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 σ 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 $f(X_1, X_2, X_3) = X_1^5 + X_2X_3$. This polynomial is unchanged if we interchange X_2 and X_3, but if we interchange X_1 and X_3 then f becomes $X_3^5 + X_2X_1$, which is not f. This polynomial is only "partially symmetric."

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

$$(T - X_1)(T - X_2) \cdots (T - X_n) = T^n - s_1T^{n-1} + s_2T^{n-2} - \cdots + (-1)^n s_n. \tag{1.1}$$

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 kth *elementary* symmetric polynomial – or kth 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 α and β are not rational, their elementary symmetric polynomials are: $s_1 = \alpha + \beta = 3$ and $s_2 = \alpha\beta = 1$.

Example 1.4. Let α, β, and γ 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 α, β, and γ 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 [4, 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.[1] The proof we will give, based on [1, 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 $\mathbf{i} = (i_1, \ldots, i_n)$ and $\mathbf{j} = (j_1, \ldots, j_n)$ in $\mathbf{N}^n$, set $\mathbf{i} < \mathbf{j}$ if, for the first index r such that $i_r \neq j_r$, we have $i_r < j_r$. Write $\mathbf{i} \leq \mathbf{j}$ if $\mathbf{i} < \mathbf{j}$ or $\mathbf{i} = \mathbf{j}$.

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

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

This way of ordering n-tuples in $\mathbf{N}^n$ 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 $\mathbf{N}$ with the same number of terms.

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

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

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

(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 $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are all equal, either $i_r < j_r$ or $j_r < k_r$ because $\mathbf{i} < \mathbf{j}$ and $\mathbf{j} < \mathbf{k}$. 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 $\mathbf{i}$ and $\mathbf{k}$ are equal. Thus $\mathbf{i} < \mathbf{k}$.

[1] 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 $\mathbf{i}$ and $\mathbf{j}$ may coincide or $\mathbf{j}$ and $\mathbf{k}$ may coincide, which makes things easier.

(3) Rather than take cases based on where $\mathbf{i}$ and $\mathbf{i}'$ may first differ or where $\mathbf{j}$ and $\mathbf{j}'$ may first differ, observe that $\mathbf{i} \leq \mathbf{i}' \Rightarrow \mathbf{i}+\mathbf{k} \leq \mathbf{i}'+\mathbf{k}$ for all $\mathbf{k}$: this is obvious when $\mathbf{i} = \mathbf{i}'$, and when $\mathbf{i} < \mathbf{i}'$ 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 $\mathbf{i} \leq \mathbf{i}'$ and $\mathbf{j} \leq \mathbf{j}'$, then

$$\mathbf{i}+\mathbf{j} \leq \mathbf{i}'+\mathbf{j} = \mathbf{j}+\mathbf{i}' \leq \mathbf{j}'+\mathbf{i}' = \mathbf{i}'+\mathbf{j}' \Longrightarrow \mathbf{i}+\mathbf{j} \leq \mathbf{i}'+\mathbf{j}'.$$

The case of $<$ in place of $\leq$ is analogous. □

A polynomial $f \in F[X_1, \dots, X_d]$ is a sum of the form

$$f = \sum_{i_1,\dots,i_n \geq 0} c_{i_1,\dots,i_n} X_1^{i_1} \cdots X_n^{i_n}$$

where $c_{i_1,\dots,i_n} \in F$ and only finitely many coefficients can be nonzero. Abbreviate this sum to multi-index form as $\sum_{\mathbf{i}} c_{\mathbf{i}} \mathbf{X}^{\mathbf{i}}$, where $\mathbf{X}^{\mathbf{i}} := X_1^{i_1} \cdots X_n^{i_n}$ for $\mathbf{i} = (i_1, \dots, i_n)$. Note $\mathbf{X}^{\mathbf{i}}\mathbf{X}^{\mathbf{j}} = \mathbf{X}^{\mathbf{i}+\mathbf{j}}$. (In the notation $\sum_{\mathbf{i}}$, 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 $F[X_1, \dots, X_n]$ as $\sum_{\mathbf{i}} c_{\mathbf{i}} \mathbf{X}^{\mathbf{i}}$. If $f \neq 0$, the *multidegree* of f is the lexicographically largest index in $\mathbf{N}^n$ of a nonzero monomial in f:

$$\operatorname{mdeg} f = \max\{\mathbf{i} : c_{\mathbf{i}} \neq \mathbf{0}\} \in \mathbf{N}^n.$$

The multidegree of the zero polynomial is not defined. If $f \neq 0$ and $\operatorname{mdeg} f = \mathbf{d}$, we call $c_{\mathbf{d}}\mathbf{X}^{\mathbf{d}}$ the *leading term* of f and $c_{\mathbf{d}}$ the *leading coefficient* of f, written $c_{\mathbf{d}} = \operatorname{lead} f$.

We can separate the leading term of a nonzero f of multidegree $\mathbf{d}$ from its other nonzero terms to get

$$f = c_{\mathbf{d}}\mathbf{X}^{\mathbf{d}} + \sum_{\mathbf{i}<\mathbf{d}} c_{\mathbf{i}} \mathbf{X}^{\mathbf{i}}.$$

Because multidegrees are totally ordered, a nonzero polynomial in $F[X_1, \dots, X_n]$ has a unique leading term and a unique leading coefficient in F.

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

Example 2.7. In $F[X_1, \dots, X_n]$, $\operatorname{mdeg}(X_1) = (1, 0, \dots, 0)$ and $\operatorname{mdeg}(X_n) = (0, 0, \dots, 1)$.

Example 2.8. The polynomials with multidegree $\mathbf{0}$ are the nonzero constants.

Example 2.9. The multidegrees of the elementary symmetric polynomials are

$$\begin{aligned}
\operatorname{mdeg}(s_1) &= (1, 0, 0, \dots, 0), \\
\operatorname{mdeg}(s_2) &= (1, 1, 0, \dots, 0), \\
&\vdots \\
\operatorname{mdeg}(s_n) &= (1, 1, 1, \dots, 1).
\end{aligned}$$

For $k = 1, \dots, n$, 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 $F[X_1, \ldots, X_n]$, $\operatorname{mdeg}(fg) = \operatorname{mdeg}(f) + \operatorname{mdeg}(g)$ in $\mathbf{N}^n$ and $\operatorname{lead}(fg) = (\operatorname{lead} f)(\operatorname{lead} g)$.*

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

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

Let $\operatorname{mdeg} f = \mathbf{n}$ and $\operatorname{mdeg} g = \mathbf{m}$, say $f = c_{\mathbf{n}}\mathbf{X}^{\mathbf{n}} + \sum_{\mathbf{i}<\mathbf{n}} c_{\mathbf{i}}\mathbf{X}^{\mathbf{i}}$ with $c_{\mathbf{n}} \neq 0$ and $g = c'_{\mathbf{m}}\mathbf{X}^{\mathbf{b}} + \sum_{\mathbf{j}<\mathbf{m}} c'_{\mathbf{j}}\mathbf{X}^{\mathbf{j}}$ 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 $c_{\mathbf{n}}c'_{\mathbf{m}}\mathbf{X}^{\mathbf{n}+\mathbf{m}}$ and every other term has multidegree $\mathbf{n}+\mathbf{j}$, $\mathbf{m}+\mathbf{i}$, or $\mathbf{i}+\mathbf{j}$ where $\mathbf{i} < \mathbf{n}$ and $\mathbf{j} < \mathbf{m}$. By Theorem 2.4, all these other multidegrees are less than $\mathbf{n} + \mathbf{m}$, so $\operatorname{mdeg}(fg) = \mathbf{n} + \mathbf{m} = \operatorname{mdeg} f + \operatorname{mdeg} g$ and $\operatorname{lead}(fg) = c_{\mathbf{n}}c'_{\mathbf{m}} = (\operatorname{lead} f)(\operatorname{lead} g)$. ☐

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 $\operatorname{mdeg}(f - g) < \operatorname{mdeg} f$. Writing the common leading term as $c_{\mathbf{n}}\mathbf{X}^{\mathbf{n}}$,

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

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

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

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

In $\mathbf{N}$ with its usual ordering, there are finitely many elements below a given element, but this is not true in $\mathbf{N}^n$ 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 $b \in \mathbf{N}$. In terms of lexicographic ordering on $F[X, Y]$, where $X > Y$, this corresponds to saying

$$1 < Y < Y^2 < Y^3 < \cdots < Y^b < \cdots < X. \tag{2.1}$$

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

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

The least element in the subset has to be unique since lexicographic ordering is a total ordering on $\mathbf{N}^n$.

Proof. Let S be a nonempty subset of $\mathbf{N}^n$. The first coordinates of elements in S are a nonempty subset of $\mathbf{N}$ and thus have a least element ℓ_1. If $n = 1$ then ℓ_1 is the least element of S. For $n \geq 2$, among the elements of S with first coordinate ℓ_1, their second coordinates are a nonempty subset of $\mathbf{N}$ and thus have a least element ℓ_2. If $n = 2$ then (ℓ_1, ℓ_2) is the least element of S. For $n \geq 3$, among the elements of S with first coordinate

ℓ_1 and second coordinate ℓ_2, their third coordinates are a nonempty subset of $\mathbf{N}$ and thus have a least element ℓ_3. Continue this way up through the nth coordinate. The n-tuple $(\ell_1, \ell_2, \ldots, \ell_n)$ in S is the least element of S in the lexicographic ordering on $\mathbf{N}^n$.

It is left to the reader to rewrite this argument as a proper proof by induction on n.

If we have an infinite strictly decreasing sequence $\mathbf{d}_1 > \mathbf{d}_2 > \mathbf{d}_3 > \ldots$ in $\mathbf{N}^n$ then it has no least element, so a strictly decreasing sequence in $\mathbf{N}^n$ must have only finitely many terms. □

3. Proof of Theorem 1.5

Here is a proof of Theorem 1.5 that uses lexicographic ordering on $F[X_1, \ldots, X_n]$ and Theorem 2.12 to show a recursive process must terminate in finitely many steps.

Proof. We want to show every symmetric polynomial in $F[X_1, \ldots, X_n]$ is a polynomial in $F[s_1, \ldots, s_n]$. This is obvious for constant polynomials, which are 0 and polynomials of multidegree $(0, 0, \ldots, 0)$.

Let f be a nonconstant symmetric polynomial in $F[X_1, \ldots, X_n]$ with multidegree $\mathbf{d} = (d_1, \ldots, d_n)$. Pull out the leading term of f:

$$f = c_{\mathbf{d}}\mathbf{X}^{\mathbf{d}} + \sum_{\mathbf{i}<\mathbf{d}} c_{\mathbf{i}}\mathbf{X}^{\mathbf{i}} = c_{\mathbf{d}} X_1^{d_1} \cdots X_n^{d_n} + \sum_{\mathbf{i}<\mathbf{d}} c_{\mathbf{i}}\mathbf{X}^{\mathbf{i}}, \tag{3.1}$$

where $c_{\mathbf{d}} \neq 0$. We will find a polynomial in $s_1, \ldots, s_n$ with the same leading term as f. Its difference with f will be symmetric with smaller multidegree than $\mathbf{d}$.

By Example 2.9 and Theorem 2.10, for nonnegative integers $a_1, \ldots, a_n$,

$$\begin{aligned} \operatorname{mdeg}(s_1^{a_1} s_2^{a_2} \cdots s_n^{a_n}) &= a_1 \operatorname{mdeg}(s_1) + a_2 \operatorname{mdeg}(s_2) + \cdots + a_n \operatorname{mdeg}(s_n) \\ &= (a_1 + a_2 + \cdots + a_n, a_2 + \cdots + a_n, \ldots, a_n). \end{aligned}$$

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

$$a_1 = d_1 - d_2, \quad a_2 = d_2 - d_3, \quad \ldots, \quad a_{n-1} = d_{n-1} - d_n, \quad a_n = d_n. \tag{3.2}$$

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 $\mathbf{d} = \operatorname{mdeg}(f)$ satisfy

$$d_1 \geq d_2 \geq \cdots \geq d_n \geq 0. \tag{3.3}$$

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

To appreciate this issue, consider $f = X_1X_2^5 + 3X_2$. 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 $f = X_1X_2^5 + X_1^5X_2$ then f is symmetric and $\operatorname{mdeg} f = (5, 1)$ does satisfy (3.3). The verification of (3.3) will depend crucially on f being symmetric.

Since $(d_1, \ldots, d_n)$ 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 d_i's permuted is *also* a multidegree of a nonzero monomial in f. (This is how the symmetry of f in the X_i's is used: under all permutations of the X_i's, f stays unchanged.) For example, if f in $F[x, y]$ is symmetric and contains the monomial cx^3y^2 then f also contains the monomial cx^2y^3 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 $\mathbf{N}^n$ 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,

$$\operatorname{lead}(c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}) = (\operatorname{lead} c_{\mathbf{d}})(\operatorname{lead} s_1)^{a_1} \cdots (\operatorname{lead} s_n)^{a_n} = c_{\mathbf{d}}.$$

So f and $c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}$ have the same leading term $c_{\mathbf{d}} X_1^{d_1} \cdots X_n^{d_n}$. Set $f_1 := f - c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}$, so f_1 is symmetric since it is a difference of symmetric polynomials and

$$f = c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n} + f_1.$$

- If $f_1 = 0$ then $f = c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}$ and we're done.
- If $f_1 \neq 0$ then
$$\operatorname{mdeg}(f_1) = \operatorname{mdeg}(f - c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}) < \operatorname{mdeg}(f)$$
by Example 2.11 (the leading terms of f and $c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}$ match).

Run through the same process with f_1 in place of f: let f_1 have leading term $c_{\mathbf{d}_1} \mathbf{X}^{\mathbf{d}_1}$. Since f_1 is symmetric, some product $s_1^{b_1} \cdots s_n^{b_n}$ has leading term $\mathbf{X}^{\mathbf{d}_1}$, so f_1 has the same leading term as $c_{\mathbf{d}_1} s_1^{b_1} \cdots s_n^{b_n}$. Set $f_2 := f_1 - c_{\mathbf{d}_1} s_1^{b_1} \cdots s_n^{b_n}$, which is symmetric since it is a difference of symmetric polynomials and

$$f = c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n} + f_1 = c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n} + c_{\mathbf{d}_1} s_1^{b_1} \cdots s_n^{b_n} + f_2$$

where $\mathbf{d}_1 < \mathbf{d}$.

- If $f_2 = 0$ then
$$f = c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n} + c_{\mathbf{d}_1} s_1^{b_1} \cdots s_n^{b_n}$$
and we're done.
- If $f_2 \neq 0$ then $\operatorname{mdeg}(f_2) < \operatorname{mdeg}(f_1)$ 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 $\operatorname{mdeg}(r_{k+1}) < \operatorname{mdeg}(r_k)$ if r_k and r_{k+1} are not 0.

If r_k is never 0 then we get an infinite decreasing sequence

$$\operatorname{mdeg}(r_1) > \operatorname{mdeg}(r_2) > \operatorname{mdeg}(r_3) \cdots$$

in $\mathbf{N}^n$. 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 $c_{\mathbf{d}} X_1^{d_1} \cdots X_{n-1}^{d_{n-1}} X_n^{d_n}$ then either $f = c_{\mathbf{d}} s_1^{d_1 - d_2} \cdots s_{n-1}^{d_{n-1} - d_n} s_n^{d_n}$ or

$$\operatorname{mdeg}(f - c_{\mathbf{d}} s_1^{d_1 - d_2} \cdots s_{n-1}^{d_{n-1} - d_n} s_n^{d_n}) < \operatorname{mdeg}(f).$$

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 $\mathbf{N}^n$.

[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

$$s_1 = X+Y+Z, \quad s_2 = XY+XZ+YZ, \quad s_3 = XYZ,$$

use lexicographic ordering with $X > Y > Z$ (that is, $X = X_1$, $Y = X_2$, and $Z = X_3$). The multidegree of $s_1^a s_2^b s_3^c$ 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 $s_1^4 = (X+Y+Z)^4$, which has leading term X^4, so set $f_1 = f - s_1^4$. By a calculation,

$$\begin{aligned} f_1 = &-4X^3Y - 4XY^3 - 4X^3Z - 4XZ^3 - 4Y^3Z - 4YZ^3 - 6X^2Y^2 - 6X^2Z^2 - 6Y^2Z^2 \\ &- 12X^2YZ - 12XY^2Z - 12XYZ^2. \end{aligned}$$

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 $-4s_1^a s_2^b s_3^c = -4s_1^2 s_2$. Set $f_2 = f_1 + 4s_1^2 s_2$:

$$f_2 = 2X^2Y^2 + 2X^2Z^2 + 2Y^2Z^2 + 8X^2YZ + 8XY^2Z + 8XYZ^2.$$

Step 3. The symmetric polynomial f_2 has leading term is $2X^2Y^2$ 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 is $4X^2YZ$, which has multidegree $(2,1,1)$. This is $(a+b+c, b+c, c)$ for $c = 1$, $b = 0$, and $a = 1$, so f_3 has the same leading term as $4s_1s_3$. Set $f_4 = f_3 - 4s_1s_3$ and this vanishes:

$$f_4 = 0.$$

Here is the decreasing sequence of multidegrees we obtained in successive steps before the process terminated:

$$(4,0,0) > .(3,1,0) > (2,2,0) > (2,1,1). \tag{3.4}$$

Putting everything back together, we get $X^4+Y^4+Z^4$ as a polynomial in s_1, s_2, s_3:

$$\begin{aligned} X^4+Y^4+Z^4 &= f \\ &= (f - f_1) + (f_1 - f_2) + (f_2 - f_3) + f_3 \\ &= s_1^4 - 4s_1^2 s_2 + 2s_2^2 + 4s_1s_3. \end{aligned} \tag{3.5}$$

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 $\mathbf{N}^n$ (Theorem 2.12). We did not write the proof of Theorem 1.5 using induction on the multidegree of symmetric polynomials, since $\mathbf{N}^n$ in the lexicographic ordering does not have finitely many multidegrees below a given multidegree when $n \geq 2$ (see (2.1)). We can describe the proof using induction if we think carefully about what makes the idea of induction work. Suppose for every $\mathbf{d}$ in $\mathbf{N}^n$ we have a proposition $P(\mathbf{d})$ (like "every symmetric polynomial in $F[X_1,\ldots,X_n]$ of multidegree $\mathbf{d}$ is in $F[s_1,\ldots,s_n]$") and

- $P(\mathbf{0})$ is true (Base Case),
- if $\mathbf{d} \neq \mathbf{0}$ and $P(\mathbf{n})$ is true for all $\mathbf{n} < \mathbf{d}$ then $P(\mathbf{d})$ is true (Inductive Step).

It follows that $P(\mathbf{d})$ is true for all $\mathbf{d}$ in $\mathbf{N}^n$. To prove that, let $S = \{\mathbf{d} \in \mathbf{N}^n : P(\mathbf{d}) \text{ is false}\}$. We want S to be *empty*: if there is no $\mathbf{d}$ in $\mathbf{N}^n$ such that $P(\mathbf{d})$ is false, then $P(\mathbf{d})$ is true for all $\mathbf{d}$. 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 $\mathbf{d}$. By the first condition above, $\mathbf{d} \neq \mathbf{0}$. When $\mathbf{n} < \mathbf{d}$, we have $\mathbf{n} \notin S$ since $\mathbf{d}$ is the least element of S, so $P(\mathbf{n})$ is true for all $\mathbf{n} < \mathbf{d}$. Therefore $P(\mathbf{d})$ is true by the second condition above. That contradicts the meaning of $\mathbf{d}$ being in S, so $S = \emptyset$.[3]

We can now think about our proof of Theorem 1.5 in terms of induction on $\mathbf{N}^n$. We started the proof by noting the theorem is true for nonzero constant polynomials, which are those of multidegree $\mathbf{0}$. Next, when f is symmetric with multidegree $\mathbf{d} > \mathbf{0}$, we found a symmetric polynomial of the form $c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}$ such that that either

- $f = c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}$ or
- $f \neq c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}$ and $\mathrm{mdeg}(f - c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}) < \mathbf{d}$.

In the first case, $f \in F[s_1, \ldots, s_n]$. In the second case, $f - c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}$ is nonzero and symmetric with multidegree less than $\mathbf{d}$, so if we *assume* all nonzero symmetric polynomials in $F[X_1, \ldots, X_n]$ of multidegree less than $\mathbf{d}$ are in $F[s_1, \ldots, s_n]$, then $f - c_{\mathbf{d}} s_1^{a_1} \cdots s_n^{a_n}$ 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 $\mathbf{N}^n$? For each $\mathbf{d} \in \mathbf{N}^n$ let $P(\mathbf{d})$ be the proposition "every symmetric polynomial in $F[X_1, \ldots, X_n]$ of multidegree $\mathbf{d}$ is in $F[s_1, \ldots, s_n]$". Then we showed $P(\mathbf{0})$ is true and we showed that if $\mathbf{d} \neq \mathbf{0}$ and $P(\mathbf{n})$ is true for all $\mathbf{n} < \mathbf{d}$ then $P(\mathbf{d})$ is true. Thus $P(\mathbf{d})$ is true for all $\mathbf{d}$ by induction on $\mathbf{N}^n$, 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 $\mathbf{N}^n$: 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, $(2, 3) > (1, 7) > (0, 5)$ in lexicographic order but $(1, 7) > (2, 3) > (0, 5)$ 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 $\mathbf{N}^n$ 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 $\mathbf{N}^n$ just by the sum of the coordinates. This makes graded lexicographic order closer to your intuition from the standard total ordering on $\mathbf{N}$. Theorem 1.5 is proved by induction using graded lexicographic order on monomials in [2, 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 $\mathbf{N}$, the argument we just gave corresponds to the proof that the Well-Ordering Principle on $\mathbf{N}$ implies the principle of mathematical induction on $\mathbf{N}$. Theorem 2.12 says $\mathbf{N}^n$ 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 $X_1, \dots, X_n$, write

$$(T - X_1^r)(T - X_2^r) \cdots (T - X_n^r) = \sum_{k=0}^{n} c_k(X_1, \dots, X_n) T^k. \tag{3.6}$$

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

$$c_k(X_1, \dots, X_n) = g_k(s_1, \dots, s_n)$$

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

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

On both sides of this equation, specialize X_i to α_i. (Each s_i is implicitly a polynomial $X_1, \dots, X_n$ and thus also gets specialized.) The elementary symmetric functions of $\alpha_1, \dots, \alpha_n$ 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 α_i's with coefficients in K lies in K, so each $g_k(s_1, \dots, s_n)$ is specialized to a value in K. □

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 rth *power sum* of $X_1, \dots, X_n$. It can be written as a polynomial in the elementary symmetric functions $s_1, \dots, s_n$. 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,

$$\begin{aligned}
X^2 + Y^2 &= (X+Y)^2 - 2XY = s_1^2 - 2s_2, \\
X^3 + Y^3 &= (X+Y)^3 - 3(X+Y)XY = s_1^3 - 3s_1 s_2 \\
X^4 + Y^4 &= (X+Y)^4 - 4(X^2+Y^2)(XY) - 6X^2Y^2 = s_1^4 - 4s_1^2 s_2 + 2s_2^2,
\end{aligned}$$

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,

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

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 α and β are irrational, their elementary symmetric functions are integers: $\boxed{s_1 = \alpha + \beta = -5}$ and $\boxed{s_2 = \alpha\beta = 2}$. Using the formulas in Example 3.5,

$$\begin{aligned}
\alpha^2 + \beta^2 &= p_2 = s_1^2 - 2s_2 = 21, \\
\alpha^3 + \beta^3 &= p_3 = s_1^3 - 3s_1s_2 = -95, \\
\alpha^4 + \beta^4 &= p_4 = s_1^4 - 4s_1^2s_2 + 2s_2^2 = 433.
\end{aligned}$$

Numerically, $\alpha \approx -.438$ and $\beta \approx -4.561$, so $p_r = \alpha^r + \beta^r$ is extremely close to β^r since $|\alpha^r| \to 0$ rapidly as r grows. In fact, p_r turns out to be the nearest integer to β^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 α and β. The power sums are determined by the elementary symmetric functions s_1 and s_2 using exact calculations with integers.

Example 3.8. Let α, β, and γ 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 α, β, and γ are all $\boxed{s_1 = 0}$, $\boxed{s_2 = -1}$, and $\boxed{s_3 = 1}$, so by the formulas in Example 3.6,

$$\begin{aligned}
\alpha^2 + \beta^2 + \gamma^2 &= p_2 = s_1^2 - 2s_2 = 2, \\
\alpha^3 + \beta^3 + \gamma^3 &= p_3 = s_1^3 - 3s_1s_2 + 3s_3 = 3, \\
\alpha^4 + \beta^4 + \gamma^4 &= p_4 = s_1^4 - 4s_1^2s_2 + 2s_2^2 + 4s_1s_3 = 2.
\end{aligned}$$

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 α, β, and γ 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 $s_k(X_1, \dots, X_n) = 0$ 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

$$\begin{aligned}
p_1 &= s_1, \\
p_2 &= s_1^2 - 2s_2, \\
p_3 &= s_1^3 - 3s_1s_2 + 3s_3, \\
p_4 &= s_1^4 - 4s_1^2s_2 + 2s_2^2 + 4s_1s_3 - 4s_4.
\end{aligned}$$

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:

$$p_2 = \det\begin{pmatrix} s_1 & 1 \\ 2s_2 & s_1 \end{pmatrix}, \quad p_3 = \det\begin{pmatrix} s_1 & 1 & 0 \\ 2s_2 & s_1 & 1 \\ 3s_3 & s_2 & s_1 \end{pmatrix}, \quad p_4 = \det\begin{pmatrix} s_1 & 1 & 0 & 0 \\ 2s_2 & s_1 & 1 & 0 \\ 3s_3 & s_2 & s_1 & 1 \\ 4s_4 & s_3 & s_2 & s_1 \end{pmatrix}.$$

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 $t \in \mathbf{Z}$. 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 $F[Y_1, \ldots, Y_n]$ (that is, each monomial in $Y_1, \ldots, Y_n$ 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

$$g(s_1, \ldots, s_n) = 0 \text{ in } F[X_1, \ldots, X_n] \Longrightarrow g(Y_1, \ldots, Y_n) = 0 \text{ in } F[Y_1, \ldots, Y_n].$$

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:

$$g(Y_1, \ldots, Y_n) = \sum_{\mathbf{i}} c_{\mathbf{i}} \mathbf{Y}^{\mathbf{i}}$$

and some $c_{\mathbf{i}}$ is nonzero. Therefore

$$g(s_1, \ldots, s_n) = \sum_{\mathbf{i}} c_{\mathbf{i}} s_1^{i_1} \cdots s_n^{i_n}. \tag{4.1}$$

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_i with s_i,

$$g(s_1, s_2) = s_1^5 + s_2^5 = (X_1 + X_2)^5 + (X_1 X_2)^5,$$

and the leading term is $(X_1 X_2)^5$, which comes from the non-leading term Y_2^5 in $g(Y_1, Y_2)$ after substituting s_i 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:

$$\begin{aligned} \operatorname{mdeg}(s_1^{i_1} \cdots s_n^{i_n}) &= \sum_{k=1}^{n} i_k \operatorname{mdeg}(s_k) \\ &= i_1(1,0,0,\ldots,0) + i_2(1,1,0,\ldots,0) + \cdots + i_n(1,1,1,\ldots,1) \\ &= (i_1 + i_2 + \cdots + i_n, i_2 + \cdots + i_n, \ldots, i_{n-1} + i_n, i_n). \end{aligned} \tag{4.2}$$

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

$$\begin{aligned} \operatorname{mdeg}(s_1^{i_1} \cdots s_n^{i_n}) &= (i_1 + i_2 + \cdots + i_n, i_2 + \cdots + i_n, \ldots, i_{n-1} + i_n, i_n), \\ \operatorname{mdeg}(s_1^{j_1} \cdots s_n^{j_n}) &= (j_1 + j_2 + \cdots + j_n, j_2 + \cdots + j_n, \ldots, j_{n-1} + j_n, j_n) \end{aligned}$$

from rightmost to the leftmost coordinates shows that if $\operatorname{mdeg}(s_1^{i_1} \cdots s_n^{i_n}) = \operatorname{mdeg}(s_1^{j_1} \cdots s_n^{j_n})$ then $\mathbf{i} = \mathbf{j}$. Therefore when the n-tuples $\mathbf{i}$ and $\mathbf{j}$ are different, the polynomials $s_1^{i_1} \cdots s_n^{i_n}$ and $s_1^{j_1} \cdots s_n^{j_n}$ in $F[X_1, \ldots, X_n]$ have different multidegrees and leading coefficients 1 (each s_i has leading coefficient 1). Therefore the terms in (4.1) with nonzero coefficient $c_{\mathbf{i}}$ 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:

$$\operatorname{mdeg}(f_1) > \operatorname{mdeg}(f_2), \ldots, \operatorname{mdeg}(f_m) \Longrightarrow \operatorname{mdeg}(f_1 + f_2 + \cdots + f_m) = \operatorname{mdeg}(f_1).$$

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

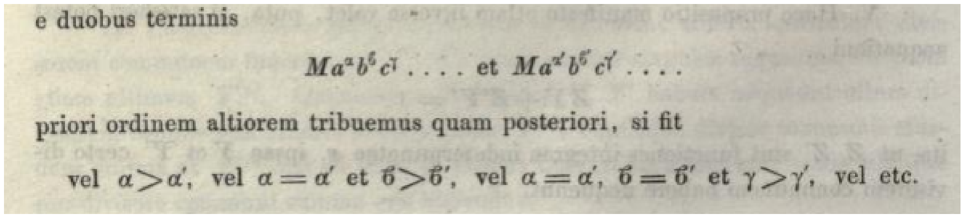
Example 4.2. The only expression of $X^4 + Y^4 + Z^4$ as a polynomial in s_1, s_2, and s_3 is the one appearing in (3.5).

In the proofs of Theorems 1.5 and 4.1, the fact that the coefficients come from a field F is not important; we never had to divide in F. The same proof shows for all nonzero commutative rings R that a symmetric polynomial in $R[X_1, \ldots, X_n]$ lies in $R[s_1, \ldots, s_n]$ in exactly one way. (There is a slight technicality to be aware of: if R is not a domain then the formula $\operatorname{mdeg}(fg) = \operatorname{mdeg} f + \operatorname{mdeg} g$ 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 relevant case of elementary symmetric polynomials $s_1, \ldots, s_n$ since their leading coefficients equal 1.)

Example 4.3. Taking α and β as in Example 3.7, their elementary symmetric functions are both integers, so every symmetric polynomial in α and β with integral coefficients is an integral polynomial in $\alpha + \beta$ and $\alpha\beta$ with integral coefficients, and thus is an integer. 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 $\mathbf{Z}$.

5. History

Lexicographic ordering on multivariable polynomials and its application to the proof of Theorem 1.5 go back to Gauss [3, 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.*".



e duobus terminis

$$Ma^{\alpha} b^{\beta} c^{\gamma} \ldots \text{ et } Ma^{\alpha'} b^{\beta'} c^{\gamma'} \ldots$$

priori ordinem altiorem tribuemus quam posteriori, si fit

vel $\alpha > \alpha'$, vel $\alpha = \alpha'$ et $\beta > \beta'$, vel $\alpha = \alpha'$, $\beta = \beta'$ et $\gamma > \gamma'$, vel etc.

Figure 1. Gauss's definition of lexicographic ordering.

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

[1] D. Cox, J. Little, D. O'Shea, "Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra," Springer-Verlag, New York, 1992.

[2] D. Cox, "Galois Theory," 2nd ed., Wiley, Hoboken, 2012.

[3] C. F. Gauss, Demonstratio nova altera theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse, *Comm. Soc. Reg. Sci. Göttingen* **3** (1816), 107–142; pp. 33–56 in *Werke*, Band III (1866). Online at https://archive.org/details/werkecarlf03gausrich/page/n43/mode/2up. English translation at http://www.paultaylor.eu/misc/gauss.pdf.

[4] S. Lang, "Algebra," revised 3rd ed., Springer-Verlag, New York, 2002.