LECTURES 5–6.

SYMMETRIC POLYNOMIALS IN SEVERAL VARIABLES

S. YAKOVENKO

1. Notation

In the situation where we have three variables, the traditional labels for them are the letters x, y, z. However, we will use a different notation, stressing the full democracy for all variables. Throughout this lecture, n will always stand for the *number of different variables*, and the variables themselves will be labeled as $x_1, \dots, x_n$. The letter x will be used to denote the tuple of all variables, so that we will deal with polynomials denoted by $p(x), q(x)$ from the ring $\mathbb{R}[x]$ which is a shorthand for $\mathbb{R}[x_1, \dots, x_n]$, etc.

A *monomial* is a product of the form $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$, where $\alpha_i \geqslant 0$ are the corresponding nonnegative powers; $\alpha_i = 0$ means that x_i is absent in the above product. We will, following the above logic, denote by $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_+^n$ the tuple of these powers and abbreviate the product above as x^α.

If $\alpha, \beta \in \mathbb{Z}_+^n$ are two such tuples (vectors), then obviously the sum $\alpha + \beta$ makes perfect sense as a vector (componentwise) sum. The rule of product means that

$$x^\alpha \cdot x^\beta = x^{\alpha+\beta} \qquad \forall \alpha, \beta \in \mathbb{Z}_+^n. \tag{1}$$

The vector α can also appear as a (multiple) index, so that a general polynomial in n variables can be written as a (finite) sum $p(x) = \sum_\alpha c_\alpha x^\alpha$. To stress the fact that the polynomial has degree d, we will use the convention that the *norm* ("length") of α is the sum of its entries:

$$\alpha = (\alpha_1, \dots, \alpha_n) \qquad |\alpha| = \alpha_1 + \cdots + \alpha_n.$$

Note that $|\alpha| \geqslant 0$, and $|\alpha| = 0$ if and only if $\alpha = (0, \dots, 0)$. Using these conventions, we can write a general polynomial in n variables of degree $\leqslant n$ with real coefficients as the sum

$$p(x) = \sum_{|\alpha| \leqslant d} c_\alpha x^\alpha, \qquad c_\alpha \in \mathbb{R}.$$

2. Symmetric polynomials

Definition 1. A polynomial $p \in \mathbb{R}[x]$ is called *symmetric*, if it remains unchanged by any permutation of variables.

To make this intuitive definition slightly more formal, consider the *full permutation group on n symbols*: by definition, each permutation is one-to-one transformation of the set $\{1, 2, \dots, n\}$. Permutations of this type form a group with the operation "composition", denoted usually by S_n (this group is often called also the *symmetric group*). A permutation $g \in S$ acts on the symbol $i \in \{1, 2, \dots, n\}$ and maps it into $g(i) \in \{1, 2, \dots, n\}$ so that $g(i) \neq g(j)$ if $i \neq j$.

The standard way to write explicitly a permutation is by a two-row matrix with columns containing i and $g(i)$ for all labels. For instance, a cyclical permutation on n symbols is the permutation

$$\begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix}$$

Permutation on symbols can be extended to act also on any objects labeled by these symbols. For instance, a permutation g can be extended as a transformation of the space $g_* \colon \mathbb{R}^n \to \mathbb{R}^n$ into itself sending the tuple (point) $(x_1, \dots, x_n)$ into the point $g_* x = (x_{g(1)}, x_{g(2)}, \dots, x_{g(n)})$. In the same way the symmetric group acts on monomials by the following rules:

$$g^*(x_i) = x_{g(i)}, \quad g^*(x_i x_j) = x_{g(i)} x_{g(j)}, \quad g^*(x_i x_j x_k) = x_{g(i)} x_{g(j)} x_{g(k)}, \quad \cdots$$
$$i, j, k, \cdots = 1, \dots, n.$$

In the abbreviated form this action can be described as follows,

$$g^*(x^\alpha) = (g_* x)^\alpha, \qquad g \in S_n, \ \alpha \in \mathbb{Z}_+^n.$$

Finally, after explaining how a "permutation of the variables" g acts on an arbitrary polynomial, we extend it by linearity:

$$p = \sum_\alpha c_\alpha x^\alpha \implies g^* p = \sum_\alpha c_\alpha (g_* x)^\alpha, \qquad g \in S_n.$$

Lemma 2. *The action g^* respects both the addition/substraction and the multiplication operations on polynomials: for any two polynomials $p, q \in \mathbb{R}[x]$, we have $g^*(p \pm q) = g^* p \pm g^* q$, $g^*(p \cdot q) = (g^* p) \cdot (g^* q)$. The action preserves the degree:* $\deg g^* p = \deg p$.

Proof. Obvious. □

Now we can give a formal definition of a symmetric polynomial.

Definition 3. A polynomial $p \in \mathbb{R}[x]$ is called symmetric, if $g^* p = p$ for all permutations $g \in S_n$.

Example 4. The polynomials $\sigma_1(x) = x_1 + \cdots + x_n$ and $\sigma_n(x) = x_1, \dots, x_n$ are obviously symmetric. Besides that, all sums of powers $s_k(x) = \sum_1^n x_i^k$, $k = 2, n, \dots$, are obviously symmetric.

Our global task is to describe all symmetric polynomials in $n \geqslant 3$ and find for them representation analogous to that in two variables.

3. Construction of symmetric polynomials

After constructing the action of the permutation group $G = S_n$ on the polynomials, we have a situation already familiar from Lecture 3. For any degree d the polynomials of degree $\leqslant d$ have the structure of a (finite-dimension) linear space over $\mathbb{R}$ and the group acts on this space by linear transformations.

Problem 5. Compute the dimension of the space of homogeneous polynomials of degree exactly d in n independent variables. Compute the dimension of polynomials of degree $\leqslant d$.

The symmetric group S_n is finite, it contains $n!$ distinct permutations. Therefore we can apply the construction of symmetrization to construct sufficiently many symmetric polynomials. We recall it.

Definition 6. The *symmetrization* of a polynomial $p \in \mathbb{R}[x]$ is the polynomial

$$\mathsf{E}p = \frac{1}{|S_n|} \sum_{g \in S_n} g^* p. \tag{2}$$

Remark 7. We use here the notation $\mathsf{E}p$, traditional for the Probability Theory, where it denotes the expectation of a random variable. This is not accidental: the polynomial $\mathsf{E}p$ can be considered as the expectation of the polynomial p by a "random permutation" of its variables (all permutations being considered as equiprobable). The result is a new polynomial.

The operator E is a projection: $\mathsf{E}(\mathsf{E}p) = \mathsf{E}p$ for any polynomial p.

Theorem 8. *The polynomial p is symmetric if and only if $\mathsf{E}p = p$.*

Proof. For the sake of better understanding, we repeat here the proof from Lecture 3.

If p is symmetric, then $g^*p = p$ for any $g \in S_n$. Since g respects all operations in the ring $\mathbb{R}[x]$, the action of any g preserves each term in the sum (2), hence $\mathsf{E}p = p$.

Conversely, if $\mathsf{E}p = p$, then, applying an arbitrary $h \in S_n$ to the sum and using the fact that h preserves all operations, we conclude that

$$h^*(\mathsf{E}p) = \frac{1}{|S_n|} \sum_g g^* p = \frac{1}{|S_n|} \sum_g h^*(g^* p) = \frac{1}{|S_n|} \sum_g (hg)^* p.$$

Yet for any fixed h if g runs over the entire group S_n, then so does the composition $hg \in S_n$. Thus the result of the application is only permutation of different terms in the sum, which does not affect the result. Thus $h^*p = h^*(\mathsf{E}p) = p$, i.e., p is symmetric. $\square$

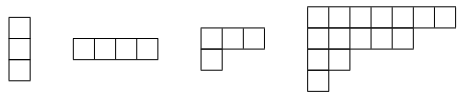
This result gives a way for bulk production of symmetric polynomials:

(1) For any $i = 1, \dots, n$, $\mathsf{E}(x_i) = \frac{1}{n!} \sum_1^n (n-1)! x_i = \frac{1}{n} \sum_i x_i = \frac{1}{n} \sigma_1(x)$.
(2) In the same way, $\mathsf{E}(x_i^k) = \frac{1}{n} \sum x_i^k = \frac{1}{n} s_k(x)$ is the sum of powers.

(3) $\mathsf{E}(x_1 \cdots x_{n-1}) = \prod_1^n x_i \cdot \frac{1}{n} \sum_i \frac{1}{x_i} = \frac{1}{n}\sigma_n(x) s_{-1}(x)$.
(4) $\mathsf{E}(x_1 x_2) = \frac{2}{n(n-1)} \sum_{i<j} x_i x_j$—a symmetric polynomial of degree 2, denoted by $\sigma_2(x)$.

The construction can be applied to any monomial x^α. The result of the averaging is the sum of monomials with a numeric coefficient. Without loss of generality we may always assume that the vector of powers α consists of *non-increasing* numbers. Such vector can be graphically represented as a so called Young tableau (table), the collection of columns of height $\alpha_1 \geqslant \alpha_2 \geqslant \cdots$; columns of height zero are not plotted.

Example 9. The Young diagrams



correspond to monomials

$$x_1^3, \qquad x_1 x_2 x_3 x_4, \qquad x_1^2 x_2 x_3, \qquad x_1^4 x_2^3 (x_3 x_4 x_5)^2 x_6 x_7.$$

The total area of the tableau is the degree d of the monomials, the width (number of columns) is *less or equal* to the number of variables n.

The result of application of the averaging operator $\mathsf{E}(x^\alpha)$ depends only on the Young diagram of the monomial. The numeric coefficient can be obtained by solving the corresponding combinatorial problem.

Example 10. The result of averaging of the monomial

$$\mathsf{E}(\square\square \cdots \square)$$

of degree $d \leqslant n$ is a sum of $\binom{n}{d}$ monomials with the reciprocal coefficient $1/\binom{n}{d}$.

Example 11. The result of averaging $\mathsf{E}(\begin{smallmatrix}\square & \square \\ \square & \end{smallmatrix})$ is the sum of $2\binom{n}{2} = n(n-1)$ monomials of the form $x_i^2 x_j$, with the reciprocal coefficient. Indeed, the label for the quadratic term can be chosen by n ways, and the remaining label by $(n-1)$ ways. The terms $x_i^2 x_j$ and $x_j^2 x_i$ are different, so no need to divide by 2.

Problem 12. Assume that the Young diagram consists of $\nu_1 \geqslant 1$ columns of maximal height, followed by $\nu_2 \geqslant 1$ columns of second biggest height, followed by ν_3 columns of third biggest height *etc*, followed by $\nu_0 \geqslant$ (invisible) zero-height columns. Prove that the number of monomials after averaging will be

$$\frac{n!}{\nu_1! \nu_2! \cdots \nu_0!}.$$

In the preceding Example we have $\nu_1 = \nu_2 = 1$, $\nu_0 = n-2$, so the answer is indeed $\frac{n!}{1!\,1!\,(n-2)!} = n(n-1)$.

The operator E is obviously linear (respects the operation $+$ on the ring $\mathbb{R}[x]$). Unfortunately, it does not respect the multiplication of polynomials, except for very special cases. This should be compared with the fact that the expectation of product of random variables is in general different from the product of their expectations, except when the variables are independent (the notion of independence is the cornerstone of the Probability theory, which makes it really different from the integration theory).

Example 13. Comparing $\mathsf{E}(x_1^p x_2^q)$ with the product $\mathsf{E}(x_1^p)\mathsf{E}(x_2^q) = \mathsf{E}(x_1^p)\mathsf{E}(x_1^q)$ we see that the latter is

$$\frac{1}{n^2}\sum_{i=1}^{n} x_i^p \cdot \sum_{j=1}^{n} x_j^q = \frac{1}{n^2}\sum_{1}^{n} x_i^{p+q} + \frac{2}{n^2}\sum_{i<j} x_i^p x_j^q = \frac{1}{n}\mathsf{E}(x_1^{p+q}) + \frac{n-1}{n}\mathsf{E}(x_1^p x_2^q).$$

In other words,

$$(n-1)\mathsf{E}(x_1^p x_2^q) = n\mathsf{E}(x_1^p)\mathsf{E}(x_2^q) - \mathsf{E}(x_1^{p+q}).$$

4. Elementary symmetric polynomials

Another (direct) way to produce symmetric polynomials is to use the obvious fact *that any monic* (i.e., with the unit leading coefficient) *polynomial in one variable λ is uniquely determined by its roots in a way non-sensitive to the permutation of these roots.*

In particular, all other (non-leading) coefficients of this polynomial are symmetric functions of the roots. These functions can easily be computed and turn out to be polynomials. The result is known as the (collection of) Vieta formulas.

Theorem 14 (Vieta theorem)**.** *The product*

$$(\lambda - x_1)(\lambda - x_2)\cdots(\lambda - x_n) = \prod_{i=1}^{n}(\lambda - x_i)$$

is the monic polynomial of degree n in the variable λ which can be expanded as

$$\lambda^n - \sigma_1\lambda^{-1} + \sigma_2\lambda^{n-2} - \sigma_3\lambda^{n-3} + \cdots + (-1)^n\sigma_n. \tag{3}$$

The coefficients $\sigma_i = \sigma_i(x) \in \mathbb{R}[x]$, $i = 1, \ldots, n$ are symmetric polynomials of degree i of the parameters (formal variables) $x_1, \ldots, x_n$. They can be described as follows:

$$\begin{aligned}\sigma_k(x) &= \binom{n}{k}\mathsf{E}(\square\square\cdots\square\square) \quad (k \text{ is the area of the diagram})\\ &= \binom{n}{k}\mathsf{E}(x_1x_2\cdots x_k), \qquad k = 1, \ldots, n.\end{aligned}$$

The proof of this theorem is obvious from combinatorial arguments.

Definition 15. The coefficients $\sigma_i(x)$ occurring in the expansion (3), are called *elementary symmetric polynomials* in the variables $x_1, \dots, n$. It is convenient to denote $\sigma_0 \equiv 1$. **It is even more convenient to assume that $\sigma_k(x) \equiv 0$ for all $k > n$**: this will simplify many formulas.

Example 16. In three dimensions we have only one new symmetric polynomial: in addition to $\sigma_1 = x+y+z$ and $\sigma_3 = xyz$, we have $\sigma_2 = xy+yz+xz$.

5. Strategic goal

What we want to do now is to find a suitable explicit description of symmetric polynomials, similar to the theorem in 2 dimensions. In other words, we want to construct a (finite) collection of symmetric polynomials $\rho_1(x), \dots, \rho_k(x) \in \mathbb{R}[x]$, which would be:

- sufficiently rich so that *any* symmetric polynomial p could be represented through them as a composition, $p(x) = P(\rho_1(x), \dots, \rho_k(x))$;
- non-redundant, so that the representation above is unique.

The Grand Conjecture is that the elementary symmetric polynomials $\sigma_1, \dots, \sigma_n$ fit this rather restrictive demand. We will justify it after some efforts.

6. Symmetry break

In a somewhat paradoxical way, the idea of the proof is to break the symmetry and introduce inequality where all variables initially enjoyed "equal rights".

Consider the lexicographic ordering of the monomials. A monomial x^α is said to be *stronger* than another monomial x^β, if:

- $\alpha_1 > \beta_1$, or
- $\alpha_1 = \beta_1$, but $\alpha_2 > \beta_2$, or
- $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$ but $\alpha_3 > \beta_3$,

$$\dots$$

- $\alpha_i = \beta_i$ for all $i = 1, \dots, n-1$, but $\alpha_n > \beta_n$.

In other words, the variable x_1 is considered as "most important", if it cannot distinguish two monomials, then its role is passed to x_2 etc. Using this notion, one can describe the result of averaging of a monomial as follows:

$$\mathsf{E}(x^\alpha) = cx^\beta + \text{weaker terms},$$

where β is the rearrangement of the components of the vector $\alpha \in \mathbb{Z}_+^n$ such that the components $\beta_1, \beta_2, \dots$ are non-increasing, and $c \in \mathbb{Q}$ is a rational coefficient which depends only on the Young diagram of α.

Lemma 17.

$$(x^\alpha + \textit{weaker terms})(x^\beta + \textit{weaker terms}) = x^{\alpha+\beta} + \textit{weaker terms}. \tag{4}$$

Proof. This follows from the following obvious observation: if x^α is stronger than x^β, then for *any* monomial x^γ, $x^{\alpha+\gamma}$ is stronger than $x^{\beta+\gamma}$. □

Can one determine the strongest (leading) term of a monomial in the "variables" σ_i after the expansion? Consider a monomial $\sigma^\beta = \sigma_1^{\beta_1}\sigma_2^{\beta_2}\cdots\sigma_n^{\beta_n}$. The leading terms of each polynomial σ_k is the product $x_1\cdots x_k$:

$$\sigma_k = x_1\cdots x_k + \text{weaker terms.}$$

Recycling Lemma 17, we conclude that the leading term of the above monomial σ^β is

$$x_1^{\beta_1}\cdot(x_1x_2)^{\beta_2}\cdots(x_1x_2\cdots x_n)^{\beta_n} = x_1^{\beta_1+\cdots+\beta_n}x_2^{\beta_2+\cdots+\beta_n}\cdots x_n^{\beta_n}.$$

Proposition 18. *For any monomial x^α with $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_n$, there exists a unique monomial σ^β such that*

$$\sigma^\beta = x^\alpha + \textit{weaker terms.}$$

Proof. One has to solve the equations

$$\begin{aligned}
\alpha_1 &= \beta_1 + \cdots + \beta_n,\\
\alpha_2 &= \beta_2 + \cdots + \beta_n,\\
\cdots &\quad \cdots\\
\alpha_n &= \beta_n
\end{aligned}$$

with respect to β_k: $\beta_1 = \alpha_1 - \alpha_2$, $\beta_2 = \alpha_2 - \alpha_3, \ldots, \beta_n = \alpha_n$. The total degree of this monomial in the variables x_i is $(\alpha_1-\alpha_2)+2(\alpha_2-\alpha_3)+3(\alpha_3-\alpha_4)+\cdots+n\alpha_n = \alpha_1+\cdots+\alpha_n$. □

Problem 19. Consider two "symmetric monomials" σ^β and σ^γ. Can one, looking only at the multi-exponents β, γ, quickly decide, which of these monomials will produce the stronger leading term when expanded in multi-powers of x?

The answer can be described in terms of the dual order. Namely, we say that the monomial $\sigma^\beta = \sigma_1^{\beta_1}\cdots\sigma_n^{\beta_n}$ is stronger than $\sigma^\gamma = \sigma_1^{\gamma_1}\cdots\sigma_n^{\gamma_n}$, if:

- $|\beta| > |\gamma|$, or
- $|\beta| = |\gamma|$, but $\beta_2+\cdots+\beta_n > \gamma_2+\cdots+\gamma_n$, or
- $|\beta| = |\gamma|$, $\beta_2+\cdots+\beta_n = \gamma_2+\cdots+\gamma_n$, but $\beta_3+\cdots+\beta_n > \gamma_3+\cdots+\gamma_n$,
  $\cdots$
- $\sum_{i=k}^n \beta_i = \sum_{i=k}^n \gamma_i$ for all $k = 1, \ldots, n-1$, but $\beta_n > \gamma_n$.

Theorem 20 (Main theorem on symmetric polynomials)**.** *Any symmetric polynomial in the variables $x_1, \ldots, x_n$ can be uniquely expressed as a polynomial in the elementary functions $\sigma_1, \ldots, \sigma_n$.*

Proof. Let $p(x) = c_\alpha x^\alpha + \text{weaker terms}$ be a symmetric polynomial with the leading term explicitly selected. By Proposition 18, there exists a monomial $c\sigma^\beta$ which has the same leading term.

The difference $p - c\sigma^\beta$ is again a symmetric polynomial of the same degree, yet all monomials in it are by construction weaker than x^α, so its leading

term is $c_{\alpha'}x^{\alpha'}$ with α' weaker than α. It remains to note that any lexicographically decreasing (strictly!) sequence of vectors $\alpha, \alpha', \alpha'', \cdots$ must eventually end by the zero vector.

The uniqueness of the expansion follows from the same argument. If a symmetric polynomial p admits two different representations through σ_i, then the difference of these representations is a nontrivial polynomial in the letters σ_i. □

7. Power sums, a.k.a. symmetric powers a.k.a. Newton sums

The above demonstration of the main theorem is algorithmic (i.e., it describes a process which in finitely many steps would transform any symmetric polynomial in x into a polynomial in the elementary symmetric functions σ. However, sometimes we need the explicit answer, especially for the simple symmetric polynomials.

The first case should be, obviously, that of symmetric (Newton) sums $s_k = \sum_{i=1}^n x_i^k$. Here the index k may take any natural value, in particular, greater than n.

Theorem 21 (Newton identities)**.**

$$\sigma_0 s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} - \sigma_3 s_{k-3} \pm \cdots + (-1)^k \sigma_k s_0 = 0. \tag{5}$$

If $k > n$, we use the convention that $\sigma_k = 0$ for $k > n$.

To prove this theorem, we introduce one of the most surprising tricks in combinatorics, namely, the notion of generating functions.

8. Generating functions

Every time we need to answer a combinatorial question or prove an identity about an infinite sequence of numbers or polynomials $\{p_0, p_1, p_2, p_3, \dots\}$, it makes sense to introduce a new fictitious variable, say, t, and instead of the infinite sequence p consider the "generating function" $P(t)$ defined formally by the Taylor series

$$P(t) = \sum_{k=0}^{\infty} p_k\, t^k.$$

This function may depend on additional variables if p_k are not numbers, but functions of this variable.

Example 22. With the sequence of the Newton sums one associates the generating function

$$S(t,x) = \sum_{k=0}^{\infty} s_k(x)\, t^k. \tag{6}$$

With the sequence of elementary symmetric functions one can associate the generating function

$$E(t,x) = \sum_{k=0}^{\infty} \sigma_k(x)\, t^k. \tag{7}$$

Note that we use the convention $s_0(x) = n$ and $\sigma_0(x) = 1$, $\sigma_k(x) = 0$ for $k > n$. Thus effectively the function $E(t)$ is a polynomial in t. Its roots are the negative reciprocals $t_i = -1/x_i$, $i = 1, \dots, n$.

The reason why the generating functions are useful is the following. Standard algebraic operations (multiplication, inversion, derivation, product) on power series and rational functions, when performed componentwise, produce combinatorial formulas for the coefficients which coincide with the recurrent identities defining the sequences. One example is given by the Vieta formulas above, yet a more natural representation is available for the generating function.

Lemma 23.

$$E(t, x) = \prod_{i=1}^{n} (1 + x_i\, t). \tag{8}$$

Proof. Obvious. The coefficient before t^k comes from all possible products of k distinct variables from the list $\{x_1, \dots, x_n\}$. □

Lemma 24.

$$S(t, x) = \sum_{i=1}^{n} \frac{1}{1 - x_i\, t}.$$

Proof. Obvious. One has to use the formula for the sum of the geometric progression $1+z+z^2+\cdots = 1/(1-z)$ in each variable $z = x_i t$ separately. □

The relationship between the functions $E(t, x)$ and $S(-t, x)$ comes from the observation that the logarithmic derivative operator $f(t) \mapsto f'(t)/f(t) = \frac{\mathrm{d}}{\mathrm{d}t} \ln f(t)$ takes the product into a sum. Applying this observation to the function E, we note (the prime denotes the derivative in the variable t) that

$$\frac{E'}{E} = \frac{\mathrm{d}}{\mathrm{d}t} \ln \prod_{i=1}^{n} (1 + x_i\, t) = \sum_{i=1}^{n} \frac{\mathrm{d}}{\mathrm{d}t} \ln(1 + x_i\, t) = \sum_{i=1}^{n} \frac{x_i}{1 + x_i\, t}.$$

The right hand side of the last identity clearly resembles the expression for the function S with some modifications. More precisely, it can be expressed as follows,

$$P(t, x) = \sum_{i=1}^{n} \frac{x_i}{1 + x_i\, t} = \frac{S(-t, x) - s_0}{-t} = \sum_{k=1}^{\infty} s_k(x)(-t)^{k-1}.$$

In other words, the function $P(t, x)$ is a generating function for the "shifted" Newton sums; P is generating for the sequence of the functions $\tilde{s}_j = s_{j+1}$, $j = 0, 1, \dots$.

Comparing these two expressions, we conclude that

$$E'(t) = P(t)E(t).$$

Proof of the Newton identities (*Theorem 21*). Substituting the formulas for the generating functions into the identity $E' = PE$ we conclude with the formal identity

$$\sum_{k=1}^{\infty} k\sigma_k\, t^{k-1} = \left(\sum_{i=0}^{\infty} s_i\,(-t)^{i-1}\right)\left(\sum_{j=0}^{\infty} \sigma_j\, t^j\right).$$

Comparing the coefficients before the powers $0, t, t^2, \dots$, we obtain the infinite series of identities:

$$\begin{aligned}
\sigma_1 &= s_1\sigma_0\\
2\sigma_2 &= s_1\sigma_1 - s_2\sigma_0,\\
3\sigma_3 &= s_1\sigma_2 - s_2\sigma_1 + s_3\sigma_0,\\
&\dots\\
n\sigma_n &= s_1\sigma_{n-1} - s_2\sigma_{n-2} + \dots + (-1)^{n-1}s_n\sigma_0,\\
0 &= s_1\sigma_n - s_2\sigma_{n-1} + s_3\sigma_{n-2} \pm \dots,
\end{aligned}$$

and so on (the left hand side is always zero from the nth step). Since $s_0 = n$, this coincides with the Newton formulas (5). □

Remark 25. The above calculations were performed without any regard to the convergence of the series. This is justified, since the only property which is required is the formal logarithmic derivative of the polynomial E. However, the series are all converging for sufficiently small values of $|x| = |x_1| + \dots + |x_n|$.

The method of generating functions is very powerful in many areas of mathematics (not just in connection with symmetric polynomials).

Example 26. Consider the generating function

$$H(t,x) = \prod_{i=1}^{n} \frac{1}{1 - x_i\, t} = \sum_{k=0}^{\infty} h_k(x)\, t^k, \tag{9}$$

where $h_k(x)$ is a *complete symmetric function of order* k, the sum of all monomials of degree k:

$$h_k(x) = \sum_{|\alpha|=k} x^\alpha, \qquad k = 0, 1, 2, \dots \tag{10}$$

The identity (9) is proved by expanding each fraction $1/(1 - x_i\, t)$ as the sum of the infinite geometric progression, and then apply the combinatorial arguments when multiplying these progressions between themselves.

Then we have the obvious identity $E(-t,x)H(t,x) \equiv 1$, cf. with (8). Computing the coefficients of the product in the left hand side, we conclude with the identities

$$\sum_{k=0}^{m} (-1)^k \sigma_k h_{m-k} = 0 \qquad \forall m = 1, 2, \dots \tag{11}$$

(recall that $\sigma_{n+1} = \sigma_{n+2} = \cdots = 0$). These identities can be used to express the complete symmetric polynomials through the elementary symmetric functions $\sigma_1, \ldots, \sigma_n$ using the recursion with the initial condition $h_0 = 1, h_1 = \sigma_1$.

Remark 27. The symmetric powers $s_0, \ldots, s_n$ can be used as a different basis for representation of all symmetric functions. Indeed, the same Newton identities can be interpreted as the recurrent formulas expressing σ_k through $s_0, \ldots, s_k$. It may be worth giving a try sometimes (cf. with Problem 32 below).

9. Miscellaneous problems

Problem 28. Factor out the polynomial $x^3 + y^3 + z^3 - 3xyz \in \mathbb{R}[x, y, z]$.

Solution. This polynomial expands as $s_3 - 3\sigma_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3 - 3\sigma_3 = \sigma_1(\sigma_1^2 - 3\sigma_2)$. □

Remark 29. Warning! In general, a symmetric polynomial can be factored into non-symmetric factors, which are permuted by the action of the symmetric group S_n. The above example is rather specific.

Problem 30. Factor out $(x + y)(y + z)(x + z) + xyz$.

Problem 31. Factor out $p = (x - y)^3 + (y - z)^3 + (z - x)^3$.

Solution. Denote $X = x-y$, $Y = y-z$, $Z = z-x$. Then $X+Y+Z = 0$, and we need to factorize the polynomial $s_3(X, Y, Z) = X^3 + Y^3 + Z^3$. Using the Newton formula, we obtain $p = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$, where σ_i are elementary symmetric functions of X, Y, Z. Since $\sigma_1 = 0$, we have $p = 3\sigma_3 = 3XYZ = 3(x - y)(y - z)(z - x)$. □

Clearly, this approach allows to simplify all symmetric expressions under symmetric assumptions on the variables. In particular, many symmetric sums s_k can be factorized under the assumption that $\sigma_1 = 0$.

Problem 32. Find $x^4 + y^4 + z^4$ if it is known that $x + y + z = 0$ and $x^2 + y^2 + z^2 = 1$.

Symmetric representation can be useful not just for solving equations and proving identities. Inequalities can also be proved. The central fact is, of course, the inequality on arithmetic and geometric means,

$$\frac{x_1 + \cdots + x_n}{n} \geqslant (x_1 \cdots x_n)^{\frac{1}{n}}, \tag{12}$$

which is valid for all positive real values of the symbols $x_1, \ldots, x_n$. In the symmetric representation it takes the form

$$\sigma_1^n - n^n \sigma_n \geqslant 0. \tag{13}$$

The "dual form" of this inequality reads as

$$\sigma_{n-1}^n - n^n \sigma_n^{n-1} \geqslant 0.$$

Prove it by considering the sum $s_{-1}(x) = \sum \frac{1}{x_i} = \frac{\sigma_{n-1}}{\sigma_n}$.

Problem 33. Prove that for real positive $x_1, \dots, x_n$

$$(x_1 + \cdots + x_n)\left(\frac{1}{x_1} + \cdots + \frac{1}{x_n}\right) \geqslant n^2.$$

Solution. After division by n^2m the left hand side is the product of the arithmetic mean of the numbers x_i by the arithmetic mean of their reciprocals $\frac{1}{x_i}$. By the Main Theorem, this is greater or equal to the product of the geometric means $\sqrt[n]{x_1 \cdots x_n}$ and $\sqrt[n]{\frac{1}{x_1 \cdots x_n}}$ which is one. $\square$

Problem 34. Prove that for positive real x, y, z

$$x^3 + y^3 + z^3 \geqslant 3xyz.$$

Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel, email: yakov@weizmann.ac.il