# 21MAB143 Rings and Polynomials: Week 7

## 1 Symmetric polynomials

**Definition 1.1.** *Let $K$ be a field. A polynomial $f \in K[x_1, \ldots, x_n]$ is* **symmetric** *if it is unchanged by every permutation of the variables $x_1, \ldots, x_n$.*

**Examples:** Let's look at a couple of examples in $\mathbf{R}[x, y, z]$.

The polynomials

$$\begin{aligned} f &= x^3 + y^3 + z^3 \\ g &= x^6yz + xy^6z + xyz^6 - 17x^7y^7z^7 \end{aligned}$$

are both symmetric. The polynomial

$$h = xy^2 + yz^2 + zx^2$$

is not symmetric: if we swap $x$ and $y$ it is transformed into the polynomial

$$h' = x^2y + xz^2 + y^2z$$

which is not equal to $h$.

**Definition 1.2.** *Fix a positive integer $n$. For a positive integer $k$, the* **$k$-th elementary symmetric polynomial in** *$x_1, \ldots, x_n$, denoted $\sigma_k(x_1, \ldots, x_n)$, is the element of $K[x_1, \ldots, x_n]$ defined as*

$$\begin{aligned} \sigma_k(x_1, \ldots, x_n) &= \sum_{1 \le j_1 < j_2 < \cdots < j_k \le n} x_{j_1} \cdots x_{j_k} && \text{for } 1 \le k \le n \\ &= 0 && \text{for } k > n. \end{aligned}$$

*In the first formula we are summing over all monomials obtained by taking the product of any chosen set of $k$ out of the $n$ variables. We may sometimes write just $\sigma_k$ instead of $\sigma_k(x_1, \ldots, x_n)$ if the value of $n$ is clear from context.*

Writing these out more explicitly, for any $n$ we have

$$\begin{aligned}
\sigma_1 &= x_1 + \cdots + x_n \\
\sigma_2 &= x_1x_2 + \cdots + x_1x_n + x_2x_3 + \cdots + x_2x_n + \cdots + x_{n-1}x_n \\
&\vdots \\
\sigma_n &= x_1x_2\cdots x_n
\end{aligned}$$

For small values of $n$ we can write out all $n$ of the nonzero elementary symmetric polynomials in full:

- **$n = 1$:** Here there is only one nonzero elementary symmetric polynomial

$$\sigma_1(x) = x$$

- **$n = 2$:** Here we have two nonzero elementary symmetric polynomials:

$$\begin{aligned}
\sigma_1(x, y) &= x + y \\
\sigma_2(x, y) &= xy
\end{aligned}$$

- **$n = 3$:** Here the three nonzero elementary symmetric polynomials are:

$$\begin{aligned}
\sigma_1(x, y, z) &= x + y + z \\
\sigma_2(x, y, z) &= xy + xz + yz \\
\sigma_3(x, y, z) &= xyz
\end{aligned}$$

The importance of the elementary symmetric polynomials is that **every** symmetric polynomial in $K[x_1, \ldots, x_n]$ can be written in terms of them. Here is the formal statement:

**Theorem 1.3** (Fundamental Theorem of Symmetric Polynomials)**.** *Let $f \in K[x_1, \ldots, x_n]$ be a symmetric polynomial. Then there is a polynomial $F \in K[x_1, \ldots, x_n]$ such that*

$$f = F(\sigma_1, \ldots, \sigma_n).$$

The proof requires a way to "order" polynomials in the polynomial ring $K[x_1, \ldots, x_n]$, which we now introduce.

**Definition 1.4.** *In the ring $K[x_1, \ldots, x_n]$, the* **multidegree** *of a monomial*

$$m = ax_1^{d_1} \cdots x_n^{d_n}$$

*(where $a \in K$ and the $d_i$ are nonnegative integers) is defined to be*

$$\operatorname{mdeg}(m) = (d_1, \ldots, d_n)\,.$$

*For two multidegrees* $(d_1, \dots, d_n)$ *and* $(e_1, \dots, e_n)$ *we write*

$$(d_1, \dots, d_n) > (e_1, \dots, e_n)$$

*if, for the first* $i$ *such that* $d_i \neq e_i$*, we have* $d_i > e_i$*.*

*For two monomials* $m_1$ *and* $m_2$ *we write* $m_1 > m_2$ *if* $\mathrm{mdeg}(m_1) > \mathrm{mdeg}(m_2)$*.*

*Finally, for a polynomial* $f \in K[x_1, \dots, x_n]$*, we define its* **leading term** $\mathrm{lt}(f)$ *to be the unique monomial* $m$ *in* $f$ *such that* $m > m'$ *for all other monomials* $m'$ *in* $f$*, and we define the multidegree of* $f$ *to be*

$$\mathrm{mdeg}(f) = \mathrm{mdeg}(\mathrm{lt}(f))$$

*where the right-hand side is defined above since* $\mathrm{lt}(f)$ *is a monomial.*

**Example:** The symmetric polynomial

$$f = x^3yz + xy^3z + xyz^3 - 2x^2 - 2y^2 - 2z^2 + 4$$

has 7 monomials with the following multidegrees:

$$\begin{aligned} \mathrm{mdeg}(x^3yz) &= (3, 1, 1) \\ \mathrm{mdeg}(xy^3z) &= (1, 3, 1) \\ \mathrm{mdeg}(xyz^3) &= (1, 1, 3) \\ \mathrm{mdeg}(-2x^2) &= (2, 0, 0) \\ \mathrm{mdeg}(-2y^2) &= (0, 2, 0) \\ \mathrm{mdeg}(-2z^2) &= (0, 0, 2) \\ \mathrm{mdeg}(4) &= (0, 0, 0) \end{aligned}$$

We can order these monomials using the order relation above:

$$4 < -2z^2 < -2y^2 < xyz^3 < xy^3z < -2x^2 < x^3yz.$$

In particular the leading term of f is

$$\mathrm{lt}(f) = x^3yz.$$

and hence

$$\begin{aligned} \mathrm{mdeg}(f) &= \mathrm{mdeg}(x^3yz) \\ &= (3, 1, 1). \end{aligned}$$

Before we start the proof, we need to prove a lemma about the multidegrees of symmetric polynomials.

**Lemma 1.5.** *Let* $f \in K[x_1, \dots, x_n]$ *be a symmetric polynomial. If*

$$\mathrm{mdeg}(f) = (d_1, \dots, d_n)$$

*then* $d_1 \geq d_2 \geq \cdots \geq d_n$.

*Proof of Lemma.* Suppose that the multidegree of f is $(d_1, \dots, d_n)$. So the leading term of f is

$$\mathrm{lt}(f) = a x_1^{d_1} \cdots x_n^{d_n}$$

for some constant $a \in K$. Since f is symmetric, it must contain all the monomials obtained by permuting the exponents in $\mathrm{lt}(f)$: that is, all the monomials of the form

$$a x_1^{e_1} \cdots x_n^{e_n}$$

where $(e_1, \dots, e_n)$ is any permutation of $(d_1, \dots, d_n)$. But by definition $\mathrm{lt}(f)$ has the biggest multidegree among all the monomials in f, so this means that $(d_1, \dots, d_n) > (e_1, \dots, e_n)$ for any nontrivial permutation of the $d_i$. This implies that

$$d_1 \geq d_2 \geq \cdots \geq d_n.$$

$\square$

*Proof of Theorem* **??**. If f is the zero polynomial there is nothing to prove. For nonzero symmetric polynomials f, we prove the theorem by induction on the multidegree. This makes sense because for a given multidegree $(d_1, \dots, d_n)$, there are only finitely many possible multidegrees $(e_1, \dots, e_n) < (d_1, \dots, d_n)$ such that $e_1 \geq e_2 \cdots \geq e_n$, and Lemma **??** the degree of any symmetric polynomial must satisfy these conditions.

The base case of multidegree $(0, \dots, 0)$ is trivial: any such polynomial is a constant $a \in K$, and so the theorem holds with $F = a$, a constant polynomial.

For the inductive step, given our symmetric polynomial f we claim that we can find a polynomial $\Phi$ such that $\Phi(\sigma_1, \dots, \sigma_n)$ has the same leading term as f. Then $f - \Phi(\sigma_1, \dots, \sigma_n)$ will be a symmetric polynomial with smaller multidegree than f, so by induction there exists a polynomial G such that

$$\begin{aligned} f - \Phi(\sigma_1, \dots, \sigma_n) &= G(\sigma_1, \dots, \sigma_n) \quad \text{hence} \\ f &= (\Phi + G)(\sigma_1, \dots, \sigma_n) \end{aligned}$$

and so the theorem is proved with $F = \Phi + G$.

It remains to prove the claim. Let

$$\mathrm{lt}(f) = a x_1^{d_1} \cdots x_n^{d_n}$$

where $a \in K$. Note that $\mathrm{mdeg}(f) = (d_1, \dots, d_n)$. Consider the polynomial

$$a \sigma_1^{\delta_1} \cdots \sigma_n^{\delta_n}$$

where

$$\delta_1 = d_1 - d_2,\ \delta_2 = d_2 - d_3, \ldots, \delta_n = d_n.$$

Lemma **??** shows that each of the $\delta_i$ is non-negative, so this is indeed a well-defined polynomial. One can check that

$$\begin{aligned} \mathrm{mdeg}(a\sigma_1^{\delta_1} \cdots \sigma_n^{\delta_n}) &= (\delta_1 + \delta_2 + \cdots + \delta_n, \delta_2 + \cdots + \delta_n, \ldots, \delta_n) \\ &= (d_1, \ldots, d_n) \end{aligned}$$

Hence

$$\begin{aligned} \mathrm{lt}(f) &= a x_1^{d_1} \cdots x_n^{d_n} \\ &= \mathrm{lt}(a\sigma_1^{\delta_1} \cdots \sigma_n^{\delta_n}) \end{aligned}$$

Therefore the difference $f - a\sigma_1^{\delta_1} \cdots \sigma_n^{\delta_n}$ has multidegree smaller than that of $f$, as required. $\square$

## 1.1 Examples

The proof of Theorem **??** has the nice property that it actually gives an algorithm to write a given symmetric polynomial in terms of elementary symmetric polynomials. Let's see how it works.

**Example:** Let's start with a simple example where $n = 2$. Here there are two elementary symmetric polynomials:

$$\sigma_1(x, y) = x + y, \quad \sigma_2(x, y) = xy.$$

Let's write the symmetric polynomial

$$f = x^4y^2 + x^2y^4$$

in terms of $\sigma_1$ and $\sigma_2$.

Acoording to the proof of Theorem **??** we need to start by finding the leading term of $f$. The terms of $f$ have the following multidegrees:

$$\begin{aligned} \mathrm{mdeg}(x^4y^2) &= (4, 2) \\ \mathrm{mdeg}(x^2y^4) &= (2, 4). \end{aligned}$$

So $\mathrm{lt}(f) = x^4y^2$. In this term we have $d_1 = 4$, $d_2 = 2$. This gives

$$\begin{aligned} \delta_1 = d_1 - d_2 &= 2 \\ \delta_2 = d_2 &= 2. \end{aligned}$$

So let

$$\begin{aligned} f_1 &= f - \sigma_1^{\delta_1}\sigma_2^{\delta_2} \\ &= f - (x+y)^2(xy)^2 \\ &= x^4y^2 + x^2y^4 - (x^4y^2 + 2x^3y^3 + x^2y^4) \\ &= -2x^3y^3. \end{aligned}$$

The multidegree of $f_1$ is now $(3,3)$, so it is smaller than that of $f$, as we want.

Now we apply the same procedure to $f_1$. The leading term of $f_1$ is $-2x^3y^3$ which has multidegree $(3,3)$. Here $d_1 = d_2 = 3$, so $\delta_1 = 0$, $\delta_2 = 3$. So let

$$\begin{aligned} f_2 &= f_1 - (-2\sigma_1^{\delta_1}\sigma_2^{\delta_2}) \\ &= f_1 + 2(xy)^3 \\ &= 0. \end{aligned}$$

That is, $f_1 = -2\sigma_2^3$.

Finally we get

$$\begin{aligned} f &= f_1 + \sigma_1^2\sigma_2^2 \\ &= -2\sigma_2^3 + \sigma_1^2\sigma_2^2 \end{aligned}$$

which is the required expresssion for $f$ in terms of elementary symmetric polynmials.

**Example:** Now let's do an example with $n = 3$. Consider the symmetric polynomial

$$f = x^2yz + xy^2z + xyz^2$$

Let's write $f$ in terms of the elementary symmetric polynomials in the variables $x, y, z$.

Recall

$$\begin{aligned} \sigma_1 &= x + y + z \\ \sigma_2 &= xy + xz + yz \\ \sigma_3 &= xyz \end{aligned}$$

We have

$$f = x^2yz + xy^2z + xyz^2$$

so the terms have multidegrees $(2,1,1)$, $(1,2,1)$, and $(1,1,2)$ respectively.

Hence the leading term is $\mathrm{lt}(f) = x^2yz$ which gives us

$$d_1 = 2,\ d_2 = 1,\ d_3 = 1$$

from which we compute

$$\begin{aligned}
\delta_1 &= d_1 - d_2 = 1 \\
\delta_2 &= d_2 - d_3 = 0 \\
\delta_3 &= d_3 = 1
\end{aligned}$$

We set

$$\begin{aligned}
f_1 &= f - \sigma_1^{\delta_1}\sigma_2^{\delta_2}\sigma_3^{\delta_3} \\
&= f - \sigma_1\sigma_3 \\
&= f - (x + y + z)(xyz) \\
&= 0
\end{aligned}$$

So $f = \sigma_1\sigma_3$.

# 2 Applications

## 2.1 Roots and coefficients of polynomials

Let $f \in \mathbf{C}[x]$ be a monic polynomial of degree $d$:

$$f = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$$

According to the Fundamental Theorem of Algebra we know that $f$ has $d$ complex roots $z_1, \ldots, z_d$ (possibly repeated) and that $f$ factors as

$$f = (x - z_1) \cdots (x - z_d).$$

Expanding out the right-hand side, we see that the coefficients of $f$ are (plus or minus) the values of the elementary symmetric polynomials in $d$ variables, evaluated on the roots of $f$. That is, the coefficient $a_k$ of $x^k$ in $f$ can be written as

$$a_k = (-1)^{d-k}\sigma_{d-k}(z_1, \ldots, z_d).$$

For example this includes the well-known equalities

$$\begin{aligned} a_{d-1} &= -\sigma_1(z_1, \ldots, z_d) \\ &= -(z_1 + \ldots + z_d) \\ a_0 &= (-1)^d \sigma_d(z_1, \ldots, z_d) \\ &= (-1)^d z_1 \cdots z_d. \end{aligned}$$

Theorem **??** then tells us the following:

**Corollary 2.1.** *For a polynomial $f \in \mathbf{C}[x]$, any symmetric polynomial function of the roots of $f$ can be expressed in terms of the coefficients of $f$.*

**Example:** In Week 4 we stated (but did not prove) one formula for the discriminant of a one-variable polynomial $f$ as follows:

$$\mathrm{Disc}(f) = a_d^{2d-2} \prod_{1 \le i < j \le d} (\lambda_i - \lambda_j)^2$$

where $a_d$ is the leading coefficient of $f$ and the $\lambda_i$ are the roots of $f$.

On the face of it, this formula doesn't make it at all clear that $\mathrm{Disc}(f)$ can actually be expressed just in terms of the coefficients of $f$. However, the right-hand side of the formula is a symmetric polynomial function of the roots $\lambda_i$: a permutation may rearrange some of the factors $(\lambda_i - \lambda_j)$ or send $(\lambda_i - \lambda_j)$ to $(\lambda_j - \lambda_i)$, but once we square and take the product we get the same thing. Therefore Corollary **??** guarantees that the expression on the right-hand side is indeed a polnyomial in the coefficients of $f$.

**Note:** Remember that our actual defintion of $\mathrm{Disc}(f)$ was in terms of the resultant of $f$ and $f'$, and this is automatically a function of the coefficients of $f$.)

## 2.2 Power sums and characteristic polynomial of a matrix

**Definition 2.2.** *Fix a positive integer* $k$. *In the ring* $\mathbb{C}[x_1, \ldots, x_n]$, *the* $k$**-th power sum polynomial** $\pi_k(x_1, \ldots, x_n)$ *is defined as*

$$\pi_k(x_1, \ldots, x_n) = x_1^k + \ldots + x_n^k$$

Clearly $\pi_k(x_1, \ldots, x_n)$ is a symmetric polynomial in the variables $x_i$, so by Theorem **??** we can write it in terms of the elementary symmetric polynomials $\sigma_i(x_1, \ldots, x_n)$. The notable feature of power sum polynomials is that the converse is also true, as shown by the following theorem.

**Theorem 2.3** (Newton's identities)**.** *Let* $\pi_1, \ldots, \pi_n$ *be the first* $n$ *power sum polynomials in* $n$ *variables, and let* $\sigma_k$ *be one of the elementary symmetric polynomials in* $n$ *variables.*

$$k\sigma_k = \sum_{i=1}^{k} (-1)^{i-1} \sigma_{k-i} \pi_i$$

Applying this recursively we get:

**Corollary 2.4.** *There is a polynomial* $B \in \mathbb{C}[x_1, \ldots, x_n]$ *such that*

$$\sigma_k = B(\pi_1, \ldots, \pi_n)$$

We won't give the proof of the theorem, but let's work out an example.
**Example:** Let's take $n = 3$. The theorem says that the elementary symmetric polynomials

$$\begin{aligned}
\sigma_1(x, y, z) &= x + y + z \\
\sigma_2(x, y, z) &= xy + xz + yz \\
\sigma_3(x, y, z) &= xyz
\end{aligned}$$

should all be expressible as polynomials in the power sum polynomials

$$\begin{aligned}
\pi_1(x, y, z) &= x + y + z \\
\pi_2(x, y, z) &= x^2 + y^2 + z^2 \\
\pi_3(x, y, z) &= x^3 + y^3 + z^3
\end{aligned}$$

Evidently $\pi_1 = \sigma_1$ so there is nothing to prove here. What about the other cases?
For $\sigma_2$ we compute:

$$\begin{aligned}
\pi_1^2 &= (x + y + z)^2 \\
&= x^2 + y^2 + z^2 + 2(xy + xz + yz) \\
&= \pi_2 + 2\sigma_2.
\end{aligned}$$

So $\sigma_2 = \frac{1}{2}(\pi_1^2 - \pi_2)$.

For $\sigma_3$ we compute:

$$\begin{aligned}
\pi_1(\sigma_2 - \pi_2) &= (x + y + z)(xy + xz + yz - x^2 - y^2 - z^2) \\
&= 3xyz - x^3 - y^3 - z^2 \\
&= 3\sigma_3 - \pi_3
\end{aligned}$$

So we have

$$\begin{aligned}
3\sigma_3 &= \pi_1\sigma_2 - \pi_1\pi_2 + \pi_3 \\
&= \pi_1(\frac{1}{2}(\pi_1^2 - \pi_2)) - \pi_1\pi_2 + \pi_3
\end{aligned}$$

which gives

$$\sigma_3 = \frac{1}{6}\pi_1^3 - \frac{1}{2}\pi_1\pi_2 + \frac{1}{3}\pi_3.$$

To apply this to linear algebra, recall that for a $d \times d$ matrix $M$, the **characteristic polynomial of** $M$ is defined as

$$p_M(\lambda) = (-1)^d \det(M - \lambda I_d)$$

where $I_d$ is the $d \times d$ identity matrix. This polynomial is important because its roots are exactly the eigenvalues of $M$.

In practice, however, as $M$ becomes large it is computationally expensive to calculate $p_M(\lambda)$ directly as a determinant. The theory of symmetric polynomials gives a shortcut to do this. Let's see how it works.

Suppose the eigenvalues of $M$ are $\lambda_1, \ldots, \lambda_d$. These are exactly the roots of $p_M(\lambda)$, which is a monic polynomial, so as explained above, the coefficients of $p_m(\lambda)$ are exactly (plus or minus) the elementary symmetric functions evaluated on the $\lambda_i$:

$$p_m(\lambda) = \lambda^d + a_{d-1}\lambda^{d-1} + \cdots + a_1\lambda + a_0$$

where

$$a_k = (-1)^{d-k}\sigma_{d-k}(\lambda_1, \ldots, \lambda_d).$$

For a nonnegative integer $k$, consider the matrix $M^k$. Its eigenvalues are $\lambda_1^k, \ldots, \lambda_d^k$. We are going to look at the **traces** of the matrices $M^k$. By definition the trace is equal to the sum of the diagonal elements, which is straightforward to compute.

On other other hand, by diagonalisation we can also see that $\operatorname{tr}(M^k)$ equals the sum of the eigenvalues. So for any nonnegative integer $k$ we have

$$\begin{aligned}\operatorname{tr}(M^k) &= \lambda_1^k + \cdots + \lambda_d^k \\ &= \pi_k(\lambda_1, \ldots, \lambda_d).\end{aligned}$$

Therefore, using the formulae for the elementary symmetric polynomials $\sigma_k$ in terms of $\pi_1, \ldots, \pi_d$, we can compute the coefficients $a_k$ of the characteristic polynomial in terms of the traces $\mathrm{tr}(M^k)$.

**Example:** Consider the $3 \times 3$ matrix

$$M = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 2 \end{pmatrix}$$

We have

$$M^2 = \begin{pmatrix} 1 & 4 & 2 \\ -1 & 1 & 3 \\ -3 & -2 & 4 \end{pmatrix}, \quad M^3 = \begin{pmatrix} -1 & 6 & 8 \\ -4 & -1 & 7 \\ -7 & -8 & 6 \end{pmatrix}.$$

So

$$\operatorname{tr}(M) = 4,\ \operatorname{tr}(M^2) = 6,\ \operatorname{tr}(M^3) = 4.$$

If the characterstic polynomial of $M$ is

$$p_M(\lambda) = \lambda^3 + a_2\lambda^2 + a_1\lambda + a_0$$

then using the identities from the previous section we see that

$$\begin{aligned}a_2 = -\sigma_1 &= -\pi_1 \\ &= -\operatorname{tr}(M) \\ &= -4 \\ a_1 = \sigma_2 &= \frac{1}{2}\left(\pi_1^2 - \pi_2\right) \\ &= \frac{1}{2}\left((\operatorname{tr}(M))^2 - \operatorname{tr}(M^2)\right) \\ &= 5 \\ a_0 = -\sigma_3 &= -\frac{1}{6}\left(\pi_1^3 - 3\pi_1\pi_2 + 2\pi_3\right) \\ &= -\frac{1}{6}\left((\operatorname{tr}(M))^3 - 3\operatorname{tr}(M)\operatorname{tr}(M^2) + 2\operatorname{tr}(M^3)\right) \\ &= 0\end{aligned}$$

So the characteristic polynomial of $M$ is

$$p_M(\lambda) = \lambda^3 - 4\lambda^2 + 5\lambda.$$

The eigenvalues of $M$ are the roots of this cubic, which are not difficult to find:

$$\lambda_1 = 0,\ \lambda_2 = 2 + i,\ \lambda_3 = 2 - i.$$