

## §38

### Symmetric Polynomials

Let  $D$  be an integral domain and let  $f(x_1, x_2, \dots, x_m) \in D[x_1, x_2, \dots, x_m]$ . For any permutation  $\sigma = \begin{pmatrix} 1 & 2 & \dots & m \\ i_1 & i_2 & \dots & i_m \end{pmatrix}$  in  $S_m$ , the value of  $f$  at  $(x_{i_1}, x_{i_2}, \dots, x_{i_m})$  is a polynomial  $f(x_{i_1}, x_{i_2}, \dots, x_{i_m})$  in  $D[x_1, x_2, \dots, x_m]$ , which we can shortly denote by  $f^\sigma$  (Definition 35.19). For example, if  $f(x, y, z) = x^2 + y^2 - xz$  in  $\mathbb{Z}[x, y, z]$ , then  $f(z, x, y) = z^2 + x^2 - zy$ ; and if  $g(x, y) = x^2 - xy + y^3$  in  $\mathbb{Z}[x, y]$ , then  $g(y, x) = y^2 - yx + x^3 \in \mathbb{Z}[x, y]$ . In general,  $f(x_{i_1}, x_{i_2}, \dots, x_{i_m})$  will be a polynomial distinct from  $f(x_1, x_2, \dots, x_m)$ .

**38.1 Definition:** Let  $D$  be an integral domain and let  $f(x_1, x_2, \dots, x_m)$  be a polynomial in  $D[x_1, x_2, \dots, x_m]$ . If  $f(x_{i_1}, x_{i_2}, \dots, x_{i_m}) = f(x_1, x_2, \dots, x_m)$  for all permutations  $\sigma = \begin{pmatrix} 1 & 2 & \dots & m \\ i_1 & i_2 & \dots & i_m \end{pmatrix}$  in  $S_m$ , then  $f = f(x_1, x_2, \dots, x_m)$  is called a *symmetric polynomial* in  $D[x_1, x_2, \dots, x_m]$ . We also say that  $f(x_1, x_2, \dots, x_m)$  is *symmetric* in the indeterminates  $x_1, x_2, \dots, x_m$ .

The polynomials  $x + y$ ,  $xy$ ,  $x^2 + y^2$ ,  $x^3 + y^3$  are symmetric polynomials in  $D[x, y]$ . Also, the polynomials  $x^2 + y^2 + z^2$  and  $xy + yz + zx$  are symmetric polynomials in  $D[x, y, z]$ .

The sum, difference and product of symmetric polynomials are symmetric polynomials. Indeed, if  $f(x_1, x_2, \dots, x_m)$  and  $g(x_1, x_2, \dots, x_m)$  are symmetric polynomials in  $D[x_1, x_2, \dots, x_m]$ , and if

$$h(x_1, x_2, \dots, x_m) = f(x_1, x_2, \dots, x_m) + g(x_1, x_2, \dots, x_m)$$

is their sum, then, for any permutation  $\begin{pmatrix} 1 & 2 & \dots & m \\ i_1 & i_2 & \dots & i_m \end{pmatrix}$  in  $S_m$ , we have

$$\begin{aligned} h(x_{i_1}, x_{i_2}, \dots, x_{i_m}) &= f(x_{i_1}, x_{i_2}, \dots, x_{i_m}) + g(x_{i_1}, x_{i_2}, \dots, x_{i_m}) \\ &= f(x_1, x_2, \dots, x_m) + g(x_1, x_2, \dots, x_m) \\ &= h(x_1, x_2, \dots, x_m), \end{aligned}$$

and so  $h(x_1, x_2, \dots, x_m)$  is a symmetric polynomial. The same argument works also when  $h = f - g$  and  $h = fg$ . This proves

**38.2 Lemma:** *Let  $D$  be an integral domain. The symmetric polynomials in  $D[x_1, x_2, \dots, x_m]$  form a subring of  $D[x_1, x_2, \dots, x_m]$ .  $\square$*

We introduce a new indeterminate  $t$  and consider the polynomial

$$f(t) = (t - x_1)(t - x_2)\dots(t - x_m) \text{ in } D[x_1, x_2, \dots, x_m][t].$$

We see that  $x_1, x_2, \dots, x_m \in D[x_1, x_2, \dots, x_m]$  are the roots of  $f(t)$ . We have

$$f(t) = t^m - \sigma_1(x_1, x_2, \dots, x_m)t^{m-1} + \sigma_2(x_1, x_2, \dots, x_m)t^{m-2} - + \dots + (-1)^m \sigma_m(x_1, x_2, \dots, x_m)$$

for some  $\sigma_1, \sigma_2, \dots, \sigma_m$  in  $D[x_1, x_2, \dots, x_m]$ . Since

$$\begin{aligned} f(t) &= (t - x_{i_1})(t - x_{i_2})\dots(t - x_{i_m}) \\ &= t^m - \sigma_1(x_{i_1}, x_{i_2}, \dots, x_{i_m})t^{m-1} + \sigma_2(x_{i_1}, x_{i_2}, \dots, x_{i_m})t^{m-2} - + \dots + (-1)^m \sigma_m(x_{i_1}, x_{i_2}, \dots, x_{i_m}) \end{aligned}$$

for any permutation  $\begin{pmatrix} 1 & 2 & \dots & m \\ i_1 & i_2 & \dots & i_m \end{pmatrix}$  in  $S_m$ , we have

$$\sigma_j(x_{i_1}, x_{i_2}, \dots, x_{i_m}) = \sigma_j(x_1, x_2, \dots, x_m) \quad \text{for all } j = 1, 2, \dots, m.$$

Thus  $\sigma_1, \sigma_2, \dots, \sigma_m$  are symmetric polynomials in  $D[x_1, x_2, \dots, x_m]$ .

**38.3 Definition:** Let  $D$  be an integral domain and let

$$\begin{aligned} &(t - x_1)(t - x_2)\dots(t - x_m) \\ &= t^m - \sigma_1(x_1, x_2, \dots, x_m)t^{m-1} + \sigma_2(x_1, x_2, \dots, x_m)t^{m-2} - + \dots + (-1)^m \sigma_m(x_1, x_2, \dots, x_m). \end{aligned}$$

The symmetric polynomials  $\sigma_1, \sigma_2, \dots, \sigma_m$  are called the *elementary symmetric polynomials* in  $D[x_1, x_2, \dots, x_m]$ .

By routine computation, we find the elementary symmetric polynomials explicitly. For example,

$$\sigma_1 = x + y, \quad \sigma_2 = xy \quad \text{in } D[x, y]$$

$$\sigma_1 = x + y + z, \quad \sigma_2 = xy + yz + zx, \quad \sigma_3 = xyz \quad \text{in } D[x,y,z]$$

$$\begin{aligned} \sigma_1 &= x + y + z + u, & \sigma_2 &= xy + xz + xu + yz + yu + zu, \\ \sigma_3 &= xyz + xyu + xzu + yzu, & \sigma_4 &= xyzu \end{aligned} \quad \text{in } D[x,y,z,u]$$

are the elementary symmetric polynomials.

Notice that  $(t - x_1)(t - x_2) \dots (t - x_m)$ , when multiplied out, is a sum of certain terms  $a_1 a_2 \dots a_m$ , where each  $a_i$  is either  $t$  or one of  $-x_1, -x_2, \dots, -x_m$ . The term  $(-1)^j \sigma_j(x_1, x_2, \dots, x_m) t^{m-j}$  is the sum of those  $a_1 a_2 \dots a_m$ 's for which exactly  $m - j$  of the  $a$ 's are equal to  $t$ . Hence  $(-1)^j \sigma_j(x_1, x_2, \dots, x_m)$  is the sum of all products  $b_1 b_2 \dots b_j$ , where  $b_1, b_2, \dots, b_j$  run independently over the set  $\{-x_1, -x_2, \dots, -x_m\}$ . In other words,  $\sigma_j(x_1, x_2, \dots, x_m)$  is the sum of all  $\binom{m}{j}$  products of  $x_1, x_2, \dots, x_m$ , taken  $j$  at a time. Thus

$$\begin{aligned} \sigma_1 &= \sum x_i \\ \sigma_2 &= \sum x_i x_j \\ \sigma_3 &= \sum x_i x_j x_k \\ &\dots\dots\dots \\ \sigma_m &= x_1 x_2 \dots x_m. \end{aligned}$$

Note that " $\sigma_j$ " stands for many polynomials.  $\sigma_j$  in  $D[x_1, x_2, \dots, x_m]$  is distinct from  $\sigma_j$  in  $D[x_1, x_2, \dots, x_n]$  when  $m \neq n$ . This ambiguity in notation will not cause any confusion if we pay attention to the number of indeterminates. When confusion is likely, we write  $\sigma_j(x_1, x_2, \dots, x_m)$  instead of  $\sigma_j$ .

Now  $\sigma_1, \sigma_2, \dots, \sigma_m$  are symmetric polynomials in  $D[x_1, x_2, \dots, x_m]$ , and, by repeated application of Lemma 38.2, we conclude that  $g(\sigma_1, \sigma_2, \dots, \sigma_m)$  is also a symmetric polynomial, where  $g$  is any polynomial in  $m$  indeterminates. Hence the set  $\{g(\sigma_1, \sigma_2, \dots, \sigma_m) : g \in D[u_1, u_2, \dots, u_m]\}$  consist only of symmetric polynomials. We will prove conversely that every symmetric polynomial is in this set (the subring of symmetric polynomials in  $D[x_1, x_2, \dots, x_m]$  is the subring of  $D[x_1, x_2, \dots, x_m]$  generated by  $\sigma_1, \sigma_2, \dots, \sigma_m$ ).

**38.4 Theorem (Fundamental theorem on symmetric polynomials):** *Let  $D$  be an integral domain and  $f(x_1, x_2, \dots, x_m)$  a symmetric poly-*

nomial in  $D[x_1, x_2, \dots, x_m]$ . Then there is a unique polynomial  $g(u_1, u_2, \dots, u_m)$  in  $D[u_1, u_2, \dots, u_m]$  such that  $f$  is the value of  $g$  at  $(\sigma_1, \sigma_2, \dots, \sigma_m)$ :

$$f(x_1, x_2, \dots, x_m) = g(\sigma_1, \sigma_2, \dots, \sigma_m) \in D[x_1, x_2, \dots, x_m].$$

Loosely speaking, every symmetric polynomial is a polynomial in the elementary symmetric polynomials  $\sigma_1, \sigma_2, \dots, \sigma_m$ . We introduced new indeterminates  $u_1, u_2, \dots, u_m$  in order to distinguish clearly between  $g$  and  $g(\sigma_1, \sigma_2, \dots, \sigma_m)$ .

For example,  $f(x, y) = x^2 + y^2 \in \mathbb{Z}[x, y]$  is a symmetric polynomial, and we have  $x^2 + y^2 = (x + y)^2 - 2xy = \sigma_1^2 - 2\sigma_2$ . Hence  $f(x, y) = g(\sigma_1, \sigma_2)$ , where  $g(u, v) = u^2 - 2v \in \mathbb{Z}[u, v]$ . Likewise, if  $f(x, y, z)$  is the symmetric polynomial  $x^2y + xy^2 + x^2z + xz^2 + y^2z + yz^2$  in  $\mathbb{Z}[x, y, z]$ , we have  $f(x, y, z) = (x + y + z)(xy + yz + zx) - 3xyz = \sigma_1\sigma_2 - \sigma_3$ . Thus  $f(x, y, z) = g(\sigma_1, \sigma_2, \sigma_3)$ , where  $g(u, v, w) = uv - 3w \in \mathbb{Z}[u, v, w]$ .

The proof of the fundamental theorem requires some preparation. First we need an ordering of  $m$ -tuples. Given any two  $m$ -tuples  $(r_1, r_2, \dots, r_m)$ ,  $(s_1, s_2, \dots, s_m)$  of nonnegative integers, we will say  $(r_1, r_2, \dots, r_m)$  is *higher than*  $(s_1, s_2, \dots, s_m)$ , or  $(s_1, s_2, \dots, s_m)$  is *lower than*  $(r_1, r_2, \dots, r_m)$  when  $r_1 > s_1$ . If  $r_1 = s_1$ , we will say  $(r_1, r_2, \dots, r_m)$  is *higher than*  $(s_1, s_2, \dots, s_m)$ , or  $(s_1, s_2, \dots, s_m)$  is *lower than*  $(r_1, r_2, \dots, r_m)$  when  $r_2 > s_2$ . If  $r_1 = s_1$  and  $r_2 = s_2$ , we will compare  $r_3$  and  $s_3$ , etc. This is very much like the ordering of words alphabetically, and will be referred to as the *alphabetical* or *lexicographical* ordering of  $m$ -tuples. Stated differently,  $(r_1, r_2, \dots, r_m)$  is higher than  $(s_1, s_2, \dots, s_m)$  if and only if the first nonzero difference among

$$r_1 - s_1, \quad r_2 - s_2, \quad \dots, \quad r_m - s_m$$

is positive. Clearly, if  $(r_1, r_2, \dots, r_m)$  is higher than  $(s_1, s_2, \dots, s_m)$  and  $(s_1, s_2, \dots, s_m)$  is higher than  $(t_1, t_2, \dots, t_m)$ , then  $(r_1, r_2, \dots, r_m)$  is higher than  $(t_1, t_2, \dots, t_m)$ .

Now let  $f$  be a polynomial in  $D[x_1, x_2, \dots, x_m]$ . So  $f$  is a sum of monomials  $ax_1^{k_1}x_2^{k_2}\dots x_m^{k_m}$ , where  $a \in D$  and  $(k_1, k_2, \dots, k_m)$  is an  $m$ -tuple of nonnegative integers. Here there may be several monomials  $ax_1^{k_1}x_2^{k_2}\dots x_m^{k_m}$ ,  $bx_1^{k_1}x_2^{k_2}\dots x_m^{k_m}$ ,  $cx_1^{k_1}x_2^{k_2}\dots x_m^{k_m}$ , etc. with the same exponent system  $(k_1, k_2, \dots, k_m)$ . In this case, we collect these monomials into a single one

$(a + b + c + \cdots)x_1^{k_1}x_2^{k_2}\cdots x_m^{k_m}$ . We assume this has been done for each of the exponent systems, so that each  $m$ -tuple  $(k_1, k_2, \dots, k_m)$  occurs as an exponent system of a monomial at most once. If, after this collection process, a monomial  $ax_1^{k_1}x_2^{k_2}\cdots x_m^{k_m}$  occurring in  $f$  has a nonzero coefficient  $a \in D$ , we will say that  $a$  *appears in*  $f$ .

Let us now assume  $f \neq 0$ . We order the monomials appearing in  $f$  by the alphabetical ordering of their exponent systems. First we write the monomial appearing in  $f$  whose exponent system is highest (i.e., higher than the exponent systems of all other monomials appearing in  $f$ ). Among the remaining monomials appearing in  $f$ , we find the one with the highest exponent system and write it in the second place. Among the remaining monomials appearing in  $f$ , the one with the highest exponent system will be written it in the third place, and so on. In this ordering of monomials, the one that is written in the first place, that is to say, the one with the highest exponent system will be called the *leading monomial* of the nonzero polynomial  $f \in D[x_1, x_2, \dots, x_m]$ . Note that the coefficients of monomials play no role in this ordering. Only the exponent systems are relevant.

For instance,  $f(x, y, z) = xz^5 + z^7 + 2x^3 + 5x^2y + 100x^2y^2 - x^2y^2z \in \mathbb{Z}[x, y, z]$  will be written as  $2x^3 - x^2y^2z + 100x^2y^2 + 5x^2y + xz^5 + z^7$  when we order the monomials in the described manner. The leading monomial of  $f(x, y, z)$  is  $2x^3$ .

**38.5 Lemma:** *Let  $D$  be an integral domain and  $f, g \in D[x_1, x_2, \dots, x_m] \setminus \{0\}$ . If  $ax_1^{k_1}x_2^{k_2}\cdots x_m^{k_m}$  is the leading monomial of  $f$  and  $bx_1^{n_1}x_2^{n_2}\cdots x_m^{n_m}$  is the leading monomial of  $g$ , then  $abx_1^{k_1+n_1}x_2^{k_2+n_2}\cdots x_m^{k_m+n_m}$  is the leading monomial of  $fg$ .*

**Proof:** By hypothesis,  $a \neq 0$ ,  $b \neq 0$ , so  $ab \neq 0$ . Now  $fg \neq 0$  and  $fg$  is the sum of all products  $(cx_1^{r_1}x_2^{r_2}\cdots x_m^{r_m})(dx_1^{s_1}x_2^{s_2}\cdots x_m^{s_m})$ , where  $cx_1^{r_1}x_2^{r_2}\cdots x_m^{r_m}$  and  $dx_1^{s_1}x_2^{s_2}\cdots x_m^{s_m}$  run through all monomials appearing in  $f$  and  $g$ , respectively. We contend that, among all these products, the highest exponent system is  $(k_1 + n_1, k_2 + n_2, \dots, k_m + n_m)$ , and that this exponent

system arises only from the product  $(ax_1^{k_1}x_2^{k_2}\dots x_m^{k_m})(bx_1^{n_1}x_2^{n_2}\dots x_m^{n_m})$ .

This will imply

$$fg = abx_1^{k_1+n_1}x_2^{k_2+n_2}\dots x_m^{k_m+n_m} + [\text{a sum of monomials, each with an exponent system lower than } (k_1 + n_1, k_2 + n_2, \dots, k_m + n_m)],$$

and, since  $ab \neq 0$ , the leading monomial of  $fg$  will be equal to  $abx_1^{k_1+n_1}x_2^{k_2+n_2}\dots x_m^{k_m+n_m}$ .

To prove our contention, let  $cx_1^{r_1}x_2^{r_2}\dots x_m^{r_m}$  be a monomial appearing in  $f$  and let  $dx_1^{s_1}x_2^{s_2}\dots x_m^{s_m}$  be one appearing in  $g$ , but assume that either  $cx_1^{r_1}x_2^{r_2}\dots x_m^{r_m}$  is distinct from  $ax_1^{k_1}x_2^{k_2}\dots x_m^{k_m}$  or  $dx_1^{s_1}x_2^{s_2}\dots x_m^{s_m}$  is distinct from  $bx_1^{n_1}x_2^{n_2}\dots x_m^{n_m}$ . We are to show that the exponent system  $(r_1 + s_1, r_2 + s_2, \dots, r_m + s_m)$  is lower than  $(k_1 + n_1, k_2 + n_2, \dots, k_m + n_m)$ . Now  $(r_1, r_2, \dots, r_m)$  is lower than  $(k_1, k_2, \dots, k_m)$  or equal to it, and  $(s_1, s_2, \dots, s_m)$  is lower than  $(n_1, n_2, \dots, n_m)$  or equal to it, but the case of simultaneous equality is excluded. Hence the first nonzero integer in

$$k_1 - r_1, k_2 - r_2, \dots, k_m - r_m$$

is positive, or  $(k_1, k_2, \dots, k_m) = (r_1, r_2, \dots, r_m)$ , and the first nonzero integer in

$$n_1 - s_1, n_2 - s_2, \dots, n_m - s_m$$

is positive, or  $(n_1, n_2, \dots, n_m) = (s_1, s_2, \dots, s_m)$ . Since simultaneous equality is excluded, there are nonzero integers in

$$(k_1 - r_1) + (n_1 - s_1), (k_2 - r_2) + (n_2 - s_2), \dots, (k_m - r_m) + (n_m - s_m)$$

and the first of them, being a sum of two positive integers or a sum of a positive integer and zero, is certainly positive. This means that

$$(k_1 + n_1, k_2 + n_2, \dots, k_m + n_m) \text{ is higher than } (r_1 + s_1, r_2 + s_2, \dots, r_m + s_m).$$

Since the exponent system  $(k_1 + n_1, k_2 + n_2, \dots, k_m + n_m)$  does arise from the product  $(ax_1^{k_1}x_2^{k_2}\dots x_m^{k_m})(bx_1^{n_1}x_2^{n_2}\dots x_m^{n_m})$ , it is indeed the highest exponent system of all the products  $(cx_1^{r_1}x_2^{r_2}\dots x_m^{r_m})(dx_1^{s_1}x_2^{s_2}\dots x_m^{s_m})$  where  $cx_1^{r_1}x_2^{r_2}\dots x_m^{r_m}$  and  $dx_1^{s_1}x_2^{s_2}\dots x_m^{s_m}$  run through all monomials appearing in  $f$  and  $g$ , respectively. This proves our contention, and also the lemma.  $\square$

By induction, we obtain

**38.6 Lemma:** Let  $D$  be an integral domain and  $f_1, f_2, \dots, f_t$  be nonzero polynomials in  $D[x_1, x_2, \dots, x_m]$ . Then the leading monomial of  $f_1 f_2 \dots f_t$  is the product of the leading monomials of  $f_1, f_2, \dots, f_t$ .  $\square$

**38.7 Lemma:** Let  $D$  be an integral domain,  $a \in D \setminus \{0\}$ , and let  $\sigma_1, \sigma_2, \dots, \sigma_m$  be the elementary symmetric polynomials in  $D[x_1, x_2, \dots, x_m]$ .

If  $k_1 \geq k_2 \geq k_3 \geq \dots \geq k_m \geq 0$  are integers, then the leading monomial of  $a \sigma_1^{k_1-k_2} \sigma_2^{k_2-k_3} \dots \sigma_{m-1}^{k_{m-1}-k_m} \sigma_m^{k_m}$  is  $a x_1^{k_1} x_2^{k_2} \dots x_{m-1}^{k_{m-1}} x_m^{k_m}$ .

**Proof:** The leading monomials of  $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \dots, \sigma_m$  are respectively  $x_1, x_1 x_2, x_1 x_2 x_3, x_1 x_2 x_3 x_4, \dots, x_1 x_2 \dots x_m$ , because  $\sigma_j$  is a sum of  $\binom{m}{j}$  monomials, each of which is a product of  $j$  indeterminates from  $x_1, x_2, \dots, x_m$ . In view of Lemma 38.6, the leading monomial of  $a \sigma_1^{k_1-k_2} \sigma_2^{k_2-k_3} \dots \sigma_{m-1}^{k_{m-1}-k_m} \sigma_m^{k_m}$  is

$$\begin{aligned} & a(x_1)^{k_1-k_2}(x_1 x_2)^{k_2-k_3}(x_1 x_2 x_3)^{k_3-k_4} \dots (x_1 x_2 x_3 \dots x_{m-1})^{k_{m-1}-k_m}(x_1 x_2 x_3 \dots x_{m-1} x_m)^{k_m} \\ &= a x_1^{k_1} x_2^{k_2} \dots x_{m-1}^{k_{m-1}} x_m^{k_m}. \end{aligned} \quad \square$$

We need one more lemma for the proof of the fundamental theorem.

**38.8 Lemma:** Let  $D$  be an integral domain and let  $f(x_1, x_2, \dots, x_m)$  be a nonzero symmetric polynomial in  $D[x_1, x_2, \dots, x_m]$ . Let  $a x_1^{k_1} x_2^{k_2} \dots x_m^{k_m}$  be the leading monomial of  $f$  (here  $a \in D, a \neq 0$  and  $k_1, k_2, \dots, k_m$  are nonnegative integers).

(1) We have  $k_1 \geq k_2 \geq \dots \geq k_{m-1} \geq k_m$ .

(2) If  $b x_1^{r_1} x_2^{r_2} \dots x_m^{r_m}$  is a monomial appearing in  $f$ , then

$$k_1 \geq r_1, k_2 \geq r_2, \dots, k_m \geq r_m.$$

**Proof:** Let  $\sigma$  be any permutation in  $S_m$  and let  $\tau$  be the inverse of  $\sigma$ .

(1) As  $a x_1^{k_1} x_2^{k_2} \dots x_m^{k_m}$  appears in  $f(x_1, x_2, \dots, x_m)$ ,

$$a x_{1\tau}^{k_1} x_{2\tau}^{k_2} \dots x_{m\tau}^{k_m} \text{ appears in } f(x_{1\tau}, x_{2\tau}, \dots, x_{m\tau}) = f^\tau = f = f(x_1, x_2, \dots, x_m),$$

$$a x_1^{k_1} x_2^{k_2} \dots x_m^{k_m} \text{ appears in } f(x_1, x_2, \dots, x_m),$$

and, since  $a x_1^{k_1} x_2^{k_2} \dots x_m^{k_m}$  is the leading monomial of  $f$ , we obtain:

$$\text{for all } \sigma \in S_m, (k_1, k_2, \dots, k_m) \text{ is higher than or equal to } (k_{1\sigma}, k_{2\sigma}, \dots, k_{m\sigma}).$$

Using this with  $\sigma = (12) \in S_m$ , we see  $(k_1, k_2, \dots, k_m)$  is higher than or equal to  $(k_2, k_1, \dots, k_m)$ , so  $k_1 \geq k_2$ . And  $\sigma = (23)$  yields that  $(k_1, k_2, k_3, \dots$

$(k_m)$  is higher than or equal to  $(k_1, k_3, k_2, \dots, k_m)$ , so  $k_2 \geq k_3$ . In like manner, when we choose  $\sigma = (34), \dots, (m-1, m) \in S_m$ , we get  $k_3 \geq k_4, \dots, k_{m-1} \geq k_m$ . This proves (1).

(2) As  $bx_1^{r_1}x_2^{r_2}\dots x_m^{r_m}$  appears in  $f(x_1, x_2, \dots, x_m)$ ,  
 $bx_{1\tau}^{r_1}x_{2\tau}^{r_2}\dots x_{m\tau}^{r_m}$  appears in  $f(x_{1\tau}, x_{2\tau}, \dots, x_{m\tau}) = f^\tau = f = f(x_1, x_2, \dots, x_m)$ ,  
 $bx_1^{r_{1\sigma}}x_2^{r_{2\sigma}}\dots x_m^{r_{m\sigma}}$  appears in  $f(x_1, x_2, \dots, x_m)$ ,

and so:

for all  $\sigma \in S_m$ ,  $(k_1, k_2, \dots, k_m)$  is higher than or equal to  $(r_{1\sigma}, r_{2\sigma}, \dots, r_{m\sigma})$ .

Thus  $k_1 \geq r_{1\sigma}$  for all  $\sigma \in S_m$ . Here  $1\sigma$  assumes all values  $1, 2, \dots, m$  as  $\sigma$  runs through  $S_m$ , and hence  $k_1 \geq r_1, k_1 \geq r_2, \dots, k_1 \geq r_m$ .  $\square$

**Proof of the fundamental theorem:** Throughout the proof, the number  $m$  of the indeterminates will be fixed. We make induction on the exponent system of the leading monomial of the symmetric polynomial. This will be explained shortly.

Let  $f$  be a nonzero symmetric polynomial in  $D[x_1, x_2, \dots, x_m]$  and let  $ax_1^{k_1}x_2^{k_2}\dots x_m^{k_m}$  be its leading monomial.

First we claim: if  $(k_1, k_2, \dots, k_m) = (0, 0, \dots, 0)$ , then there is a polynomial  $g$  in  $m$  indeterminates  $u_1, u_2, \dots, u_m$  over  $D$  such that  $f(x_1, x_2, \dots, x_m)$  is equal to  $g(\sigma_1, \sigma_2, \dots, \sigma_m)$ . This is very easy to prove. Indeed, if  $(k_1, k_2, \dots, k_m) = (0, 0, \dots, 0)$ , then, by Lemma 38.2(2), the exponent system of any monomial appearing in  $f$  is  $(0, 0, \dots, 0)$ , so  $f$  is the constant polynomial  $a$  in  $D[x_1, x_2, \dots, x_m]$ . Then of course  $f(x_1, x_2, \dots, x_m) = g(\sigma_1, \sigma_2, \dots, \sigma_m)$ , where  $g$  is the constant polynomial  $a$  in  $D[u_1, u_2, \dots, u_m]$ .

Now suppose that  $(k_1, k_2, \dots, k_m)$  is higher than  $(0, 0, \dots, 0)$  and that, for any nonzero symmetric polynomial  $f_1 \in D[x_1, x_2, \dots, x_m]$  whose leading monomial has a lower exponent system than  $(k_1, k_2, \dots, k_m)$ , there is a polynomial  $g_1$  in  $D[u_1, u_2, \dots, u_m]$  such that  $f_1(x_1, x_2, \dots, x_m) = g_1(\sigma_1, \sigma_2, \dots, \sigma_m)$ . Under this assumption, we will prove the existence of a polynomial  $g$  in  $D[u_1, u_2, \dots, u_m]$  with  $f(x_1, x_2, \dots, x_m) = g(\sigma_1, \sigma_2, \dots, \sigma_m)$ . This will establish the fundamental theorem because  $(0, 0, \dots, 0)$  is the lowest possible exponent system and the theorem has been proved in this case above. Moreover,

as there are only a finite number of  $m$ -tuples lower than  $(k_1, k_2, \dots, k_m)$ , the method of proof can be used effectively to find the polynomial  $g$  explicitly in concrete cases. [Basically, we write the  $m$ -tuples  $L_1, L_2, L_3, \dots$  in alphabetical order and prove that (1) the theorem is true for all nonzero symmetric polynomials whose leading monomials have the exponent system  $L_1 = (0, 0, \dots, 0)$  and that (2) for any  $s > 1$ , if the theorem is true for all nonzero symmetric polynomials whose leading monomials have exponent systems equal to one of  $L_1, L_2, \dots, L_{s-1}$ , then the theorem is also true for all nonzero symmetric polynomials whose leading monomials have the exponent system  $L_s$ . Once the leading monomial of a symmetric polynomial is given, there can be only a finite number of exponent systems of monomials appearing in that symmetric polynomial (Lemma 38.8(2).]

By Lemma 38.8(1), the integers  $k_1 - k_2, k_2 - k_3, \dots, k_m - k_{m-1}, k_m$  are non-negative and, by Lemma 38.7, the polynomial  $a\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\dots\sigma_{m-1}^{k_{m-1}-k_m}\sigma_m^{k_m}$  has the same leading monomial as  $f$ . Let  $f_1 = f - a\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\dots\sigma_{m-1}^{k_{m-1}-k_m}\sigma_m^{k_m}$ . Thus  $f_1$  is a symmetric polynomial in  $D[x_1, x_2, \dots, x_m]$ . If  $f_1 = 0$ , then  $f = a\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\dots\sigma_{m-1}^{k_{m-1}-k_m}\sigma_m^{k_m}$  and  $f = g(\sigma_1, \sigma_2, \dots, \sigma_m)$ , where  $g = au_1^{k_1-k_2}u_2^{k_2-k_3}\dots u_{m-1}^{k_{m-1}-k_m}u_m^{k_m} \in D[u_1, u_2, \dots, u_m]$ , and the proof is completed in this case. If  $f_1 \neq 0$ , then  $f_1$  has a leading monomial. The exponent system of this leading monomial of  $f_1$  is the exponent system of a monomial appearing in  $f$  or in  $a\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\dots\sigma_{m-1}^{k_{m-1}-k_m}\sigma_m^{k_m}$  (or in both). This exponent system is distinct from  $(k_1, k_2, \dots, k_m)$ . Since it arises from a monomial appearing in  $f$  or in  $a\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\dots\sigma_{m-1}^{k_{m-1}-k_m}\sigma_m^{k_m}$ , it is lower than the common exponent system  $(k_1, k_2, \dots, k_m)$  of the leading monomials of  $f$  and  $a\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\dots\sigma_{m-1}^{k_{m-1}-k_m}\sigma_m^{k_m}$ . By hypothesis, there is a polynomial  $g_1$  in  $D[u_1, u_2, \dots, u_m]$  such that  $f_1(x_1, x_2, \dots, x_m) = g_1(\sigma_1, \sigma_2, \dots, \sigma_m)$ . Hence

$$\begin{aligned} f &= f_1 + a\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\dots\sigma_{m-1}^{k_{m-1}-k_m}\sigma_m^{k_m} \\ &= g_1(\sigma_1, \sigma_2, \dots, \sigma_m) + a\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\dots\sigma_{m-1}^{k_{m-1}-k_m}\sigma_m^{k_m} \end{aligned}$$

and there is a polynomial  $g$  in  $D[u_1, u_2, \dots, u_m]$ , namely

$$g_1 + au_1^{k_1-k_2}u_2^{k_2-k_3}\dots u_{m-1}^{k_{m-1}-k_m}u_m^{k_m},$$

such that  $f(x_1, x_2, \dots, x_m) = g(\sigma_1, \sigma_2, \dots, \sigma_m)$ .

This completes the proof of the existence of  $g$ . It remains to show the uniqueness of  $g$ . Suppose now  $f$  is a nonzero symmetric polynomial in  $D[x_1, x_2, \dots, x_m]$  and assume that  $g, h \in D[u_1, u_2, \dots, u_m]$  with  $g(\sigma_1, \sigma_2, \dots, \sigma_m) = f(x_1, x_2, \dots, x_m) = h(\sigma_1, \sigma_2, \dots, \sigma_m)$ . If  $g$  were distinct from  $h$ , then  $g - h \neq 0$

would have a leading monomial which we may write in the form  $u_1^{s_1-s_2}u_2^{s_2-s_3}\dots u_{m-1}^{s_{m-1}-s_m}u_m^{s_m}$ , where  $s_1 \geq s_2 \geq \dots \geq s_{m-1} \geq s_m$ . Then  $0 = f - f = g(\sigma_1, \sigma_2, \dots, \sigma_m) - h(\sigma_1, \sigma_2, \dots, \sigma_m)$  in  $D[x_1, x_2, \dots, x_m]$  would have a leading monomial  $bx_1^{s_1}x_2^{s_2}\dots x_{m-1}^{s_{m-1}}x_m^{s_m}$ , a contradiction. Hence  $g = h$ , as was to be proved.  $\square$

The fundamental theorem can be summarized by saying that the substitution mapping

$$T: D[x_1, x_2, \dots, x_m] \longrightarrow S$$

$$g(u_1, u_2, \dots, u_m) \rightarrow g(\sigma_1, \sigma_2, \dots, \sigma_m)$$

is a ring isomorphism, where  $S$  is the subring of  $D[x_1, x_2, \dots, x_m]$  consisting of the symmetric polynomials in  $D[x_1, x_2, \dots, x_m]$ .

### 38.9 Examples: (a) We express the polynomial

$$f(x, y, z) = 5xyz + x^2y + xy^2 + xz^2 + yz^2 + y^2z + x^2z \in \mathbb{Z}[x, y, z]$$

in terms of  $\sigma_1, \sigma_2, \sigma_3$ .

We first arrange the monomials appearing in  $f$  in the alphabetical order of their exponent systems:

$$f(x, y, z) = x^2y + x^2z + xy^2 + 5xyz + xz^2 + y^2z + yz^2.$$

The leading monomial of  $f$  is  $1x^2y^1z^0$ . We therefore subtract  $1\sigma_1^{2-1}\sigma_2^{1-0}\sigma_3^0$  from  $f$  and get

$$f - \sigma_1\sigma_2 = (x^2y + x^2z + xy^2 + 5xyz + xz^2 + y^2z + yz^2) - (x + y + z)(xy + yz + zx)$$

$$= 2xyz.$$

The leading monomial of  $f - \sigma_1\sigma_2$  is  $2x^1y^1z^1$ . So we subtract  $2\sigma_1^{1-1}\sigma_2^{1-1}\sigma_3^1$  from  $f - \sigma_1\sigma_2$  and get

$$(f - \sigma_1\sigma_2) - 2\sigma_3 = 2xyz - 2xyz = 0.$$

Hence  $f(x, y, z) = \sigma_1\sigma_2 + 2\sigma_3$ .

### (b) We express

$$f(x, y, z, w) = x^3 + y^3 + z^3 + w^3 \in \mathbb{Z}[x, y, z]$$

in terms of  $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ . The monomials are in alphabetical order, and the leading monomial of  $f$  is  $1x^3y^0z^0w^0$ . So we subtract  $1\sigma_1^{3-0}\sigma_2^{0-0}\sigma_3^{0-0}\sigma_4^0$  from  $f$  and get

$$f - \sigma_1^3 = (x^3 + y^3 + z^3 + w^3) - (x + y + z + w)^3$$

$$= \dots$$

$$= -3x^2y - 3xy^2 - 3x^2z - 3xz^2 - 3x^2w - 3xw^2 - 3y^2z - 3yz^2$$

$$- 3y^2w - 3yw^2 - 3z^2w - 3zw^2 - 6xyz - 6xyw - 6xzw - 6yzw.$$

The leading monomial of  $f - \sigma_1^3$  is  $-3x^2y = -3x^2y^1z^0w^0$ . We therefore subtract  $-3\sigma_1^{2-1}\sigma_2^{1-0}\sigma_3^{0-0}\sigma_4^{0-0}$  from  $f - \sigma_1^3$  and get

$$\begin{aligned} (f - \sigma_1^3) - (-3\sigma_1\sigma_2) &= (f - \sigma_1^3) + 3(x + y + z + w)(xy + xz + xw + yz + yw + zw) \\ &= \dots \\ &= 3xyz + 3xyw + 3xzw + 3yzw \\ &= 3\sigma_3. \end{aligned}$$

Hence  $f(x,y,z,w) = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$ .

We now derive formulas connecting the sum of the  $k$ -th powers of  $x_1, x_2, \dots, x_m$  with the elementary symmetric polynomials. These formulas are due to I. Newton (1642-1727).

**38.10 Theorem (Newton):** *Let  $D$  be an integral domain and  $x_1, x_2, \dots, x_m$  indeterminates over  $D$ . For  $k = 1, 2, 3, \dots$ , we put  $s_k = x_1^k + x_2^k + \dots + x_m^k$ , so that  $s_k \in D[x_1, x_2, \dots, x_m]$ . Then*

$$\begin{aligned} 0 &= s_1 - \sigma_1 \\ 0 &= s_2 - \sigma_1 s_1 + 2\sigma_2 \\ 0 &= s_3 - \sigma_1 s_2 + \sigma_2 s_1 - 3\sigma_3 \\ &\dots \\ 0 &= s_{m-1} - \sigma_1 s_{m-2} + \sigma_2 s_{m-3} + \dots + (-1)^{m-2} \sigma_{m-2} s_1 + (-1)^{m-1} (m-1) \sigma_{m-1} \end{aligned}$$

and

$$\begin{aligned} 0 &= s_m - \sigma_1 s_{m-1} + \sigma_2 s_{m-2} + \dots + (-1)^{m-2} \sigma_{m-2} s_2 + (-1)^{m-1} \sigma_{m-1} s_1 + (-1)^m m \sigma_m \\ 0 &= s_{m+1} - \sigma_1 s_m + \sigma_2 s_{m-1} + \dots + (-1)^{m-2} \sigma_{m-2} s_3 + (-1)^{m-1} \sigma_{m-1} s_2 + (-1)^m \sigma_m s_1 \\ 0 &= s_{m+2} - \sigma_1 s_{m+1} + \sigma_2 s_m + \dots + (-1)^{m-2} \sigma_{m-2} s_4 + (-1)^{m-1} \sigma_{m-1} s_3 + (-1)^m \sigma_m s_2 \\ 0 &= s_{m+3} - \sigma_1 s_{m+2} + \sigma_2 s_{m+1} + \dots + (-1)^{m-2} \sigma_{m-2} s_5 + (-1)^{m-1} \sigma_{m-1} s_4 + (-1)^m \sigma_m s_3 \\ &\dots \end{aligned}$$

**Proof:** We make use of the polynomial  $f(t) = (t - x_1)(t - x_2)\dots(t - x_m)$ . We know that

$$f(t) = t^m - \sigma_1 t^{m-1} + \sigma_2 t^{m-2} - \dots + (-1)^{m-1} \sigma_{m-1} t + (-1)^m \sigma_m$$

and that  $x_1, x_2, \dots, x_m$  are the roots of  $f(t) \in D[x_1, x_2, \dots, x_m]$ . Hence

$$0 = x_i^m - \sigma_1 x_i^{m-1} + \sigma_2 x_i^{m-2} - \dots + (-1)^{m-1} \sigma_{m-1} x_i + (-1)^m \sigma_m$$

for all  $i = 1, 2, \dots, m$ . Multiplying both sides of this equation by  $x_i^j$ , where  $j = 0, 1, 2, 3, \dots$ , we get

$$0 = x_i^{m+j} - \sigma_1 x_i^{m+j-1} + \sigma_2 x_i^{m+j-2} - \dots + (-1)^{m-1} \sigma_{m-1} x_i^{j+1} + (-1)^m \sigma_m x_i^j$$

for all  $i = 1, 2, \dots, m$ . Adding these  $m$  equations side by side, we obtain 0 =

$$\sum_{i=1}^m x_i^{m+j} - \sigma_1 \sum_{i=1}^m x_i^{m+j-1} + \sigma_2 \sum_{i=1}^m x_i^{m+j-2} - \dots + (-1)^{m-1} \sigma_{m-1} \sum_{i=1}^m x_i^{j+1} + (-1)^m \sigma_m \sum_{i=1}^m x_i^j$$

i.e.,

$$0 = s_{m+j} - \sigma_1 s_{m+j-1} + \sigma_2 s_{m+j-2} + \dots + (-1)^{m-2} \sigma_{m-2} s_{j+2} + (-1)^{m-1} \sigma_{m-1} s_{j+1} + (-1)^m \sigma_m s_j.$$

This establishes all the equations except the first  $m - 1$  of them. The first  $m - 1$  equations will be established by a similar reasoning. This time we make use of the derivative of  $f(t)$ . By Lemma 35.16(2), we have

$$\begin{aligned} f'(t) &= (t-x_2)(t-x_3)\dots(t-x_m) + (t-x_1)(t-x_3)\dots(t-x_m) + \dots + (t-x_1)(t-x_2)\dots(t-x_{m-1}) \\ &= \frac{f(t)}{t-x_1} + \frac{f(t)}{t-x_2} + \dots + \frac{f(t)}{t-x_m}. \end{aligned}$$

For  $i = 1, 2, \dots, m$ , we put

$$\frac{f(t)}{t-x_i} = q_{m-1}^{(i)} t^{m-1} + q_{m-2}^{(i)} t^{m-2} + \dots + q_1^{(i)} t + q_0^{(i)}.$$

$$\begin{aligned} \text{Hence } m t^{m-1} - (m-1) \sigma_1 t^{m-2} + (m-2) \sigma_2 t^{m-3} - \dots + (-1)^{m-1} \sigma_{m-1} &= f'(t) \\ = \sum_{i=1}^m \frac{f(t)}{t-x_i} &= \sum_{i=1}^m (q_{m-1}^{(i)} t^{m-1} + q_{m-2}^{(i)} t^{m-2} + \dots + q_1^{(i)} t + q_0^{(i)}) \\ &= \left( \sum_{i=1}^m q_{m-1}^{(i)} \right) t^{m-1} + \left( \sum_{i=1}^m q_{m-2}^{(i)} \right) t^{m-2} + \dots + \left( \sum_{i=1}^m q_1^{(i)} \right) t + \left( \sum_{i=1}^m q_0^{(i)} \right), \end{aligned}$$

so that

$$m = \sum_{i=1}^m q_{m-1}^{(i)}, \quad -(m-1)\sigma_1 = \sum_{i=1}^m q_{m-2}^{(i)}, \quad \dots, \quad (-1)^{m-2}2\sigma_{m-2} = \sum_{i=1}^m q_1^{(i)},$$

$$(-1)^{m-1}\sigma_{m-1} = \sum_{i=1}^m q_0^{(i)}. \quad (*)$$

On the other hand,  $t^m - \sigma_1 t^{m-1} + \sigma_2 t^{m-2} - \dots + (-1)^{m-1}\sigma_{m-1}t + (-1)^m\sigma_m$

$$= f(t) = (t - x_i)(q_{m-1}^{(i)}t^{m-1} + q_{m-2}^{(i)}t^{m-2} + \dots + q_1^{(i)}t + q_0^{(i)})$$

$$= q_{m-1}^{(i)}t^m + q_{m-2}^{(i)}t^{m-1} + q_{m-3}^{(i)}t^{m-2} + \dots + q_1^{(i)}t^2 + q_0^{(i)}t$$

$$- q_{m-1}^{(i)}x_i t^{m-1} - q_{m-2}^{(i)}x_i t^{m-2} - \dots - q_2^{(i)}x_i t^2 - q_1^{(i)}x_i t - q_0^{(i)}x_i.$$

Comparing the coefficients of powers of  $t$  on both sides, we get

$$1 = q_{m-1}^{(i)}$$

$$-\sigma_1 = q_{m-2}^{(i)} - q_{m-1}^{(i)}x_i$$

$$+\sigma_2 = q_{m-3}^{(i)} - q_{m-2}^{(i)}x_i$$

$$-\sigma_3 = q_{m-4}^{(i)} - q_{m-3}^{(i)}x_i$$

.....

$$(-1)^{m-1}\sigma_{m-1} = q_0^{(i)} - q_1^{(i)}x_i$$

$$(-1)^m\sigma_m = -q_0^{(i)}x_i,$$

which may be written

$$q_{m-1}^{(i)} = 1$$

$$q_{m-2}^{(i)} = -\sigma_1 + q_{m-1}^{(i)}x_i$$

$$q_{m-3}^{(i)} = +\sigma_2 + q_{m-2}^{(i)}x_i$$

$$q_{m-4}^{(i)} = -\sigma_3 + q_{m-3}^{(i)}x_i$$

.....

$$q_0^{(i)} = (-1)^{m-1}\sigma_{m-1} + q_1^{(i)}x_i$$

$$0 = (-1)^m\sigma_m + q_0^{(i)}x_i.$$

So, for each  $i = 1, 2, \dots, m,$

$$q_{m-2}^{(i)} = -\sigma_1 + x_i$$

$$(1)$$

$$q_{m-3}^{(i)} = +\sigma_2 + (-\sigma_1 + x_i)x_i = \sigma_2 - \sigma_1 x_i + x_i^2 \quad (2)$$

$$q_{m-4}^{(i)} = -\sigma_3 + (\sigma_2 - \sigma_1 x_i + x_i^2)x_i = -\sigma_3 + \sigma_2 x_i - \sigma_1 x_i^2 + x_i^3 \quad (3)$$

.....

$$q_0^{(i)} = (-1)^{m-1}\sigma_{m-1} + ((-1)^{m-2}\sigma_{m-2} + (-1)^{m-3}\sigma_{m-3}x_i + \dots + (-1)\sigma_1 x_i^{m-3} + x_i^{m-2})x_i$$

$$= (-1)^{m-1}\sigma_{m-1} + (-1)^{m-2}\sigma_{m-2}x_i + (-1)^{m-3}\sigma_{m-3}x_i^2 + \dots + (-1)\sigma_1x_i^{m-2} + x_i^m. \quad (m - 1)$$

We now the  $m$  equations (1), the  $m$  equations (2), the  $m$  equations (3),... , the  $m$  equations ( $m - 1$ ) (for  $i = 1, 2, \dots, m$ ). Using (\*), we get

$$\begin{aligned} -(m - 1)\sigma_1 &= -m\sigma_1 + s_1 \\ +(m - 2)\sigma_2 &= +m\sigma_2 - \sigma_1s_1 + s_2 \\ -(m - 3)\sigma_3 &= -m\sigma_3 + \sigma_2s_1 - \sigma_1s_2 + s_3 \\ &\dots \end{aligned}$$

$$(-1)^{m-1}\sigma_{m-1} = (-1)^{m-1}m\sigma_{m-1} + (-1)^{m-2}\sigma_{m-2}s_1 + (-1)^{m-3}\sigma_{m-3}s_2 + \dots + (-1)\sigma_1s_{m-2} + s_{m-1},$$

which are equivalent to

$$\begin{aligned} s_1 - \sigma_1 &= 0 \\ s_2 - \sigma_1s_1 + 2\sigma_2 &= 0 \\ s_3 - \sigma_1s_2 + \sigma_2s_1 - 3\sigma_3 &= 0 \\ &\dots \\ s_{m-1} - \sigma_1s_{m-2} + \sigma_2s_{m-3} + \dots + (-1)^{m-2}\sigma_{m-2}s_1 + (-1)^{m-1}(m - 1)\sigma_{m-1} &= 0. \end{aligned}$$

This completes the proof. □

Newton's formulas express  $s_k$  recursively in terms of  $s_1, s_2, \dots, s_{k-1}$  and of  $\sigma_1, \sigma_2, \dots, \sigma_m$ . We can eliminate  $s_1, s_2, \dots, s_{k-1}$  and write  $s_k$  solely in terms of  $\sigma_1, \sigma_2, \dots, \sigma_m$ . For instance:

$$\begin{aligned} s_1 &= \sigma_1 \\ s_2 &= \sigma_1s_1 - 2\sigma_2 = \sigma_1\sigma_1 - 2\sigma_2 = \sigma_1^2 - 2\sigma_2 \\ s_3 &= \sigma_1s_2 - \sigma_2s_1 + 3\sigma_3 = \sigma_1(\sigma_1^2 - 2\sigma_2) - \sigma_2\sigma_1 + 3\sigma_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3 \\ s_4 &= \sigma_1s_3 - \sigma_2s_2 + \sigma_3s_1 - 4\sigma_4 = \sigma_1(\sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3) - \sigma_2(\sigma_1^2 - 2\sigma_2) + \sigma_3\sigma_1 - 4\sigma_4 \\ &= \sigma_1^4 - 4\sigma_1^2\sigma_2 + 4\sigma_1\sigma_3 + 2\sigma_2^2 - 4\sigma_4 \end{aligned}$$

(here  $\sigma_j$  should be replaced by 0 when  $j > m$ ).

Now let  $D, E$  be integral domains and  $D \subseteq E$ . Let

$$p(t) = c_0t^m + c_1t^{m-1} + \dots + c_{m-1}t + c_m$$

be a nonzero polynomial of degree  $m$  in  $D[t]$ , and assume that there are exactly  $m$  roots  $a_1, a_2, \dots, a_m$  of  $p$  in  $E$  (counted with multiplicities). Then

$$p(t) = c_0(t - a_1)(t - a_2)\dots(t - a_m) \quad \text{in } E[t].$$

Hence  $p(t) \in E[t]$  is obtained from

$$c_0 f(t) = c_0(t - x_1)(t - x_2)\dots(t - x_m) \in D[x_1, x_2, \dots, x_m][t]$$

by substituting  $a_i$  for  $x_i$  ( $i = 1, 2, \dots, m$ ). Now  $c_0 f(t) =$

$$c_0(t^m - \sigma_1(x_1, x_2, \dots, x_m)t^{m-1} + \sigma_2(x_1, x_2, \dots, x_m)t^{m-2} - \dots + (-1)^m \sigma_m(x_1, x_2, \dots, x_m))$$

and, since substitution is a homomorphism (Lemma 35.20), we have

$$p(t) = c_0(t^m - \sigma_1(a_1, a_2, \dots, a_m)t^{m-1} + \sigma_2(a_1, a_2, \dots, a_m)t^{m-2} - \dots + (-1)^m \sigma_m(a_1, a_2, \dots, a_m)).$$

Therefore

$$\begin{aligned} c_1 &= -c_0 \sigma_1(a_1, a_2, \dots, a_m) \\ c_2 &= +c_0 \sigma_2(a_1, a_2, \dots, a_m) \\ c_3 &= -c_0 \sigma_3(a_1, a_2, \dots, a_m) \\ &\dots\dots\dots \\ c_{m-1} &= (-1)^{m-1} \sigma_{m-1}(a_1, a_2, \dots, a_m) \\ c_m &= (-1)^m \sigma_m(a_1, a_2, \dots, a_m); \end{aligned}$$

in words: the values of the elementary symmetric polynomials at the roots of a polynomial are equal to the coefficients of that polynomial, except for a factor  $\mp c_0$ , where  $c_0$  is the leading coefficient of the polynomial. The equations above tell us that (i)  $\sigma_i(a_1, a_2, \dots, a_m)$  belong to  $D$  if  $c_0$  is a unit in  $D$ ; (ii)  $\sigma_i(a_1, a_2, \dots, a_m)$  belong to the field of fractions of  $D$  in any case; (iii) in particular,  $\sigma_i(a_1, a_2, \dots, a_m)$  belong to  $D$  if  $D$  is a field. Moreover, if  $h(x_1, x_2, \dots, x_m) \in D[x_1, x_2, \dots, x_m]$  is a symmetric polynomial, then  $h(x_1, x_2, \dots, x_m) = g(\sigma_1, \sigma_2, \dots, \sigma_m)$  for some polynomial in  $m$  indeterminates over  $D$ , and substitution yields

$$h(a_1, a_2, \dots, a_m) = g(\sigma_1(a_1, a_2, \dots, a_m), \sigma_2(a_1, a_2, \dots, a_m), \dots, \sigma_m(a_1, a_2, \dots, a_m))$$

so that (i)  $h(a_1, a_2, \dots, a_m)$  belongs to  $D$  if  $c_0$  is a unit in  $D$ ; (ii)  $h(a_1, a_2, \dots, a_m)$  belongs to the field of fractions of  $D$  in any case; (iii)  $h(a_1, a_2, \dots, a_m)$  belongs to  $D$  if  $D$  is a field. We summarize this discussion in the next theorem.

**38.11 Theorem:** *Let  $D$  be an integral domain and let*

*$p(t) = c_0 t^m + c_1 t^{m-1} + \dots + c_{m-1} t + c_m$  a polynomial over  $D$ . Assume that  $p(t)$  has exactly  $m$  roots  $a_1, a_2, \dots, a_m$  in an integral domain containing  $D$ .*

(1)  $c_i = (-1)^i \sigma_m(a_1, a_2, \dots, a_m)$  for  $i = 1, 2, \dots, m$ .

- (2) If  $h$  is any symmetric polynomial in  $m$  indeterminates over  $D$ , then  $h(a_1, a_2, \dots, a_m)$ , which is an element of the integral domain containing the roots of  $p(t)$ , is in fact an element of the field of fractions of  $D$ .
- (3) If, in addition, the leading coefficient of  $p(t)$  is a unit in  $D$ , then  $h(a_1, a_2, \dots, a_m)$  belongs to  $D$ .
- (4) If, in particular,  $D$  is a field, then  $h(a_1, a_2, \dots, a_m)$  belongs to  $D$ .  $\square$

It is true that any nonzero polynomial of degree  $m$  over an integral domain  $D$  has exactly  $m$  roots in some integral domain containing  $D$ . This will be proved later (Theorem 53.6). In the following examples, we will assume that the polynomials have as many roots as their degrees in some integral domain.

**Examples: (a)** Let us evaluate  $a^2b^2 + a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 + c^2d^2$ , where  $a, b, c, d$  are the roots of  $t^4 - t^2 + 1 \in \mathbb{Z}[t]$ . To this end, we express the symmetric polynomial  $x^2y^2 + x^2z^2 + x^2u^2 + y^2z^2 + y^2u^2 + z^2u^2$  in terms of  $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ . Subtracting  $1\sigma_1^{2-2}\sigma_2^{2-0}\sigma_3^{0-0}\sigma_4^{0-0}$  from this polynomial, we get

$$(x^2y^2 + \dots + z^2u^2) - \sigma_2^2 = -2x^2yz - \dots - 6xyzu,$$

$$\text{and } (x^2y^2 + \dots + z^2u^2) - \sigma_2^2 - (-2\sigma_1^{2-1}\sigma_2^{1-1}\sigma_3^{1-0}\sigma_4^0) = \dots = 0,$$

$$\text{so } x^2y^2 + x^2z^2 + x^2u^2 + y^2z^2 + y^2u^2 + z^2u^2 = \sigma_2^2 - \sigma_1\sigma_3.$$

$$\text{Then } a^2b^2 + a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 + c^2d^2$$

$$= (ab + ac + ad + bc + bd + cd)^2 - 2(a + b + c + d)(abc + abd + acd + bcd).$$

Here  $a, b, c, d$  are the roots of  $t^4 - t^2 + 1$ , so

$$a + b + c + d = -0, \quad ab + ac + ad + bc + bd + cd = +(-1),$$

$$abc + abd + acd + bcd = -0, \quad abcd = +1$$

$$\text{and therefore } a^2b^2 + a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 + c^2d^2 = (-1)^2 - 2(0)(0) = 1.$$

**(b)** We find a polynomial of degree three in  $\mathbb{Z}[t]$  whose roots are the cubes of the roots of  $t^3 + 2t^2 + 3t + 4 \in \mathbb{Z}[t]$ . Let us denote the roots of this polynomial by  $a, b, c$ , so that  $a + b + c = -2$ ,  $ab + ac + bc = 3$ ,  $abc = 4$ .

$$\text{We put } t^3 + q_1t^2 + q_2t + q_3 = (t - a^3)(t - b^3)(t - c^3).$$

From Theorem 38.11, we know that

$$q_1 = a^3 + b^3 + c^3, \quad q_2 = a^3b^3 + a^3c^3 + b^3c^3, \quad q_3 = a^3b^3c^3.$$

Since  $s_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$ , we conclude

$$q_1 = a^3 + b^3 + c^3$$

$$\begin{aligned}
&= (a + b + c)^3 - 3(a + b + c)(ab + ac + bc) + 3(abc) \\
&= (-2)^3 - 3(-2)(3) + 3(-4) \\
&= -2.
\end{aligned}$$

We find easily that  $x^3y^3 + x^3z^3 + y^3z^3 = \sigma_2^3 - 3\sigma_1\sigma_2\sigma_3 + 3\sigma_3^2$ ; hence

$$\begin{aligned}
q_2 &= a^3b^3 + a^3c^3 + b^3c^3 \\
&= (ab + ac + bc)^3 - 3(a + b + c)(ab + ac + bc)(abc) + 3(abc)^2 \\
&= (3)^3 - 3(-2)(3)(-4) + 3(-4)^2 \\
&= 3.
\end{aligned}$$

Finally,  $q_3 = a^3b^3c^3$

$$\begin{aligned}
&= (abc)^3 \\
&= (-4)^3 \\
&= -64.
\end{aligned}$$

Thus

$$t^3 - (-2)t^2 + (3)t - (-64) = t^3 + 2t^2 + 3t + 64 \in \mathbb{Z}[x]$$

is a polynomial whose roots are the cubes of the roots of  $t^3 + 2t^2 + 3t + 4$ .

## Exercises

1. Express the following symmetric polynomials over  $\mathbb{Z}$  in terms of the elementary symmetric polynomials:

- (a)  $x^3y^2 + x^2y^3 + x^3z^2 + x^2z^3 + y^3z^2 + y^2z^3$ ;
- (b)  $x^2y^2 + x^2z^2 + x^2u^2 + y^2z^2 + y^2u^2 + z^2u^2$ ;
- (c)  $x^5 + y^5 + x^5 + x^4y + y^4x + x^4z + z^4x + y^4z + z^4y$ .

2. Find a polynomial over  $\mathbb{Z}$  whose roots are the

- (a) squares of the roots of  $t^3 + 5t^2 + 7t + 1 \in \mathbb{Z}[t]$ ;
- (b) squares of the roots of  $t^5 + 5t^4 - 6t^3 + t^2 - 7t - 4 \in \mathbb{Z}[t]$ ;
- (c) cubes of the roots of  $t^4 - 3t^3 + 2t^2 + 2 \in \mathbb{Z}[t]$ .

3. Let  $K$  be a field. A rational function  $\frac{f(x_1, x_2, \dots, x_m)}{g(x_1, x_2, \dots, x_m)}$  in  $K[x_1, x_2, \dots, x_m]$

is said to be a *symmetric rational function over  $K$*  if

$$\frac{f(x_{1\sigma}, x_{2\sigma}, \dots, x_{m\sigma})}{g(x_{1\sigma}, x_{2\sigma}, \dots, x_{m\sigma})} = \frac{f(x_1, x_2, \dots, x_m)}{g(x_1, x_2, \dots, x_m)}$$

for all  $\sigma \in S_n$ . Prove that a symmetric rational function over  $K$  can be expressed as a fraction of two symmetric polynomials over  $K$ . Conclude that any symmetric rational function over  $K$  can be written as

$$\frac{p(\sigma_1, \sigma_2, \dots, \sigma_m)}{q(\sigma_1, \sigma_2, \dots, \sigma_m)}$$

with suitable polynomials  $p, q$  in  $K[u_1, u_2, \dots, u_m]$ . (Loosely speaking, any symmetric rational function is a rational function of the elementary symmetric polynomials.)

4. Express the following rational functions over  $\mathbb{Z}$  in terms of the elementary symmetric polynomials:

(a)  $\frac{x}{y} + \frac{y}{x} + \frac{x}{z} + \frac{z}{x} + \frac{y}{z} + \frac{z}{y};$

(b)  $\frac{x^2}{yz} + \frac{y^2}{xz} + \frac{z^2}{xy};$

(c)  $\frac{1}{1-x} + \frac{1}{1-y} + \frac{1}{1-z}.$

5. Prove: for any symmetric polynomial  $f(x_1, x_2, \dots, x_m)$  over  $\mathbb{Z}$ , there is a polynomial  $h(u_1, u_2, \dots, u_m)$  in  $\mathbb{Q}[u_1, u_2, \dots, u_m]$  such that  $f(x_1, x_2, \dots, x_m) = h(s_1, s_2, \dots, s_m)$ , where  $s_j$  are the power sums of  $x_i$ .

6. Write the symmetric polynomials in Ex. 1 as polynomials in  $s_j$  over  $\mathbb{Q}$ .