SYMMETRIC POLYNOMIALS

KEITH CONRAD

1. Introduction

Let $F$ be a field. A polynomial $f(X_1, \ldots, X_n) \in F[X_1, \ldots, X_n]$ is called symmetric if it is unchanged by all permutations of its variables:

$$f(X_1, \ldots, X_n) = f(X_{\sigma(1)}, \ldots, X_{\sigma(n)})$$

for every permutation $\sigma$ of $\{1, \ldots, n\}$.

**Example 1.1.** The sum $X_1 + \cdots + X_n$ and product $X_1 \cdots X_n$ are both symmetric, as are the power sums $X_1^r + \cdots + X_n^r$ for all $r \geq 1$.

**Example 1.2.** Let $f(X_1, X_2, X_3) = X_1^5 + X_2 X_3$. This polynomial is unchanged if we interchange $X_2$ and $X_3$, but if we interchange $X_1$ and $X_3$ then $f$ becomes $X_3^5 + X_2 X_1$, which is not $f$. This polynomial is only "partially symmetric."

An important collection of symmetric polynomials occurs as the coefficients in the polynomial

$$(1.1) \quad (T - X_1)(T - X_2) \cdots (T - X_n) = T^n - s_1 T^{n-1} + s_2 T^{n-2} - \cdots + (-1)^n s_n.$$  

Here $s_1$ is the sum of the $X_i$'s, $s_n$ is their product, and more generally

$$s_k = \sum_{1 \leq i_1 < \cdots < i_k \leq n} X_{i_1} \cdots X_{i_k}$$

is the sum of the products of the $X_i$'s taken $k$ terms at a time. The polynomial $s_k$ is symmetric in $X_1, \ldots, X_n$ and is called the $k$th elementary symmetric polynomial – or $k$th elementary symmetric function – in $X_1, \ldots, X_n$.

**Example 1.3.** Let $\alpha = \frac{3 + \sqrt{5}}{2}$ and $\beta = \frac{3 - \sqrt{5}}{2}$. Although $\alpha$ and $\beta$ are not rational, their elementary symmetric polynomials are: $s_1 = \alpha + \beta = 3$ and $s_2 = \alpha \beta = 1$.

**Example 1.4.** Let $\alpha, \beta, \gamma$ be the three roots of $T^3 - T - 1$, so

$$T^3 - T - 1 = (T - \alpha)(T - \beta)(T - \gamma).$$

Multiplying out the right side and equating coefficients on both sides, the elementary symmetric functions of $\alpha, \beta,$ and $\gamma$ are $s_1 = \alpha + \beta + \gamma = 0$, $s_2 = \alpha \beta + \alpha \gamma + \beta \gamma = -1$, and $s_3 = \alpha \beta \gamma = 1$.

Our goal is to prove the following theorem.

**Theorem 1.5.** The set of symmetric polynomials in $F[X_1, \ldots, X_n]$ is $F[s_1, \ldots, s_n]$. That is, every symmetric polynomial in $n$ variables is a polynomial in the elementary symmetric functions of those $n$ variables.
**Example 1.6.** In two variables, the polynomial $X^3 + Y^3$ is symmetric in $X$ and $Y$. As a polynomial in $s_1 = X + Y$ and $s_2 = XY$, 

$$X^3 + Y^3 = (X + Y)^3 - 3XY(X + Y) = s_1^3 - 3s_1s_2.$$ 

There are several ways to prove Theorem 1.5. For example, it is proved in [5, Chap. IV, Sect. 6] by a double induction on $n$ and on the degree of the polynomial. The theorem can also be proved using Galois theory, transcendental field extensions, and integral ring extensions. The proof we will give, based on [2, Sect. 7.1], provides an explicit algorithm that turns a symmetric polynomial in $X_1, \ldots, X_n$ into a polynomial in $s_1, \ldots, s_n$.

2. **Lexicographic ordering on $F[X_1, \ldots, X_n]$**

In $F[X]$, many theorems are proved using induction on the degree of polynomials. The degree is a nonnegative integer associated to each nonzero polynomial $f(X)$: it is the largest $n \geq 0$ such that $f(X)$ contains a monomial $a_nX^n$ where $a_n \neq 0$ in $F$. The monomials $1, X, X^2, X^3, \ldots$ (all with coefficient 1) are ordered by degree as $1 < X < X^2 < \ldots$, and $\deg(f)$ is the largest monomial appearing in $f$ with a nonzero coefficient.

On $F[X_1, \ldots, X_n]$ we will use a total ordering on the monomials $X_1^{i_1} \cdots X_n^{i_n}$ and associate to that ordering an analogue on $F[X_1, \ldots, X_n]$ of the degree on $F[X]$.

**Definition 2.1.** For $i = (i_1, \ldots, i_n)$ and $j = (j_1, \ldots, j_n)$ in $\mathbb{N}^n$, set $i < j$ if, for the first index $r$ such that $i_r \neq j_r$, we have $i_r < j_r$. Write $i \leq j$ if $i < j$ or $i = j$.

**Example 2.2.** In $\mathbb{N}^4$, $(3, 0, 2, 4) < (5, 1, 1, 3)$ and $(3, 0, 2, 4) < (3, 0, 3, 1)$.

**Example 2.3.** In $\mathbb{N}^n$, $0 < i$ for all $i \neq 0$.

This way of ordering $n$-tuples in $\mathbb{N}^n$ is called the lexicographic (i.e., dictionary) ordering since it resembles the way words are ordered in the dictionary alphabetically if we think of one word as “less” than another if it comes earlier in the dictionary. The “greater” word comes later. Alphabetical order first compares words by the first letter, if the first letters are the same then the words are compared by the second letter, and so on. While words in a dictionary have varying length, we are using lexicographic ordering only to compare sequences in $\mathbb{N}$ with the same number of terms.

**Theorem 2.4.** Lexicographic ordering on $\mathbb{N}^n$ has the following properties.

1. (Total ordering) For all $i$ and $j$, exactly one of $i = j$ or $i < j$ or $j < i$ holds.
2. (Transitivity) If $i < j$ and $j < k$ then $i < k$. The same is true with $\leq$ in place of $<$.
3. (Compatibility with addition) If $i \leq i'$ and $j \leq j'$ then $i + j \leq i' + j'$, and if either inequality in the hypothesis is strict then the inequality in the conclusion is strict.

**Proof.**

1. If $i \neq j$, then there is an $r$ where $i_r \neq j_r$ in $\mathbb{N}$. Let $r$ be the least index where this happens. If $i_r < j_r$ then $i < j$, and if $j_r < i_r$ then $j < i$.

2. Let $r$ be the least index where $i_r, j_r$, and $k_r$ are not all equal. We must have $i_r \neq j_r$ or $j_r \neq k_r$ (if both were equalities then $i_r = j_r = k_r$, which isn’t true). Since earlier coordinates in $i$, $j$, and $k$ are all equal, either $i_r < j_r$ or $j_r < k_r$ because $i < j$ and $j < k$. Therefore $i_r \leq j_r \leq k_r$ with at least one inequality being strict, so $i_r < k_r$ and earlier coordinates in $i$ and $k$ are equal. Thus $i < k$.

\footnote{Historically, Theorem 1.5 used to be part of the mathematical development leading to Galois theory, which would make a proof of Theorem 1.5 by Galois theory circular, but Galois theory in its modern form does not require Theorem 1.5.}
This result for \( \leq \) is the same argument as with \(<\) except we have the extra cases where \(i\) and \(j\) may coincide or \(i\) and \(k\) may coincide, which makes things easier.

(3) Rather than take cases based on where \(i\) and \(i'\) may first differ or where \(j\) and \(j'\) may first differ, observe that \(i \leq i' \Rightarrow i + k \leq i' + k\) for all \(k\): this is obvious when \(i = i'\), and when \(i < i'\) the only way \(i_r + k_r\) differs from \(i'_r + k_r\) is if \(i_r \neq i'_r\), and the first time this happens we have \(i_r < i'_r\), so \(i_r + k_r < i'_r + k_r\).

Now we use that twice together with the transitivity in (2). If \(i \leq i'\) and \(j \leq j'\), then
\[
i + j \leq i' + j = j + i' \leq j' + i' = i' + j' \implies i + j \leq i' + j'.
\]

The case of \(<\) in place of \(\leq\) is analogous. \(\square\)

A polynomial \(f \in F[X_1, \ldots, X_d]\) is a sum of the form
\[
f = \sum_{i_1, \ldots, i_n \geq 0} c_{i_1, \ldots, i_n} X_1^{i_1} \cdots X_n^{i_n}
\]
where \(c_{i_1, \ldots, i_n} \in F\) and only finitely many coefficients can be nonzero. Abbreviate this sum to multi-index form as \(\sum c_i X^i\), where \(X^i := X_1^{i_1} \cdots X_n^{i_n}\) for \(i = (i_1, \ldots, i_n)\). Note \(X^i X^j = X^{i+j}\). (In the notation \(\sum c_i\), only finitely many terms are nonzero.) When \(f \neq 0\), lexicographic ordering lets us compare the different nonzero monomials appearing in \(f\), which leads to the following concepts that generalize degree and leading terms on \(F[X]\).

**Definition 2.5.** Write \(f\) in \(F[X_1, \ldots, X_n]\) as \(\sum c_i X^i\). If \(f \neq 0\), the multidegree of \(f\) is the lexicographically largest index in \(\mathbb{N}^n\) of a nonzero monomial in \(f\):
\[
mdeg f = \max\{i : c_i \neq 0\} \in \mathbb{N}^n.
\]
The multidegree of the zero polynomial is not defined. If \(f \neq 0\) and \(\text{mdeg} f = d\), we call \(c_d X^d\) the leading term of \(f\) and \(c_d\) the leading coefficient of \(f\), written \(c_d = \text{lead} f\).

We can separate the leading term of a nonzero \(f\) of multidegree \(d\) from its other nonzero terms to get
\[
f = c_d X^d + \sum_{i < d} c_i X^i.
\]
Because multidegrees are totally ordered, a nonzero polynomial in \(F[X_1, \ldots, X_n]\) has a unique leading term and a unique leading coefficient in \(F\).

**Example 2.6.** In \(\mathbb{Q}[X_1, X_2]\), let \(f = 7X_1X_2^5 + 3X_2^8 + 9\). Since \(\max((1, 5), (0, 8), (0, 0)) = (1, 5)\) in \(\mathbb{N}^2\), \(\text{mdeg}(f) = (1, 5)\) and \(\text{lead}(7X_1X_2^5 + 3X_2^8 + 9) = 7\).

**Example 2.7.** In \(F[X_1, \ldots, X_n]\), \(\text{mdeg}(X_1) = (1, 0, \ldots, 0)\) and \(\text{mdeg}(X_n) = (0, 0, \ldots, 1)\).

**Example 2.8.** The polynomials with multidegree \(0\) are the nonzero constant.

**Example 2.9.** The multidegrees of the elementary symmetric polynomials are
\[
\text{mdeg}(s_1) = (1, 0, 0, \ldots, 0),
\]
\[
\text{mdeg}(s_2) = (1, 1, 0, \ldots, 0),
\]
\[\vdots\]
\[
\text{mdeg}(s_n) = (1, 1, 1, \ldots, 1).
\]
For \(k = 1, \ldots, n\), the leading term of \(s_k\) is \(X_1 \cdots X_k\), so the leading coefficient of \(s_k\) is 1.
Our definition of multidegree is specific to calling $X_1$ the “first” variable and $X_n$ the “last” variable. Despite its ad hoc nature (there is nothing intrinsic about making $X_1$ the “first” variable), the multidegree is useful since it permits us to prove theorems about all multivariable polynomials by ordering them according to their multidegree.

The following theorem shows that a number of standard properties of the degree of polynomials in one variable carry over to multidegrees of multivariable polynomials.

**Theorem 2.10.** For nonzero $f$ and $g$ in $F[X_1, \ldots, X_n]$, $mdeg(fg) = mdeg(f) + mdeg(g)$ in $\mathbb{N}^n$ and $\text{lead}(fg) = (\text{lead } f)(\text{lead } g)$.

For $f$ and $g$ in $F[X_1, \ldots, X_n]$, $mdeg(f + g) \leq \max(mdeg f, mdeg g)$ and if $mdeg f < mdeg g$ then $mdeg(f + g) = mdeg g$.

**Proof.** We will prove the first result and leave the second to the reader.

Let $mdeg f = \mathbf{n}$ and $mdeg g = \mathbf{m}$, say $f = c_\mathbf{n}X^n + \sum_{i < \mathbf{n}} c_iX^i$ with $c_\mathbf{n} \neq 0$ and $g = c'_\mathbf{m}X^\mathbf{m} + \sum_{j < \mathbf{m}} c'_jX^j$ with $c'_\mathbf{m} \neq 0$. This amounts to pulling out the top multidegree terms of $f$ and $g$. Then $fg$ has a nonzero term $c_\mathbf{n}c'_\mathbf{m}X^{n+\mathbf{m}}$ and every other term has multidegree $\mathbf{n} + \mathbf{j} + \mathbf{m} + \mathbf{i}$, or $i + j$ where $i < \mathbf{n}$ and $j < \mathbf{m}$. By Theorem 2.4, all these other multidegrees are less than $\mathbf{n} + \mathbf{m}$, so $mdeg(fg) = \mathbf{n} + \mathbf{m} = mdeg f + mdeg g$ and $\text{lead}(fg) = c_\mathbf{n}c'_\mathbf{m} = (\text{lead } f)(\text{lead } g)$. \qed

**Example 2.11.** If $f$ and $g$ in $F[X_1, \ldots, X_n]$ are different polynomials with the same leading term then let’s show $mdeg(f - g) < mdeg f$. Writing the common leading term as $c_\mathbf{n}X^n$,

$$f = c_\mathbf{n}X^n + \sum_{i < \mathbf{n}} a_iX^i, \quad g = c_\mathbf{n}X^n + \sum_{i < \mathbf{n}} b_iX^i$$

where the sums over $i < \mathbf{n}$ have finitely many nonzero terms. Subtracting,

$$f - g = \sum_{i < \mathbf{n}} (a_i - b_i)X^i$$

where the right side has finitely many nonzero terms. Each $i$ on the right side is less than $\mathbf{n}$, so $mdeg(f - g) < \mathbf{n}$.

In $\mathbb{N}$ with its usual ordering, there are finitely many elements below a given element, but this is not true in $\mathbb{N}^n$ for $n \geq 2$ with lexicographic ordering: there can be infinitely many $n$-tuples below some $n$-tuple. For instance, $(0, b) < (1, 0)$ for all $b \in \mathbb{N}$. In terms of lexicographic ordering on $F[X, Y]$, where $X > Y$, this corresponds to saying

$$(2.1) \quad 1 < Y < Y^2 < Y^3 < \cdots < Y^b < \cdots < X.$$

The lexicographic ordering on $\mathbb{N}^n$ shares with $\mathbb{N}$ the following important ordering property.

**Theorem 2.12.** When $\mathbb{N}^n$ has lexicographic ordering, each nonempty subset of $\mathbb{N}^n$ has a least element. In particular, every strictly decreasing sequence $\mathbf{d}_1 > \mathbf{d}_2 > \mathbf{d}_3 > \cdots$ in $\mathbb{N}^n$ has finite length.

For example, even though infinitely many elements of $\mathbb{N}^2$ are lexicographically less than $(1, 0)$, a strictly decreasing sequence that starts at $(1, 0)$ is finite: the second term in the sequence must be some $(0, b)$ and only finitely many elements of $\mathbb{N}^2$ are less than $(0, b)$.

**Proof.** Let $S$ be a nonempty subset of $\mathbb{N}^n$. The first coordinates of elements in $S$ are a nonempty subset of $\mathbb{N}$ and thus have a least element $\ell_1$. If $n = 1$ then $\ell_1$ is the least element of $S$. For $n \geq 2$, among the elements of $S$ with first coordinate $\ell_1$, their second coordinates are a nonempty subset of $\mathbb{N}$ and thus have a least element $\ell_2$. If $n = 2$ then
(ℓ1, ℓ2) is the least element of S. For n ≥ 3, among the elements of S with first coordinate ℓ1 and second coordinate ℓ2, their third coordinates are a nonempty subset of N and thus have a least element ℓ3. Continue this way up through the nth coordinate. The n-tuple (ℓ1, ℓ2, . . . , ℓn) in S is the least element of S in the lexicographic ordering on Nn. It is left to the reader to rewrite this argument as a proper proof by induction on n.

An infinite strictly decreasing sequence d1 > d2 > d3 > · · · in Nn then it has no least element, so a strictly decreasing sequence in Nn must have only finitely many terms. □

The least element in a nonempty subset of Nn is unique since lexicographic ordering on Nn is a total ordering.

3. Proof of Theorem 1.5

Here is a proof of Theorem 1.5 that uses lexicographic ordering on F[X1, . . . , Xn] and Theorem 2.12 to show a recursive process must terminate in finitely many steps.

Proof. We want to show every symmetric polynomial in F[X1, . . . , Xn] is a polynomial in F[s1, . . . , sn]. This is obvious for constant polynomials, which are 0 and polynomials of multidegree (0, 0, . . . , 0).

Let f be a nonconstant symmetric polynomial in F[X1, . . . , Xn] with multidegree d = (d1, . . . , dn). Pull out the leading term of f:

\[ f = c_d X^d + \sum_{i < d} c_i X_i^i = c_d X_1^{d_1} \cdots X_n^{d_n} + \sum_{i < d} c_i X_i^i, \]

where \( c_d \neq 0 \). We will find a polynomial in s1, . . . , sn with the same leading term as f. Its difference with f will be symmetric with smaller multidegree than d.

By Example 2.9 and Theorem 2.10, for nonnegative integers a1, . . . , an,

\[ \text{mdeg}(s_1^{a_1} s_2^{a_2} \cdots s_n^{a_n}) = a_1 \text{mdeg}(s_1) + a_2 \text{mdeg}(s_2) + \cdots + a_n \text{mdeg}(s_n) = (a_1 + a_2 + \cdots + a_n, a_2 + \cdots + a_n, \ldots, a_n). \]

The ith coordinate here is \( a_i + a_{i+1} + \cdots + a_n \). To make this equal \( d = (d_1, \ldots, d_n) \), we must set

\[ a_1 = d_1 - d_2, \quad a_2 = d_2 - d_3, \quad \ldots, \quad a_{n-1} = d_{n-1} - d_n, \quad a_n = d_n. \]

Does this make sense? That is, are \( d_1 - d_2, d_2 - d_3, \ldots, d_{n-1} - d_n, d_n \) all nonnegative? If not then we have a problem. We need to show the coordinates of d = mdeg(f) satisfy

\[ d_1 \geq d_2 \geq \cdots \geq d_n \geq 0. \]

Why does an n-tuple that is the multidegree of a symmetric polynomial satisfy (3.3)?

To appreciate this issue, consider \( f = X_1^5 X_2 + 3X_2 \). The multidegree of f is (1, 5), so the exponents don’t satisfy (3.3). This f is not symmetric, and that is the key point. If we took \( f = X_1^5 X_2 + X_1^5 X_2 \) then f is symmetric and mdeg f = (5, 1) does satisfy (3.3). The verification of (3.3) will depend crucially on f being symmetric.

Since (d1, . . . , dn) is the multidegree of a nonzero monomial in f (the multidegree of the leading term of f), and f is symmetric, every n-tuple with the di’s permuted is also a multidegree of a nonzero monomial in f. (This is how the symmetry of f in the Xi’s is used: under all permutations of the Xi’s, f stays unchanged.) For example, if f in F[x, y] is symmetric and contains the monomial cx^2y^3 then f also contains the monomial cx^3y^2 since permuting x and y turns the first monomial into the second while not changing f.
The multidegree of \( f \) is the largest multidegree of all monomials in \( f \), so \((d_1, \ldots, d_n)\) must be larger in \( \mathbb{N}^n \) than all of its nontrivial permutations\(^2\), which means
\[
d_1 \geq d_2 \geq \cdots \geq d_n \geq 0.
\]
That shows the definition of \( a_1, \ldots, a_n \) in (3.2) has nonnegative values, so \( s_1^{a_1} \cdots s_n^{a_n} \) is a polynomial. Its multidegree is the same as that of \( f \) by (3.2). Moreover, by Theorem 2.10,
\[
\text{lead}(c_d s_1^{a_1} \cdots s_n^{a_n}) = (\text{lead } c_d)(\text{lead } s_1)^{a_1} \cdots (\text{lead } s_n)^{a_n} = c_d.
\]
So \( f \) and \( c_d s_1^{a_1} \cdots s_n^{a_n} \) have the same leading term \( c_d X_1^{d_1} \cdots X_n^{d_n} \). Set \( f_1 := f - c_d s_1^{a_1} \cdots s_n^{a_n} \), so \( f_1 \) is symmetric since it is a difference of symmetric polynomials and
\[
f = c_d s_1^{a_1} \cdots s_n^{a_n} + f_1.
\]
- If \( f_1 = 0 \) then \( f = c_d s_1^{a_1} \cdots s_n^{a_n} \) and we’re done.
- If \( f_1 \neq 0 \) then
  \[
  \text{mdeg}(f_1) = \text{mdeg}(f - c_d s_1^{a_1} \cdots s_n^{a_n}) < \text{mdeg}(f)
  \]
by Example 2.11 (the leading terms of \( f \) and \( c_d s_1^{a_1} \cdots s_n^{a_n} \) match).

Run through the same process with \( f_1 \) in place of \( f \): let \( f_1 \) have leading term \( c_d, X^{d_1} \). Since \( f_1 \) is symmetric, some product \( s_1^{b_1} \cdots s_n^{b_n} \) has leading term \( X^{d_1} \), so \( f_1 \) has the same leading term as \( c_d, s_1^{b_1} \cdots s_n^{b_n} \). Set \( f_2 := f_1 - c_d, s_1^{b_1} \cdots s_n^{b_n} \), which is symmetric since it is a difference of symmetric polynomials and
\[
f = c_d s_1^{a_1} \cdots s_n^{a_n} + f_1 = c_d s_1^{a_1} \cdots s_n^{a_n} + c_d, s_1^{b_1} \cdots s_n^{b_n} + f_2
\]
where \( d_1 < d \).
- If \( f_2 = 0 \) then
  \[
  f = c_d s_1^{a_1} \cdots s_n^{a_n} + c_d, s_1^{b_1} \cdots s_n^{b_n}
  \]
  and we’re done.
- If \( f_2 \neq 0 \) then \( \text{mdeg}(f_2) < \text{mdeg}(f_1) \) by Example 2.11.

In this way, we obtain a sequence of formulas
\[
f = g_k(s_1, \ldots, s_n) + r_k
\]
where \( g_k(s_1, \ldots, s_n) \) is a polynomial in \( s_1, \ldots, s_n \) and \( r_k \) is a polynomial with \( \text{mdeg}(r_{k+1}) < \text{mdeg}(r_k) \) if \( r_k \) and \( r_{k+1} \) are not 0.

If \( r_k \) is never 0 then we get an infinite decreasing sequence
\[
\text{mdeg}(r_1) > \text{mdeg}(r_2) > \text{mdeg}(r_3) \cdots
\]
in \( \mathbb{N}^n \). This is impossible by Theorem 2.12, so some \( r_k \) is 0. Thus \( f = g_k(s_1, \ldots, s_n) \), which shows \( f \) is a polynomial in \( s_1, \ldots, s_n \). \( \square \)

Let’s summarize the recursive step: if \( f \) is symmetric in \( X_1, \ldots, X_n \) that is not 0 and its leading term is \( c_d, X_1^{d_1} \cdots X_n^{d_n} \) then either \( f = c_d, X_1^{d_1-1} \cdots X_n^{d_n} \) or
\[
\text{mdeg}(f - c_d, s_1^{d_1} \cdots s_n^{d_n}) < \text{mdeg}(f).
\]
By repeatedly subtracting off appropriate scalar multiples of products of powers of \( s_1, \ldots, s_n \) we get a sequence of symmetric polynomials with decreasing multidegree if they are never 0, so they must at some point be 0 because there is no infinite decreasing sequence in \( \mathbb{N}^n \).

\(^2\)A trivial permutation is one that exchanges equal coordinates, like \((2, 2, 1)\) and \((2, 2, 1)\).
Example 3.1. In three variables, let \( f(X, Y, Z) = X^4 + Y^4 + Z^4 \). To write \( f \) as a polynomial in the elementary symmetric polynomials in \( X, Y, \) and \( Z \), which are
\[
s_1 = X + Y + Z, \quad s_2 = XY + XZ + YZ, \quad s_3 = XYS,
\]
use lexicographic ordering with \( X > Y > Z \) (that is, \( X = X_1, Y = X_2, \) and \( Z = X_3 \)). The multidegree of \( s_1 s_2 s_3 \) is \((a + b + c, b + c, c)\).

Step 1. The leading term of \( f \) is \( X^4 \), with multidegree \((4, 0, 0)\). This is the multidegree of \( s_1^4 \), which has leading term \( X^4 \), so set \( f_1 = f - s_1^4 \). By a calculation,
\[
f_1 = -4X^3Y - 4XY^3 - 4XZ^3 - 4YZ^3 - 4Y^3Z - 6X^2Y^2 - 6X^2Z^2 - 6Y^2Z^2 - 12X^2YZ - 12XY^2Z - 12XYZ^2.
\]

Step 2. The symmetric polynomial \( f_1 \) has leading term \(-4X^3Y\), with multidegree \((3, 1, 0)\). This is \((a + b + c, b + c, c)\) when \( c = 0, \ b = 1, \ a = 2 \), so \( f_1 \) has the same leading term as \(-4s_1^3 s_2 s_3 = -4s_1^3 s_2\). Set \( f_2 = f_1 + 4s_1^3 s_2\):
\[
f_2 = 2X^2Y^2 + 2X^2Z^2 + 2Y^2Z^2 + 8X^2YZ + 8XY^2Z + 8XYZ^2.
\]

Step 3. The symmetric polynomial \( f_2 \) has leading term \( 2X^2Y^2 \) with multidegree \((2, 2, 0)\). This is \((a + b + c, b + c, c)\) when \( c = 0, \ b = 2, \ a = 0 \), so \( f_2 \) has the same leading term as \( 2s_2^2 \). Set \( f_3 = f_2 - 2s_2^2\):
\[
f_3 = 4X^2YZ + 4XY^2Z + 4XYZ^2.
\]

Step 4. The symmetric polynomial \( f_3 \) has leading term \( 4X^2YZ \), which has multidegree \((2, 1, 1)\). This is \((a + b + c, b + c, c)\) for \( c = 1, \ b = 0, \ a = 1 \), so \( f_3 \) has the same leading term as \( 4s_1 s_3 \). Set \( f_4 = f_3 - 4s_1 s_3 \) and this vanishes:
\[
f_4 = 0.
\]

Here is the decreasing sequence of multidegrees we obtained in successive steps before the process terminated:
\[
(3.4) \quad (4,0,0) > (3,1,0) > (2,2,0) > (2,1,1).
\]

Putting everything back together, we get \( X^4 + Y^4 + Z^4 \) as a polynomial in \( s_1, s_2, s_3 \):
\[
X^4 + Y^4 + Z^4 = f
\]
\[
= (f - f_1) + (f_1 - f_2) + (f_2 - f_3) + f_3
\]
\[
= s_1^4 - 4s_1^3 s_2 + 2s_2^2 + 4s_1 s_3.
\]

Remark 3.2. The finiteness that made the procedure in the proof of Theorem 1.5 terminate is the finiteness of strictly decreasing sequences in \( \mathbb{N}^n \) (Theorem 2.12). We did not write the proof of Theorem 1.5 using induction on the multidegree of symmetric polynomials, since \( \mathbb{N}^n \) in the lexicographic ordering does not have finitely many multidegrees below a given multidegree when \( n \geq 2 \) (see (2.1), where \((0, b) < (1, 0)\) for all \( b \geq 0 \)). We can describe the proof of Theorem 1.5 using induction if we think carefully about what makes the idea of induction work. Suppose for every \( d \) in \( \mathbb{N}^n \) we have a proposition \( P(d) \) (like “every symmetric polynomial in \( F[X_1, \ldots, X_n] \) of multidegree \( d \) is in \( F[s_1, \ldots, s_n] \))” and
- \( P(0) \) is true (Base Case),
- if \( d \neq 0 \) and \( P(n) \) is true for all \( n < d \) then \( P(d) \) is true (Inductive Step).
It follows that \( P(d) \) is true for all \( d \) in \( \mathbb{N}^n \). To prove that, let \( S = \{ d \in \mathbb{N}^n : P(d) \text{ is false} \} \). We want \( S \) to be empty: if there is no \( d \) in \( \mathbb{N}^n \) such that \( P(d) \) is false, then \( P(d) \) is true for all \( d \). To show \( S = \emptyset \) when the two conditions above hold, assume \( S \neq \emptyset \). Then \( S \) has a least element by Theorem 2.12, say \( d \). By the first condition above, \( d \neq 0 \). When \( n < d \), we have \( n \notin S \) since \( d \) is the least element of \( S \), so \( P(n) \) is true for all \( n < d \). Therefore \( P(d) \) is true by the second condition above. That contradicts the meaning of \( d \) being in \( S \), so \( S = \emptyset \).

Let’s see how to think about the proof of Theorem 1.5 in terms of induction on \( \mathbb{N}^n \). We started the proof by noting the theorem is true for nonzero constant polynomials, which are those of multidegree \( 0 \). Next, when \( f \) is symmetric with multidegree \( d > 0 \), we found a symmetric polynomial of the form \( c_d s_1^{a_1} \cdots s_n^{a_n} \) such that that either

- \( f = c_d s_1^{a_1} \cdots s_n^{a_n} \) or
- \( f \neq c_d s_1^{a_1} \cdots s_n^{a_n} \) and \( \text{mdeg}(f - c_d s_1^{a_1} \cdots s_n^{a_n}) < d \).

In the first case, \( f \in F[s_1, \ldots, s_n] \). In the second case, \( f - c_d s_1^{a_1} \cdots s_n^{a_n} \) is nonzero and symmetric with multidegree less than \( d \), so if we assume all nonzero symmetric polynomials in \( F[X_1, \ldots, X_n] \) of multidegree less than \( d \) are in \( F[s_1, \ldots, s_n] \), then \( f - c_d s_1^{a_1} \cdots s_n^{a_n} \) is in \( F[s_1, \ldots, s_n] \), so \( f \in F[s_1, \ldots, s_n] \). Do you see how this fits the format of induction on \( \mathbb{N}^n \)? For each \( d \in \mathbb{N}^n \) let \( P(d) \) be the proposition “every symmetric polynomial in \( F[X_1, \ldots, X_n] \) of multidegree \( d \) is in \( F[s_1, \ldots, s_n] \)”. Then we showed \( P(0) \) is true and we showed that if \( d \neq 0 \) and \( P(n) \) is true for all \( n < d \) then \( P(d) \) is true. Thus \( P(d) \) is true for all \( d \) by induction on \( \mathbb{N}^n \), so Theorem 1.5 is proved for all nonzero symmetric polynomials in \( F[X_1, \ldots, X_n] \), and it is obvious for the polynomial \( 0 \), so the theorem is proved for all symmetric polynomials.

Another way to prove Theorem 1.5 by induction that might feel less abstract is to modify the method of ordering \( \mathbb{N}^n \): we could compare \( n \)-tuples first by the sum of their coordinates, and in case of a tie break the tie by lexicographic order. This is called graded lexicographic order. For example, \((2, 3) \succ (1, 7) \succ (0, 5)\) in lexicographic order but \((1, 7) \succ (2, 3) \succ (0, 5)\) in graded lexicographic order. The ordering in (3.4) is valid for graded lexicographic order since all the triples there have the same sum of coordinates (all equal 4), so the comparison is made by lexicographic order. On nonzero monomials, graded lexicographic order amounts to comparing first by total degree (sum of exponents) and using lexicographic order only in case of equal total degree. When \( \mathbb{N}^n \) is arranged in graded lexicographic order, there are only finitely many \( n \)-tuples below a given \( n \)-tuple since that is already the case when we order \( \mathbb{N}^n \) just by the sum of the coordinates. This makes graded lexicographic order closer to your intuition from the standard total ordering on \( \mathbb{N} \). Theorem 1.5 is proved by induction using graded lexicographic order on monomials in [3, Theorem 2.2.2].

**Corollary 3.3.** Let \( L/K \) be a field extension and \( f(T) \in K[T] \) factor as

\[
(T - \alpha_1)(T - \alpha_2) \cdots (T - \alpha_n)
\]

in \( L[T] \). For each positive integer \( r \),

\[
(T - \alpha_1^r)(T - \alpha_2^r) \cdots (T - \alpha_n^r) \in K[T].
\]

\(^3\)In the case of \( \mathbb{N} \), the argument we just gave corresponds to the proof that the Well-Ordering Principle on \( \mathbb{N} \) implies the principle of mathematical induction on \( \mathbb{N} \). Theorem 2.12 says \( \mathbb{N}^n \) with the lexicographic ordering is a well-ordered set even though it has infinitely many elements below some elements if \( n \geq 2 \).
Proof. For indeterminates $X_1, \ldots, X_n$, write

\[(3.6) \quad (T - X_1^r)(T - X_2^r) \cdots (T - X_n^r) = \sum_{k=0}^{n} c_k(X_1, \ldots, X_n) T^k.\]

We can view each $c_k(X_1, \ldots, X_n)$ in $K[X_1, \ldots, X_n]$ (its coefficients are all 0 and 1) and it is symmetric in $X_1, \ldots, X_n$ since the left side of (3.6) is symmetric in $X_1, \ldots, X_n$. Therefore $c_k(X_1, \ldots, X_n)$ is a polynomial over $K$ in the elementary symmetric functions of $X_1, \ldots, X_n$, say

\[c_k(X_1, \ldots, X_n) = g_k(s_1, \ldots, s_n)\]

where $g_k$ is a polynomial in $n$ indeterminates over $K$. Then

\[(T - X_1^r)(T - X_2^r) \cdots (T - X_n^r) = \sum_{k=0}^{n} g_k(s_1, \ldots, s_n) T^k.\]

On both sides of this equation, specialize $X_i$ to $\alpha_i$. (Each $s_i$ is implicitly a polynomial $X_1, \ldots, X_n$ and thus also gets specialized.) The elementary symmetric functions of $\alpha_1, \ldots, \alpha_n$ are the non-leading coefficients of $f(T)$ (up to a sign), so they lie in $K$ since the coefficients of $f(T)$ are in $K$. Therefore a polynomial in the elementary symmetric functions of the $\alpha_i$’s with coefficients in $K$ lies in $K$, so each $g_k(s_1, \ldots, s_n)$ is specialized to a value in $K$. \hfill \Box

Remark 3.4. The same reasoning shows for each $h(X) \in K[X]$, not just $h(X) = X^r$, that

\[(T - h(\alpha_1))(T - h(\alpha_2)) \cdots (T - h(\alpha_n)) \in K[T].\]

The coefficient of $T^{n-1}$ in (3.6) is $-(X_1^r + \cdots + X_n^r)$. We call $X_1^r + \cdots + X_n^r$ the $r$th power sum of $X_1, \ldots, X_n$. It can be written as a polynomial in the elementary symmetric functions $s_1, \ldots, s_n$. The next example presents some of these polynomials for $n = 2$ and 3.

Example 3.5. For the second, third, and fourth powers sums in two variables,

\[
\begin{align*}
X^2 + Y^2 &= (X + Y)^2 - 2XY = s_1^2 - 2s_2, \\
X^3 + Y^3 &= (X + Y)^3 - 3(X + Y)XY = s_1^3 - 3s_1 s_2 \\
X^4 + Y^4 &= (X + Y)^4 - 4(X^2 + Y^2)(XY) - 6X^2 Y^2 = s_1^4 - 4s_1^2 s_2 + 2s_2^2,
\end{align*}
\]

where we use the formula $X^2 + Y^2 = s_1^2 - 2s_2$ in the last calculation.

Example 3.6. For the second, third, and fourth powers sums in three variables,

\[
\begin{align*}
X^2 + Y^2 + Z^2 &= (X + Y + Z)^2 - 2(XY + XZ + YZ) = s_1^2 - 2s_2, \\
X^3 + Y^3 + Z^3 &= (X + Y + Z)^3 - 3(XY + XZ + YZ)(X + Y + Z) + 3XYZ \\
&= s_1^3 - 3s_1 s_2 + 3s_3 \\
X^4 + Y^4 + Z^4 &= s_1^4 - 4s_1^2 s_2 + 2s_2^2 + 4s_1 s_3 \text{ by (3.5).}
\end{align*}
\]

Example 3.7. Let $f(T) = T^2 + 5T + 2 = (T - \alpha)(T - \beta)$ where $\alpha = (-5 + \sqrt{17})/2$ and $\beta = (-5 - \sqrt{17})/2$. Although $\alpha$ and $\beta$ are irrational, their elementary symmetric functions
are integers: \( s_1 = \alpha + \beta = -5 \) and \( s_2 = \alpha\beta = 2 \). Using the formulas in Example 3.5,
\[
\begin{align*}
\alpha^2 + \beta^2 &= p_2 = s_1^2 - 2s_2 = 21, \\
\alpha^3 + \beta^3 &= p_3 = s_1^3 - 3s_1s_2 = -95, \\
\alpha^4 + \beta^4 &= p_4 = s_1^4 - 4s_1^2s_2 + 2s_2^2 = 433.
\end{align*}
\]

Numerically, \( \alpha \approx -4.383 \) and \( \beta \approx -4.561 \), so \( p_r = \alpha^r + \beta^r \) is extremely close to \( \beta^r \) since \( |\alpha^r| \to 0 \) rapidly as \( r \) grows. In fact, \( p_r \) turns out to be the nearest integer to \( \beta^r \): compare \( \beta^2 \approx 20.807 \), \( \beta^3 \approx -94.915 \), and \( \beta^4 \approx 432.963 \). The way we computed the power sums above needed no approximations or formulas for \( \alpha \) and \( \beta \). The power sums are determined by the elementary symmetric functions \( s_1 \) and \( s_2 \) using exact calculations with integers.

**Example 3.8.** Let \( \alpha, \beta, \) and \( \gamma \) be the three roots of \( T^3 - T - 1 \), so
\[
T^3 - T - 1 = (T - \alpha)(T - \beta)(T - \gamma).
\]
The elementary symmetric functions of \( \alpha, \beta, \) and \( \gamma \) are all \( s_1 = 0 \), \( s_2 = -1 \), and \( s_3 = 1 \), so by the formulas in Example 3.6,
\[
\begin{align*}
\alpha^2 + \beta^2 + \gamma^2 &= p_2 = s_1^2 - 2s_2 = 2, \\
\alpha^3 + \beta^3 + \gamma^3 &= p_3 = s_1^3 - 3s_1s_2 + 3s_3 = 3, \\
\alpha^4 + \beta^4 + \gamma^4 &= p_4 = s_1^4 - 4s_1^2s_2 + 2s_2^2 + 4s_1s_3 = 2.
\end{align*}
\]
The next few power sums are \( p_5 = 5 \), \( p_6 = 5 \), \( p_7 = 7 \), and \( p_8 = 10 \). We do not need formulas for \( \alpha, \beta, \) and \( \gamma \) to make these power sum calculations. All of them are determined by the elementary symmetric function values \( s_1, s_2, \) and \( s_3 \), which come from the coefficients of \( T^3 - T - 1 \).

Comparing the formulas in Examples 3.5 and 3.6 for the same exponent, they match in degree 2 but look different in degrees 3 and 4. However, they are the same formula in degree 3 by taking \( s_3 = 0 \) for two variables and they are the same formula in degree 4 by also taking \( s_4 = 0 \) for two variables. If we let \( s_k(X_1, \ldots, X_n) = 0 \) when \( k > n \) then there are “universal” formulas for power sums \( p_r = X_1^r + \cdots + X_n^r \) in terms of elementary symmetric polynomials in any number of variables, starting out as
\[
\begin{align*}
p_1 &= s_1, \\
p_2 &= s_1^2 - 2s_2, \\
p_3 &= s_1^3 - 3s_1s_2 + 3s_3, \\
p_4 &= s_1^4 - 4s_1^2s_2 + 2s_2^2 + 4s_1s_3 - 4s_4.
\end{align*}
\]

There is some apparent regularity in these formulas: the first two terms in the formula for \( p_r \) appear to be \( s_1^r - rs_1^{r-2}s_2 \). A nice regularity for all \( r \) can be seen in determinant formulas:
\[
\begin{align*}
p_2 &= \det \begin{pmatrix} s_1 & 1 \\ 2s_2 & s_1 \end{pmatrix}, \\
p_3 &= \det \begin{pmatrix} s_1 & 1 & 0 \\ 2s_2 & s_1 & 1 \\ 3s_3 & s_2 & s_1 \end{pmatrix}, \\
p_4 &= \det \begin{pmatrix} s_1 & 1 & 0 & 0 \\ 2s_2 & s_1 & 1 & 0 \\ 3s_3 & s_2 & s_1 & 1 \\ 4s_4 & s_3 & s_2 & s_1 \end{pmatrix}.
\end{align*}
\]
4. Uniqueness for Theorem 1.5

The proof of Theorem 1.5 leads to a specific way of writing a symmetric polynomial \( f \) in \( X_1, \ldots, X_n \) as a polynomial in \( s_1, \ldots, s_n \), but that does not automatically mean there can only be one such representation of \( f \) as a polynomial in \( s_1, \ldots, s_n \); just because an algorithm has a well-determined final result does not mean the problem it is solving only has one answer. For instance, if we apply Euclid’s algorithm to solve \( 18x + 5y = 1 \) in integers, we get the specific solution \( (x, y) = (2, -7) \), but the equation \( 18x + 5y = 1 \) has infinitely many integral solutions: \( (x, y) = (2 + 5t, -7 - 18t) \) for all \( t \in \mathbb{Z} \). It turns out that symmetric polynomials in \( X_1, \ldots, X_n \) do have a unique representation as a polynomial in \( s_1, \ldots, s_n \).

**Theorem 4.1.** Every symmetric polynomial in \( F[X_1, \ldots, X_n] \) can be written in only one way as a polynomial in the elementary symmetric functions of \( X_1, \ldots, X_n \).

**Proof.** Let \( s_1, \ldots, s_n \) be the elementary symmetric functions of \( X_1, \ldots, X_n \). Each polynomial in \( F[X_1, \ldots, X_n] \) that is symmetric in the \( X_i \)’s is in \( F[s_1, \ldots, s_n] \) by Theorem 1.5.

To prove uniqueness of this representation, suppose \( g(Y_1, \ldots, Y_n) \) and \( h(Y_1, \ldots, Y_n) \) in \( F[Y_1, \ldots, Y_n] \) satisfy \( g(s_1, \ldots, s_n) = h(s_1, \ldots, s_n) \) in \( F[X_1, \ldots, X_n] \). We want to show \( g = h \) in \( F[Y_1, \ldots, Y_n] \). For each monomial in \( Y_1, \ldots, Y_n \), \( g \) has the same coefficients in \( g \) and \( h \).

By passing to the difference \( g(Y_1, \ldots, Y_n) - h(Y_1, \ldots, Y_n) \), we are reduced to showing that

\[
g(s_1, \ldots, s_n) = 0 \implies g(Y_1, \ldots, Y_n) = 0.
\]

We will prove the contrapositive: if \( g(Y_1, \ldots, Y_n) \neq 0 \) in \( F[Y_1, \ldots, Y_n] \) then \( g(s_1, \ldots, s_n) \neq 0 \) in \( F[X_1, \ldots, X_n] \).

Write the nonzero \( g(Y_1, \ldots, Y_n) \) as a sum of finitely many monomials in the \( Y_i \)’s:

\[
g(Y_1, \ldots, Y_n) = \sum_i c_i Y_i^{t_i}
\]

and some \( c_i \) is nonzero. Therefore

\[
g(s_1, \ldots, s_n) = \sum_i c_i s_1^{t_i} \cdots s_n^{t_n}.
\]

Different terms in this sum, as a polynomial in \( X_1, \ldots, X_n \), could share some monomials and thus lead to cancellation when like monomials are added together. To show \( g(s_1, \ldots, s_n) \neq 0 \), we want at least one monomial in the \( X_i \)'s not to cancel out in \( g(s_1, \ldots, s_n) \).

The subtle issue here is that the leading term of \( g(Y_1, \ldots, Y_n) \) need not contain the leading term of \( g(s_1, \ldots, s_n) \) as a polynomial in \( X_1, \ldots, X_n \). For example, suppose \( g(Y_1, Y_2) = Y_1^5 + Y_2^5 \), which has leading term \( Y_1^5 \). Replacing \( Y_1 \) with \( s_1 \),

\[
g(s_1, s_2) = s_1^5 + s_2^5 = (X_1 + X_2)^5 + (X_1X_2)^5,
\]

and the leading term is \( (X_1X_2)^5 \), which comes from the non-leading term \( Y_2^5 \) in \( g(Y_1, Y_2) \) after substituting \( s_1 \) for \( Y_i \). In the general case, the leading term of \( g(s_1, \ldots, s_n) \) as a polynomial in \( X_1, \ldots, X_n \) comes from somewhere in \( g(Y_1, \ldots, Y_n) \) but it’s not clear where.
We can compute the leading term of $s_1^{i_1} \cdots s_n^{i_n}$ as a polynomial in $X_1, \ldots, X_n$ by Example 2.9 and Theorem 2.10 because the multidegree tells us the exponents for the leading term:

$$mdeg(s_1^{i_1} \cdots s_n^{i_n}) = \sum_{k=1}^n i_k mdeg(s_k)$$

$$= i_1(1, 0, \ldots, 0) + i_2(1, 1, 0, \ldots, 0) + \cdots + i_n(1, 1, 1, \ldots, 1)$$

$$= (i_1 + i_2 + \cdots + i_n, i_2 + \cdots + i_n, \ldots, i_n - 1 + i_n, i_n).$$

(4.2)

For $i$ and $j$ in $\mathbb{N}^n$, comparing the formulas

$$mdeg(s_1^{i_1} \cdots s_n^{i_n}) = (i_1 + i_2 + \cdots + i_n, i_2 + \cdots + i_n, \ldots, i_n - 1 + i_n, i_n),$$

$$mdeg(s_1^{j_1} \cdots s_n^{j_n}) = (j_1 + j_2 + \cdots + j_n, j_2 + \cdots + j_n, \ldots, j_n - 1 + j_n, j_n)$$

from rightmost to the leftmost coordinates shows that if $mdeg(s_1^{i_1} \cdots s_n^{i_n}) = mdeg(s_1^{j_1} \cdots s_n^{j_n})$ then $i = j$. Therefore when the $n$-tuples $i$ and $j$ are different, the polynomials $s_1^{i_1} \cdots s_n^{i_n}$ and $s_1^{j_1} \cdots s_n^{j_n}$ have different multidegrees and leading coefficients 1 (each $s_i$ has leading coefficient 1). Therefore the terms in (4.1) with nonzero coefficient $c_i$ have different multidegrees, so one of them (we don’t know which!) is greatest. The last part of Theorem 2.10 extends to more than two polynomials:

$$mdeg(f_1) > mdeg(f_2), \ldots, mdeg(f_m) \implies mdeg(f_1 + f_2 + \cdots + f_m) = mdeg(f_1).$$

Apply this when $f_1, \ldots, f_m$ are the nonzero terms in (4.1): $mdeg(g(s_1, \ldots, s_n))$ is the largest multidegree of a nonzero term in (4.1). In particular, $g(s_1, \ldots, s_n) \neq 0$.

**Example 4.2.** The only expression of $X^4 + Y^4 + Z^4$ as a polynomial in $s_1, s_2,$ and $s_3$ is the one appearing in (3.5).

In the proofs of Theorems 1.5 and 4.1, the fact that the coefficients come from a field $F$ rather than a ring is not important since we never had to divide in $F$. The same proofs show for a nonzero commutative ring $R$ that each symmetric polynomial in $R[X_1, \ldots, X_n]$ lies in $R[s_1, \ldots, s_n]$ in exactly one way.4 For instance, Theorems 1.5 and 4.1 together using integer coefficients take the following form.

**Theorem 4.3.** The set of all symmetric polynomials in $\mathbb{Z}[X_1, \ldots, X_n]$ is $\mathbb{Z}[s_1, \ldots, s_n]$, and each element of $\mathbb{Z}[s_1, \ldots, s_n]$ is a polynomial in $s_1, \ldots, s_n$ in exactly one way.

**Example 4.4.** Taking $\alpha$ and $\beta$ as in Example 3.7, their elementary symmetric functions are both integers: $\alpha + \beta = -5$ and $\alpha \beta = 2$. If $f(X, Y)$ is a symmetric polynomial with integer coefficients then $f(X, Y) = g(X + Y, XY)$ where $g$ is a polynomial with integer coefficients. Therefore $f(\alpha, \beta) = g(-5, 2) \in \mathbb{Z}$. This implies $(T - \alpha^r)(T - \beta^r)$, whose coefficients are $\alpha^r + \beta^r$ and $\alpha^r \beta^r$, has integral coefficients and not just rational coefficients. That is why the calculations with $r = 2, 3,$ and 4 in Example 3.7 have coefficients in $\mathbb{Z}$.

**Remark 4.5.** Theorem 4.3 is used in the proof that $\pi$ is transcendental [1, Section 5.2], [6, Chapter 3].

---

4 There is a slight technicality to be aware of: if $R$ is not a domain then the formula $mdeg(fg) = mdeg f + mdeg g$ is true only as long as the leading coefficients of $f$ and $g$ are both not zero-divisors in $R$, and that is true for the case of elementary symmetric polynomials $s_1, \ldots, s_n$ since their leading coefficients equal 1.
5. History

Lexicographic ordering on multivariable polynomials and its application to the proof of Theorem 1.5 go back to Gauss [4, pp. 36–37], who used Theorem 1.5 in his second proof of the Fundamental Theorem of Algebra. His description of lexicographic ordering is shown in Figure 1, where the Latin text says “from the two terms $Ma^\alpha b^\beta c^\gamma \cdots$ and $Ma^{\alpha'} b^{\beta'} c^{\gamma'} \cdots$, call the first one of higher order than the second if $\alpha > \alpha'$ or if $\alpha = \alpha'$ and $\beta > \beta'$, or if $\alpha = \alpha'$, $\beta = \beta'$, and $\gamma > \gamma'$, etc.”.

![Figure 1. Gauss’s definition of lexicographic ordering.](image)

The lexicographic ordering is just one example of a total ordering on monomials in multivariable polynomial rings. These orderings are part of computer algebra systems that are widely used in computational commutative algebra and algebraic geometry.

References