15. Symmetric polynomials

1. *The theorem*

Let S_n be the group of permutations of $\{1, \ldots, n\}$, also called the **symmetric group** on n things.

For indeterminates x_i, let $p \in S_n$ act on $\mathbb{Z}[x_1, \ldots, x_n]$ by

$$p(x_i) = x_{p(i)}$$

A polynomial $f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ is **invariant** under S_n if for all $p \in S_n$

$$f(p(x_1), \ldots, p(x_n)) = f(x_1, \ldots, x_n)$$

The **elementary symmetric polynomials** in $x_1, \ldots, x_n$ are

$$\begin{array}{lllll} s_1 & = & s_1(x_1, \ldots, x_n) & = & \sum_i x_i \\ s_2 & = & s_2(x_1, \ldots, x_n) & = & \sum_{i<j} x_i x_j \\ s_3 & = & s_3(x_1, \ldots, x_n) & = & \sum_{i<j<k} x_i x_j x_k \\ s_4 & = & s_4(x_1, \ldots, x_n) & = & \sum_{i<j<k<\ell} x_i x_j x_k x_\ell \\ & \ldots & & & \\ s_t & = & s_t(x_1, \ldots, x_n) & = & \sum_{i_1<i_2<\ldots<i_t} x_{i_1} x_{i_2} \ldots x_{i_t} \\ & \ldots & & & \\ s_n & = & s_n(x_1, \ldots, x_n) & = & x_1 x_2 x_3 \ldots x_n \end{array}$$

[1.0.1] **Theorem:** A polynomial $f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ is invariant under S_n if and only if it is a polynomial in the *elementary* symmetric functions $s_1, \ldots, s_n$.

[1.0.2] **Remark:** In fact, the proof shows an algorithm which determines the expression for a given S_n-invariant polynomial in terms of the elementary ones.

Proof: Let $f(x_1, \ldots, x_n)$ be S_n-invariant. Let

$$q : \mathbb{Z}[x_1, \ldots, x_{n-1}, x_n] \longrightarrow \mathbb{Z}[x_1, \ldots, x_{n-1}]$$

be the map which kills off x_n, that is

$$q(x_i) = \begin{cases} x_i & (1 \le i < n) \\ 0 & (i = n) \end{cases}$$

If $f(x_1, \ldots, x_n)$ is S_n-invariant, then

$$q(f(x_1, \ldots, x_{n-1}, x_n)) = f(x_1, \ldots, x_{n-1}, 0)$$

is S_{n-1}-invariant, where we take the copy of S_{n-1} inside S_n that fixes n. And note that

$$q(s_i(x_1, \ldots, x_n)) = \begin{cases} s_i(x_1, \ldots, x_{n-1}) & (1 \le i < n) \\ 0 & (i = n) \end{cases}$$

By induction on the number of variables, there is a polynomial P in $n-1$ variables such that

$$q(f(x_1, \ldots, x_n)) = P(s_1(x_1, \ldots, x_{n-1}), \ldots, s_{n-1}(x_1, \ldots, x_{n-1}))$$

Now use the same polynomial P but with the elementary symmetric functions augmented by insertion of x_n, by

$$g(x_1, \ldots, x_n) = P(s_1(x_1, \ldots, x_n), \ldots, s_{n-1}(x_1, \ldots, x_n))$$

By the way P was chosen,

$$q(f(x_1, \ldots, x_n) - g(x_1, \ldots, x_n)) = 0$$

That is, mapping $x_n \longrightarrow 0$ sends the difference $f - g$ to 0. Using the unique factorization in $\mathbb{Z}[x_1, \ldots, x_n]$, this implies that x_n divides $f - g$. The S_n-invariance of $f - g$ implies that every x_i divides $f - g$. That is, by unique factorization, $s_n(x_1, \ldots, x_n)$ divides $f - g$.

The **total degree** of a monomial $c\, x_1^{e_1} \ldots x_n^{e_n}$ is the sum of the exponents

$$\text{total degree } (c\, x_1^{e_1} \ldots x_n^{e_n}) = e_1 + \ldots + e_n$$

The total degree of a polynomial is the maximum of the total degrees of its monomial summands.

Consider the polynomial

$$\frac{f-g}{s_n} = \frac{f(x_1, \ldots, x_n) - g(x_1, \ldots, x_n)}{s_n(x_1, \ldots, x_n)}$$

It is of lower total degree than the original f. By induction on total degree $(f-g)/s_n$ is expressible in terms of the elementary symmetric polynomials in $x_1, \ldots, x_n$. ///

[1.0.3] **Remark:** The proof also shows that if the total degree of an S_n-invariant polynomial $f(x_1, \ldots, x_{n-1}, x_n)$ in n variables is less than or equal the number of variables, then the expression for $f(x_1, \ldots, x_{n-1}, 0)$ in terms of $s_i(x_1, \ldots, x_{n-1})$ gives the correct formula in terms of $s_i(x_1, \ldots, x_{n-1}, x_n)$.

---

2. *First examples*

[2.0.1] **Example:** Consider

$$f(x_1, \ldots, x_n) = x_1^2 + \ldots + x_n^2$$

The induction on n and the previous remark indicate that the general formula will be found if we find the formula for $n = 2$, since the total degree is 2. Let $q : \mathbb{Z}[x, y] \longrightarrow \mathbb{Z}[x]$ be the $\mathbb{Z}$-algebra map sending $x \longrightarrow x$ and $y \longrightarrow 0$. Then

$$q(x^2 + y^2) = x^2 = s_1(x)^2$$

Then, following the procedure of the proof of the theorem,

$$(x^2 + y^2) - s_1(x, y)^2 = (x^2 + y^2) - (x + y)^2 = -2xy$$

Dividing by $s_2(x, y) = xy$ we obtain -2. (This is visible, anyway.) Thus,

$$x^2 + y^2 = s_1(x, y)^2 - 2s_2(x, y)$$

The induction on the number of variables gives

$$x_1^2 + \ldots + x_n^2 = s_1(x_1, \ldots, x_n)^2 - s_2(x_1, \ldots, x_n)$$

[2.0.2] **Example:** Consider

$$f(x_1, \ldots, x_n) = \sum_i x_i^4$$

Since the total degree is 4, as in the remark just above it suffices to determine the pattern with just 4 variables x_1, x_2, x_3, x_4. Indeed, we start with just 2 variables. Following the procedure indicated in the theorem, letting q be the $\mathbb{Z}$-algebra homomorphism which sends y to 0,

$$q(x^4 + y^4) = x^4 = s_1(x)^4$$

so consider

$$(x^4 + y^4) - s_1(x, y)^4 = -4x^3y - 6x^2y^2 - 4xy^3 = -s_1(x, y) \cdot (4x^2 + 6xy + 4y^2)$$

The latter factor of lower total degree is analyzed in the same fashion:

$$q(4x^2 + 6xy + 4y^2) = 4x^2 = 4s_1(x)^2$$

so consider

$$(4x^2 + 6xy + 4y^2) - 4s_1(x, y)^2 = -2xy$$

Going backward,

$$x^4 + y^4 = s_1(x, y)^4 - s_1(x, y) \cdot (4s_1(x, y)^2 - 2s_2(x, y))$$

Passing to three variables,

$$q(x^4 + y^4 + z^4) = x^4 + y^4 = s_1(x, y)^4 - s_1(x, y) \cdot (4s_1(x, y)^2 - 2s_2(x, y))$$

so consider

$$(x^4 + y^4 + z^4) - \big(s_1(x, y, z)^4 - s_1(x, y, z) \cdot (4s_1(x, y, z)^2 - 2s_2(x, y, z))\big)$$

Before expanding this, dreading the 15 terms from the $(x + y + z)^4$, for example, recall that the only terms which will *not* be cancelled are those which involve *all* of x, y, z. Thus, this is

$$-12x^2yz - 12y^2xz - 12z^2xy + (xy + yz + zx) \cdot (4(x + y + z)^2 - 2(xy + yz + zx)) + (\text{irrelevant})$$

$$= -12x^2yz - 12y^2xz - 12z^2xy + (xy + yz + zx) \cdot (4x^2 + 4y^2 + 4z^2 + 6xy + 6yz + 6zx) + (\text{irrelevant})$$

$$= -12x^2yz - 12y^2xz - 12z^2xy + 4xyz^2 + 4yzx^2 + 4zxy^2 + 6xy^2z$$
$$+ 6x^2yz + 6x^2yz + 6xyz^2 + 6xy^2z + 6xyz^2$$

$$= 4xyz(x+y+z) = 4s_3(x,y,z) \cdot s_1(x,y,z)$$

Thus, with 3 variables,

$$x^4 + y^4 + z^4$$

$$= s_1(x,y,z)^4 - s_2(x,y,z) \cdot (4s_1(x,y,z)^2 - 2s_2(x,y,z)) + 4s_3(x,y,z) \cdot s_1(x,y,z)$$

Abbreviating $s_i = s_i(x,y,z,w)$, we anticipate that

$$x^4 + y^4 + z^4 + w^4 - \left(s_1^4 - 4s_1^2 s_2 + 2s_2^2 + 4s_1 s_3\right) = \text{ constant } \cdot xyzw$$

We can save a little time by evaluating the constant by taking $x = y = z = w = 1$. In that case

$$\begin{array}{rcl} s_1(1,1,1,1) & = & 4 \\ s_2(1,1,1,1) & = & 6 \\ s_3(1,1,1,1) & = & 4 \end{array}$$

and

$$1 + 1 + 1 + 1 - \left(4^4 - 4 \cdot 4^2 \cdot 6 + 2 \cdot 6^2 + 4 \cdot 4 \cdot 4\right) = \text{ constant}$$

or

$$\text{constant } = 4 - (256 - 384 + 72 + 64) = -4$$

Thus,

$$x^4 + y^4 + z^4 + w^4 = s_1^4 - 4s_1^2 s_2 + 2s_2^2 + 4s_1 s_3 - 4s_4$$

By the remark above, since the total degree is just 4, this shows that for arbitrary n

$$x_1^4 + \ldots + x_n^4 = s_1^4 - 4s_1^2 s_2 + 2s_2^2 + 4s_1 s_3 - 4s_4$$

3. *A variant: discriminants*

Let $x_1, \ldots, x_n$ be indeterminates. Their **discriminant** is

$$D = D(x_1, \ldots, x_n) = \prod_{i<j} (x_i - x_j)$$

Certainly the sign of D depends on the ordering of the indeterminates. But

$$D^2 = \prod_{i \neq j} (x_i - x_j)^2$$

is *symmetric*, that is, is *invariant* under all permutations of the x_i. Therefore, D^2 has an expression in terms of the elementary symmetric functions of the x_i.

[3.0.1] **Remark:** By contrast to the previous low-degree examples, the discriminant (squared) has as high a degree as possible.

[3.0.2] **Example:** With just 2 indeterminates x, y, we have the familiar

$$D^2 = (x-y)^2 = x^2 - 2xy + y^2 = (x+y)^2 - 4xy = s_1^2 - 4s_2$$

Rather than compute the general version in higher-degree cases, let's consider a more accessible variation on the question. Suppose that $\alpha_1, \ldots, \alpha_n$ are roots of an equation

$$X^n + aX + b = 0$$

in a field k, with $a, b \in k$. For simplicity suppose $a \neq 0$ and $b \neq 0$, since otherwise we have even simpler methods to study this equation. Let $f(X) = x^n + aX + b$. The discriminant

$$D(\alpha_1, \ldots, \alpha_n) = \prod_{i<j} (\alpha_i - \alpha_j)$$

vanishes if and only if any two of the α_i coincide. On the other hand, $f(X)$ has a repeated factor in $k[X]$ if and only if $\gcd(f, f') \neq 1$. Because of the sparseness of this polynomial, we can in effect execute the Euclidean algorithm explicitly. Assume that the characteristic of k does not divide $n(n-1)$. Then

$$(X^n + aX + b) - \frac{X}{n} \cdot (nX^{n-1} + a) = a(1 - \frac{1}{n})X + b$$

That is, any repeated factor of $f(X)$ divides $X + \frac{bn}{(n-1)a}$, and the latter linear factor divides $f'(X)$. Continuing, the remainder upon dividing $nX^{n-1} + a$ by the linear factor $X + \frac{bn}{(n-1)a}$ is simply the value of $nX^{n-1} + a$ obtained by evaluating at $\frac{-bn}{(n-1)a}$, namely

$$n\left(\frac{-bn}{(n-1)a}\right)^{n-1} + a = \left(n^n(-1)^{n-1}b^{n-1} + (n-1)^{n-1}a^n\right) \cdot ((n-1)a)^{1-n}$$

Thus, (constraining a to be non-zero)

$$n^n(-1)^{n-1}b^{n-1} + (n-1)^{n-1}a^n = 0$$

if and only if some $\alpha_i - \alpha_j = 0$.

We obviously want to say that with the constraint that all the symmetric functions of the α_i being 0 except the last two, we have computed the discriminant (up to a less interesting constant factor).

A relatively graceful approach would be to show that $R = \mathbb{Z}[x_1, \ldots, x_n]$ admits a *universal* $\mathbb{Z}$-algebra homomorphism $\varphi : R \longrightarrow \Omega$ for some ring Ω that sends the first $n-2$ elementary symmetric functions

$$\begin{array}{ccccc} s_1 & = & s_1(x_1, \ldots, x_n) & = & \sum_i x_i \\ s_2 & = & s_2(x_1, \ldots, x_n) & = & \sum_{i<j} x_i\, x_j \\ s_3 & = & s_3(x_1, \ldots, x_n) & = & \sum_{i<j<k} x_i\, x_j\, x_k \\ \ldots & & & & \\ s_\ell & = & s_\ell(x_1, \ldots, x_n) & = & \sum_{i_1<\ldots<i_\ell} x_{i_1} \ldots x_{i_\ell} \\ \ldots & & & & \\ s_{n-2} & = & s_{n-2}(x_1, \ldots, x_n) & = & \sum_{i_1<\ldots<i_{n-2}} x_{i_1} \ldots x_{i_{n-2}} \end{array}$$

to 0, but imposes no unnecessary further relations on the images

$$a = (-1)^{n-1}\varphi(s_{n-1}) \qquad b = (-1)^n \varphi(s_n)$$

We do not have sufficient apparatus to do this nicely at this moment. [1] Nevertheless, the computation above does tell us something.

Exercises

15.[3.0.1] Express $x_1^3 + x_2^3 + \ldots + x_n^3$ in terms of the elementary symmetric polynomials.

15.[3.0.2] Express $\sum_{i \neq j} x_i\, x_j^2$ in terms of the elementary symmetric polynomials.

15.[3.0.3] Let α, β be the roots of a quadratic equation $ax^2 + bx + c = 0$, Show that the *discriminant*, defined to be $(\alpha - \beta)^2$, is $b^2 - 4ac$.

15.[3.0.4] Consider $f(x) = x^3 + ax + b$ as a polynomial with coefficients in $k(a, b)$ where k is a field not of characteristic 2 or 3. By computing the greatest common divisor of f and f', give a condition for the roots of $f(x) = 0$ to be distinct.

15.[3.0.5] Express $\sum_{i,j,k \text{ distinct}} x_i\, x_j\, x_k^2$ in terms of elementary symmetric polynomials.

---

[1] The key point is that $\mathbb{Z}[x_1, \ldots, x_n]$ is *integral* over $\mathbb{Z}[s_1, s_2, \ldots, s_n]$ in the sense that each x_i is a root of the *monic* equation $X^n - s_1 X^{n-2} + s_2 X^{n-2} - \ldots + (-1)^{n-1} s_{n-1} X + (-1)^n s_n = 0$ It is true that for R an *integral extension* of a ring S, any homomorphism $\varphi_o : S \longrightarrow \Omega$ to an algebraically closed field Ω extends (probably in more than one way) to a homomorphism $\varphi : R \longrightarrow \Omega$. This would give us a justification for our hope that, given $a, b \in \Omega$ we can require that $\varphi_o(s_1) = \varphi_o(s_2) = \ldots = \varphi_o(s_{n-2}) = 0$ while $\varphi_o(s_{n-1}) = (-1)^{n-1} a \quad \varphi_o(s_n) = (-1)^n b$.