop a satisfactory theory of what is called the *dual curve* $C^\vee$ of a curve $C \subset k^2$. One can also define the *dual variety* $V^\vee$ of a given irreducible variety $V \subset k^n$.