Symmetric polynomials

§1. Statement of the main theorem

Let R be a commutative ring, and let $R[x_1,\dots,x_n]$ be the polynomial ring over R in n variables $x_1,\dots,x_n$. The *elementary symmetric polynomials* in $x_1,\dots,x_n$ are defined by the equality

$$\sum_{i=0}(-1)^i s_i T^{n-i} = T^n - s_1 T^{n-1} + \cdots + (-1)^{n-1} s_{n-1} T + (-1)^n s_n = \prod_{j=1}^{n}(T - x_j)$$

in the polynomial ring $R[x_1,\dots,x_n][T]$. In other words $s_0 = 1$ and

$$s_i = s_i(\underline{x}) = s_i(x_1,\cdots,x_n) = \sum_{I \in S_n\cdot(1,\dots,i)\subset \mathbb{N}^n} \underline{x}^I, \quad 1 \le i \le n$$

where $S_n \cdot (1,\dots,i)$ is the orbit in $\mathbb{N}^n$ of the standard action of the permutation group S_n on $\mathbb{N}^n$ and $\underline{x}^I := x_1^{i_1}\cdots x_n^{i_n}$ for all $I = (i_1,\dots,i_n) \in \mathbb{N}^n$.

The group S_n operates on $R[x_1,\dots,x_n]$ by permuting the variables. Polynomials in $R[x_1,\dots,x_n]$ fixed by all elements of S_n are called *symmetric polynomials* (in variables $x_1,\dots,x_n$ with coefficients in R.) Let $S = R[x_1,\dots,x_n]^{S_n}$ be the subring of $R[x_1,\dots,x_n]$ consisting of all symmetric polynomials. Clearly S is the direct sum of all homogenous symmetric polynomials, and $s_1,\dots,s_n$ are elements of S.

The main theorem on symmetric polynomials asserts that S is a polynomial ring in $s_1,\dots,s_n$, and $R[x_1,\dots,x_n]$ is a free S-module of rank $n!$.

(1.1) THEOREM *Let R be a commutative ring, and let S be the subring of $R[x_1,\dots,x_n]$ consisting of all symmetric polynomials in $R[x_1,\dots,x_n]$.*

(a) *Let $R[y_1,\dots,y_n]$ be the polynomial ring in variables $y_1,\dots,y_n$. The R-algebra homomorphism*

$$\alpha\colon R[y_1,\dots,y_n] \longrightarrow S$$

which sends Y_i to s_i is an isomorphism.

(b) *The polynomial ring $R[x_1,\dots,x_n]$ is a free module of rank $n!$ over S, and the set of monomials*

$$x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n} \quad \text{with} \quad 0 \le i_\nu \le \nu - 1 \quad \forall \nu = 1,\dots,n$$

form a set of free generators.

Recall that the *lexicographic order* on monomials in $R[x_1,\dots,x_n]$ is the linearly order on the set of all monic monomials (or *terms*) such that $x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n} \prec x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$ if either $i_\nu = j_\nu$ for all ν, or if there exists an natural number ν_0 with $1 \le \nu_0 \le n$ such that $i_\nu = j_\nu$ for all $\nu < \nu_0$. and $i_{\nu_0} < j_{\nu_0}$.

§2. Proof of the main theorem 1.1

The proof of 1.1 (a) is quite easy using the lexicographic order. We will prove 1.1 (b) by induction on n. The key observation is that the R-subalgebra of $R[x_1,\dots,x_n]$ generated by $s_1,\dots,s_n$ and x_n is equal to the subring $R[x_1,\dots,x_n]^{S_{n-1}}$, consisting of all polynomials fixed under all permutations of the first $n-1$ variables $x_1,\dots,x_n$.

(2.1) PROOF OF 1.1 (a). The largest term in a monomial $s_1^{j_1}\cdots s_n^{j_n}$ with respect to the lexicographic ordering is

$$x_1^{j_1+j_2+\cdots+j_n}x_2^{j_2+\cdots+j_n}x_{n-1}^{j_{n-1}+j_n}x_n^{j_n}.$$

It follows that any two distinct monomials in the elementary symmetric polynomials $s_1,\ldots,s_n$ have distinct highest terms. Therefore the R-algebra homomorphism α is injective.

The surjectivity of α also follows from the above consideration: Given any non-zero symmetric polynomial $f(x_1,\ldots,x_n)\in S$, there exists a unique monomial $a_J\cdot s_1^{j_1}\cdots s_n^{j_n}$ in $s_1,\ldots,s_n$ with $a_J\in R$ and $(j_1,\ldots,j_n)\in\mathbb{N}^n$ such that the highest term in $f(x_1,\ldots,x_n)-a_J\cdot s_1^{j_1}\cdots s_n^{j_n}$ is strictly smaller than the largest term in $f(x_1,\ldots,x_n)$. The surjectivity of α follows from induction (on the lexicographic ordering of the highest term of $f(x_1,\ldots,x_n)$.)

(2.1.1) **REMARK** We could have used other linear orders of monic monomials, for instance the *degree lexicographic order*: $x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}\prec x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$ under the degree lexicographic order if either $i_\nu=j_\nu$ for all ν, or if $\sum_\nu i_\nu<\sum_\nu j_\nu$, or if $\sum_\nu i_\nu=\sum_\nu j_\nu$ and there exists an natural number ν_0 with $1\le\nu_0\le n$ such that $i_\nu=j_\nu$ for all $\nu<\nu_0$. and $i_{\nu_0}<j_{\nu_0}$. Using the *degree lexicographic order* means that in the above proof we assume that $f(x_1,\ldots,x_n)$ is homogeneous of degree m.

(2.2) PROOF OF 1.1 (b). Let $t_1,\ldots,t_{n-1}$ be the elementary symmetric polynomials of $x_1,\ldots,x_{n-1}$. We know from part (a) that $R[x_1,\ldots,x_n]^{S_{n-1}}=R[t_1,\ldots,t_{n-1},x_n]$, a polynomial ring over R in n variables. By induction on n, we may and do assume that $R[x_1,\ldots,x_n]$ is a free module over $R[t_1,\ldots,t_{n-1},x_n]$ of rank $(n-1)!$ with generators

$$\left\{x_1^{i_1}x_2^{i_2}\cdots x_{n-1}^{i_{n-1}}\,:\,0\le i_\nu\le\nu-1\quad\forall\nu=1,\ldots,n-1\right\},$$

so it suffices to show that $R[t_1,\ldots,t_{n-1},x_n]$ is a free module over $R[s_1,\ldots,s_n]$ of rank n with free generators $\left\{x_n^j\,:\,0\le j\le n-1\right\}$. We will produce an automorphism β of the R-algebra $R[t_1,\ldots,t_{n-1},x_n]$ such that

$$(\dagger)\quad \beta(x_n)=x_n,\quad \beta(s_i)=t_i\ \text{ for }\ i=1,\ldots,n-1$$

and $\beta(s_n)$ is a polynomial of degree n in x_n and the coefficient of x_n^n is ±1. The desired conclusion follows immediately from the existence of such an automorphism β.

Clearly we have

$$s_i=t_i+t_{i-1}x_n\quad\forall\ 1\le i\le n-1\quad\textit{and}\quad s_n=t_{n-1}x_n.$$

An easy induction gives

$$(*)\quad t_k=(-1)^kx_n^k+\sum_{j=0}^{k-1}(-1)^j s_{k-j}x_n^j\quad\text{for }\ k=1,2,\ldots,n-1.$$

(If follows immediately that the R-subalgebra of $R[x_1,\ldots,x_n]$ generated by $s_1,\ldots,s_{n-1}$ and x_n is equal to $R[t_1,\ldots,t_{n-1},x_n]$, but we will not use this in the rest of the proof.)

How do we construct β? Suppose we have an R-algebra homomorphism β which satisfies condition $(\dagger)$. Apply β to the last displayed equations would than give

$$(\ddagger)\quad \beta(t_k)=(-1)^kx_n^k+\sum_{j=0}^{k-1}(-1)^j t_{k-j}x_n^j\quad\text{for }\ k=1,2,\ldots,n-1.$$

So we *define* β to be the unique $R[x_n]$-algebra homomorphism from $R[t_1,\dots,t_{n-1},x_n]$ to itself which satisfies (‡). An easy calculation shows that indeed β satisfies the equations (†), and

$$\beta(s_n) = (-1)^{n-1}x_n^n + \sum_{1\le k\le n-1}(-1)^{k-1}t_{n-k}x_n^k.$$

We have proved 1.1 (b).

(2.2.1) REMARK (a) It is quite easy to show that $R[t_1,\dots,t_{n-1},x_n] = R[s_1,\dots,s_{n-1},x_n]$ is generated by $1,x_n,x_n^2,\dots,x_n^{n-1}$ as an $R[s_1,\dots,s_n]$-module, simply because $x_n^n + \sum_{1\le i\le n}(-1)^i s_i x_n^{n-i} = 0$. Proving that there is no non-trivial linear relation between $1,x_n,x_n^2,\dots,x_n^{n-1}$ over $R[s_1,\dots,s_{n-1},x_n]$ is requires more effort.
(b) We sketch an alternative proof of 1.1 (b). First prove it when $R=\mathbb{Z}$; the general case follows immediately, by taking the tensor product with R over $\mathbb{Z}$. It suffices to show that there is no non-trivial linear relation between $1,x_n,x_n^2,\dots,x_n^{n-1}$ over $\mathbb{Q}[s_1,\dots,s_n]$. There are at least two ways: one can either use the fact that $T^n + \sum_{1\le i\le n}(-1)^i s_i T^{n-i}$ is an irreducible polynomial in th polynomial ring $\mathbb{Q}[s_1,\dots,s_n,T]$, or invoke Galois theory noting that $[S_n : S_{n-1}] = n$.

(2.2.2) REMARK When R is a field, the phenomenon describe in 1.1 is the following. Starting with the monic polynomial $f(T) = T^n + \sum_{1\le i\le n}(-1)^i s_i T^{n-i}$ over the fraction field $K = R(s_1,\dots,s_n)$ of the polynomial ring $R[s_1,\dots,s_n]$, which is irreducible over K. Adjoint a root x_n of $f(T)$ to K, the polynomial $f(T)$ factors into a product $f(T) = (T-x_n)\cdot(T^{n-1} + \sum_{j=1}^{n-1}(-1)^j t_j T^{n-1-j})$ over $K_1 := R(s_1,\dots,s_n,x_n) = R(x_n)(t_1,\dots,t_{n-1})$. Moreover $\{1,x_n,x_n^2,\dots,x_n^{n-1}\}$ is a basis of K_1 over K. Now we can do the same thing for the polynomial $f_1(T) := T^{n-1} + \sum_{j=1}^{n-1}(-1)^j t_j T^{n-1-j}$ over the field K_1. Iterating this procedure $n-1$ times gives us a field $K_{n-1} = R(x_1,\dots,x_n)$, of degree $n!$ over K. The basis of K_1 over K obtained from this inductive procedure is the one described in theorem 1.1 (b).