

Math 701 Fall 2026, Lecture 4: Monomial symmetric polynomials

website: <https://www.cip.ifi.lmu.de/~grinberg/t/26fs/>

1.3. Partitions and monomial symmetric polynomials

Partitions are one of the most classical combinatorial objects. They have intrinsic interest (see, e.g., [21s, Chapter 4] for some of their properties) and crucial connections to number theory and representation theory, but we will be chiefly interested in them as indexing sets for (K -module) bases of $\mathcal{S}$. Let us define two variants of them:

Definition 1.3.1.

- (a) An **(integer) partition** means a weakly decreasing finite tuple of positive integers, i.e., a tuple $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ with $\lambda_1, \lambda_2, \dots, \lambda_k \in \{1, 2, 3, \dots\}$ and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$. The **length** of this partition is k , while its **size** is $|\lambda| := \lambda_1 + \lambda_2 + \dots + \lambda_k$. More generally, the size of any k -tuple $a = (a_1, a_2, \dots, a_k)$ of integers is defined to be $|a| := a_1 + a_2 + \dots + a_k$.
- (b) An **N -partition** means a weakly decreasing N -tuple of nonnegative integers, i.e., a tuple $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N)$ with $\lambda_1, \lambda_2, \dots, \lambda_N \in \mathbb{N}$ and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$. The **size** of this N -partition is $\lambda_1 + \lambda_2 + \dots + \lambda_N$, and is denoted by $|\lambda|$.

For example, $(4, 2, 2, 1)$ is a partition, whereas $(4, 2, 2, 1, 0)$ is not a partition but is a 5-partition (since N -partitions are allowed to contain zeroes, but partitions aren't). Both of these have size 9. Note that the empty tuple $()$ is also a partition (of length 0 and size 0).

The two above versions of partitions are very close to each other:

Proposition 1.3.2. There is a bijection

$$\{\text{partitions of length } \leq N\} \rightarrow \{N\text{-partitions}\},$$

$$(\lambda_1, \lambda_2, \dots, \lambda_k) \mapsto \left(\lambda_1, \lambda_2, \dots, \lambda_k, \underbrace{0, 0, \dots, 0}_{N-k \text{ zeroes}} \right),$$

which preserves the size of the respective partition. Namely, this bijection pads the partition with zeroes (at its end) in order to increase its length to N . The inverse map simply removes all zeroes at the end of the given N -partition.

Definition 1.3.3. We shall use this bijection to identify the N -partitions with the partitions of length $\leq N$. Thus, for example, the partition $(3, 1, 1)$ becomes identified with the 3-partition $(3, 1, 1)$, with the 4-partition $(3, 1, 1, 0)$, with the 5-partition $(3, 1, 1, 0, 0)$, and so on. For another example, the empty partition $()$ is identified with the N -partition $(0, 0, \dots, 0)$ for each $N \in \mathbb{N}$.

We will also regard partitions as infinite sequences by padding them with infinitely many zeroes at the end. So the partition $(3, 1, 1)$ becomes identified with the infinite sequence $(3, 1, 1, 0, 0, 0, \dots)$. Thus, each partition can be viewed uniquely as an infinite weakly decreasing sequence $(\lambda_1, \lambda_2, \lambda_3, \dots)$ of nonnegative integers that has only finitely many nonzero entries (which necessarily are all clustered at its start because it is weakly decreasing).

We next define a few basic notations related to monomials:

Definition 1.3.4. Let m be a monomial, and let f be a polynomial or power series. Then, $[m]f$ denotes the m -coefficient of f .

For instance, $[x^2](1+x)^4 = 6$.

Definition 1.3.5. Let $a = (a_1, a_2, \dots, a_N) \in \mathbb{N}^N$ be an N -tuple of nonnegative integers.

- (a) We let x^a be the monomial $x_1^{a_1} x_2^{a_2} \cdots x_N^{a_N}$.
- (b) We let $\text{sort } a$ be the N -partition obtained from a by sorting the entries in weakly decreasing order.

For instance, $\text{sort}(1, 0, 2, 4, 1) = (4, 2, 1, 1, 0)$ and $x^{(1,0,2,4,1)} = x_1^1 x_2^0 x_3^2 x_4^4 x_5^1 = x_1 x_3^2 x_4^4 x_5$.

The following simple lemma helps us feel at home with the notations we just introduced:

Lemma 1.3.6. Let $f \in \mathcal{P}$ be a polynomial. Then:

- (a) For any N -tuple $a = (a_1, a_2, \dots, a_N) \in \mathbb{N}^N$ and any permutation $\sigma \in S_N$, we have

$$[x^a](\sigma \cdot f) = [x^{a \circ \sigma}]f,$$

where $a \circ \sigma$ denotes the N -tuple $(a_{\sigma(1)}, a_{\sigma(2)}, \dots, a_{\sigma(N)})$.

- (b) The polynomial f is symmetric (i.e., belongs to $\mathcal{S}$) if and only if all $a \in \mathbb{N}^N$ satisfy

$$[x^a]f = [x^{\text{sort } a}]f.$$

Proof. **(a)** Let $a = (a_1, a_2, \dots, a_N) \in \mathbb{N}^N$ be any N -tuple. The polynomial $\sigma \cdot f$ is obtained from f by substituting $x_{\sigma(i)}$ for each indeterminate x_i ; this substitution clearly transforms each monomial $x_1^{b_1} x_2^{b_2} \cdots x_N^{b_N}$ into $x_{\sigma(1)}^{b_1} x_{\sigma(2)}^{b_2} \cdots x_{\sigma(N)}^{b_N}$. In particular, this substitution transforms the monomial

$$x^{a \circ \sigma} = x_1^{a_{\sigma(1)}} x_2^{a_{\sigma(2)}} \cdots x_N^{a_{\sigma(N)}} \quad \left(\text{since } a \circ \sigma = (a_{\sigma(1)}, a_{\sigma(2)}, \dots, a_{\sigma(N)}) \right)$$

into

$$\begin{aligned} x_{\sigma(1)}^{a_{\sigma(1)}} x_{\sigma(2)}^{a_{\sigma(2)}} \cdots x_{\sigma(N)}^{a_{\sigma(N)}} &= x_1^{a_1} x_2^{a_2} \cdots x_N^{a_N} && \text{(since } \sigma \text{ is a permutation of } [N]) \\ &= x^a \end{aligned}$$

(and no other monomials get transformed into x^a , since this substitution is clearly bijective). Thus, the x^a -coefficient of $\sigma \cdot f$ is precisely the $x^{a \circ \sigma}$ -coefficient of f . That is, we have $[x^a](\sigma \cdot f) = [x^{a \circ \sigma}]f$. This proves Lemma 1.3.6 **(a)**.

(b) $\implies$: Assume that f is symmetric. Let $a \in \mathbb{N}^N$ be an N -tuple. For any $\sigma \in S_N$, define an N -tuple $a \circ \sigma$ as in part **(a)**. Now, $\text{sort } a$ is the N -tuple obtained by sorting the entries of a in weakly decreasing order; in particular, it is obtained by permuting the entries of a . That is, $\text{sort } a = a \circ \sigma$ for some permutation $\sigma \in S_N$. Consider this σ . From part **(a)**, we know that $[x^a](\sigma \cdot f) = [x^{a \circ \sigma}]f$. Since $\sigma \cdot f = f$ (because f is symmetric) and $a \circ \sigma = \text{sort } a$, we can rewrite this as $[x^a]f = [x^{\text{sort } a}]f$. Thus, we have shown that all $a \in \mathbb{N}^N$ satisfy $[x^a]f = [x^{\text{sort } a}]f$. This proves the “ $\implies$ ” direction of Lemma 1.3.6 **(b)**.

$\impliedby$: Assume that all $a \in \mathbb{N}^N$ satisfy

$$[x^a]f = [x^{\text{sort } a}]f. \quad (1)$$

We must show that f is symmetric.

Let $\sigma \in S_N$. Let $a \in \mathbb{N}^N$ be an N -tuple. Then, the N -tuple $a \circ \sigma$ (defined as in part **(a)**) is obtained by permuting the entries of a ; thus it has the same entries as a , just in a different order. Consequently, the N -tuples $a \circ \sigma$ and a become identical if we sort them both into weakly decreasing order. In other words, $\text{sort}(a \circ \sigma) = \text{sort } a$. However, applying (1) to $a \circ \sigma$ instead of a , we find $[x^{a \circ \sigma}]f = [x^{\text{sort}(a \circ \sigma)}]f = [x^{\text{sort } a}]f$ (since $\text{sort}(a \circ \sigma) = \text{sort } a$). Comparing this with (1), we obtain $[x^{a \circ \sigma}]f = [x^a]f$. Meanwhile, Lemma 1.3.6 **(a)** yields $[x^a](\sigma \cdot f) = [x^{a \circ \sigma}]f = [x^a]f$ (since we just proved $[x^{a \circ \sigma}]f = [x^a]f$).

Forget that we fixed a . We thus have shown that $[x^a](\sigma \cdot f) = [x^a]f$ for each $a \in \mathbb{N}^N$. In other words, each coefficient of the polynomial $\sigma \cdot f$ equals the corresponding coefficient of f . Hence, these two polynomials are equal. That is, $\sigma \cdot f = f$. We have proved this for each $\sigma \in S_N$; that is, f is symmetric. This proves the “ $\impliedby$ ” direction of Lemma 1.3.6 **(b)**. $\square$

Definition 1.3.7. Let λ be any N -partition. Then, the λ -th **monomial symmetric polynomial** is defined to be

$$m_\lambda := \sum_{\substack{a \in \mathbb{N}^N; \\ \text{sort } a = \lambda}} x^a \in \mathcal{S}.$$

Example 1.3.8. Let $N = 3$. Then,

$$\begin{aligned} m_{(2,1,0)} &= \sum_{\substack{a \in \mathbb{N}^N; \\ \text{sort } a = (2,1,0)}} x^a \\ &= x^{(2,1,0)} + x^{(1,2,0)} + x^{(1,0,2)} + x^{(2,0,1)} + x^{(0,2,1)} + x^{(0,1,2)} \\ &= x^2y + xy^2 + xz^2 + x^2z + y^2z + yz^2 \end{aligned}$$

(again using the shorthands x, y, z for x_1, x_2, x_3) and

$$\begin{aligned} m_{(3,2,1)} &= x^3y^2z + x^2y^3z + x^2yz^3 + x^3yz^2 + xy^3z^2 + xy^2z^3 \\ &= x^3y^2z + (\text{all other 5 permutations of this monomial}) \end{aligned}$$

and

$$m_{(2,2,1)} = x^2y^2z + x^2yz^2 + xy^2z^2$$

and

$$m_{(2,2,2)} = x^2y^2z^2.$$

As we saw in this example, for each N -partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N)$, the polynomial m_λ is the sum of all monomials in the S_N -orbit of the single monomial $x^\lambda = x_1^{\lambda_1} x_2^{\lambda_2} \dots x_N^{\lambda_N}$. Thus, the m_λ are also known as **orbit sums**. And nowadays, they are often just called “the m ’s”, since the notation m_λ has become so standard that everyone can be expected to understand this shorthand.

The definition of m_λ shows that all coefficients of m_λ are 0 or 1. More precisely: each N -partition λ and each N -tuple $a \in \mathbb{N}^N$ satisfy

$$[x^a] m_\lambda = \begin{cases} 1, & \text{if } \text{sort } a = \lambda; \\ 0, & \text{if } \text{sort } a \neq \lambda. \end{cases} \quad (2)$$

In particular, for each N -partition μ , we have

$$\begin{aligned} [x^\mu] m_\lambda &= \begin{cases} 1, & \text{if } \text{sort } \mu = \lambda; \\ 0, & \text{if } \text{sort } \mu \neq \lambda \end{cases} \\ &= \begin{cases} 1, & \text{if } \mu = \lambda; \\ 0, & \text{if } \mu \neq \lambda \end{cases} \end{aligned} \quad (3)$$

(because $\text{sort } \mu = \mu$, since μ is already weakly decreasing).

It is also clear that each $a \in \mathbb{N}^N$ satisfies

$$\deg(x^a) = |a| = |\text{sort } a| \quad (4)$$

(since $\text{sort } a$ has the same entries as a up to order). Thus, each monomial symmetric polynomial m_λ is homogeneous of degree $|\lambda|$.

Our symmetric polynomials e_n, h_n, p_n are all easily rewritten in terms of m 's:

Proposition 1.3.9.

(a) For each $n \in \{0, 1, \dots, N\}$, we have

$$e_n = m_{(1, 1, \dots, 1, 0, 0, \dots, 0)},$$

where $(1, 1, \dots, 1, 0, 0, \dots, 0)$ is the N -tuple that begins with n many 1's and ends with $N - n$ many 0's. This N -tuple will be written as $(1^n, 0^{N-n})$ later on.

(b) For each $n \in \mathbb{N}$, we have

$$h_n = \sum_{\substack{\lambda \text{ is an } N\text{-partition;} \\ |\lambda| = n}} m_\lambda.$$

(c) Assume that $N > 0$. Then, for each $n \in \mathbb{N}$, we have

$$p_n = m_{(n, 0, 0, \dots, 0)}.$$

Under our identification of N -partitions with partitions of length $\leq N$, we can (and will) simply write this as

$$p_n = m_{(n)}.$$

Proof. LTTR. (This means “left to the reader”. In this case, the proof is so easy that just starting at the definitions should suffice.) $\square$

Next, we shall prove the following (fairly simple) theorem:

Theorem 1.3.10 (monomial basis of $\mathcal{S}$).

(a) The family $(m_\lambda)_{\lambda \text{ is an } N\text{-partition}}$ is a basis of the K -module $\mathcal{S}$.

(b) Each symmetric polynomial $f \in \mathcal{S}$ satisfies

$$f = \sum_{\substack{\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N) \\ \text{is an } N\text{-partition}}} \underbrace{\left([x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_N^{\lambda_N}] f \right)}_{\text{the } x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_N^{\lambda_N}\text{-coefficient of } f} m_\lambda.$$

Or, to put this more compactly:

$$f = \sum_{\lambda \text{ is an } N\text{-partition}} \left([x^\lambda] f \right) m_\lambda. \quad (5)$$

(c) Let $n \in \mathbb{N}$. Let

$$\mathcal{S}_n := \{\text{homogeneous symmetric polynomials } f \in \mathcal{P} \text{ of degree } n\}$$

(where the zero polynomial 0 counts as homogeneous of every degree). Then, $\mathcal{S}_n$ is a K -submodule of $\mathcal{S}$.

(d) The family $(m_\lambda)_{\lambda \text{ is an } N\text{-partition}}$ such that $|\lambda|=n$ is a basis of the K -module $\mathcal{S}_n$.

Proof. (b) Let $f \in \mathcal{S}$. We must prove that

$$f = \sum_{\lambda \text{ is an } N\text{-partition}} \left([x^\lambda] f \right) m_\lambda. \quad (6)$$

To do this, it suffices to show that

$$[x^a] f = [x^a] \left(\sum_{\lambda \text{ is an } N\text{-partition}} \left([x^\lambda] f \right) m_\lambda \right) \quad (7)$$

for all $a \in \mathbb{N}^N$ (because if two polynomials agree in all their coefficients, then they must be equal). But f is symmetric (since $f \in \mathcal{S}$), and each $a \in \mathbb{N}^N$ satisfies

$$\begin{aligned} & [x^a] \left(\sum_{\lambda \text{ is an } N\text{-partition}} \left([x^\lambda] f \right) m_\lambda \right) \\ &= \sum_{\lambda \text{ is an } N\text{-partition}} \left([x^\lambda] f \right) \underbrace{[x^a] m_\lambda}_{\substack{= \begin{cases} 1, & \text{if } \text{sort } a = \lambda; \\ 0, & \text{if } \text{sort } a \neq \lambda \end{cases} \\ \text{(by (2))}}} \\ &= \sum_{\lambda \text{ is an } N\text{-partition}} \left([x^\lambda] f \right) \begin{cases} 1, & \text{if } \text{sort } a = \lambda; \\ 0, & \text{if } \text{sort } a \neq \lambda \end{cases} \\ &= [x^{\text{sort } a}] f \quad \left(\begin{array}{l} \text{since all addends of the sum are obviously 0} \\ \text{except for the addend for } \lambda = \text{sort } a \end{array} \right) \\ &= [x^a] f \quad (\text{by Lemma 1.3.6 (b), since } f \text{ is symmetric}). \end{aligned}$$

This proves (7), and thus (6) follows. This completes the proof of part (b).

(a) The equality (5) from part (b) shows that the family $(m_\lambda)_\lambda$ is an N -partition spans the K -module $\mathcal{S}$. It remains to prove that this family is K -linearly independent.

Well, if a K -linear combination $\sum_{\lambda \text{ is an } N\text{-partition}} a_\lambda m_\lambda$ (with coefficients $a_\lambda \in K$) equals 0, then each N -partition μ satisfies

$$\begin{aligned} [x^\mu] \left(\sum_{\lambda \text{ is an } N\text{-partition}} a_\lambda m_\lambda \right) &= \sum_{\lambda \text{ is an } N\text{-partition}} a_\lambda \underbrace{[x^\mu] m_\lambda}_{\substack{= \begin{cases} 1, & \text{if } \mu = \lambda; \\ 0, & \text{if } \mu \neq \lambda \end{cases} \\ \text{(by (3))}}} \\ &= \sum_{\lambda \text{ is an } N\text{-partition}} a_\lambda \begin{cases} 1, & \text{if } \mu = \lambda; \\ 0, & \text{if } \mu \neq \lambda \end{cases} \\ &= a_\mu \end{aligned}$$

and therefore

$$a_\mu = [x^\mu] \underbrace{\left(\sum_{\lambda \text{ is an } N\text{-partition}} a_\lambda m_\lambda \right)}_{=0} = 0;$$

that is, all coefficients a_μ of this linear combination $\sum_{\lambda \text{ is an } N\text{-partition}} a_\lambda m_\lambda$ are 0.

This says precisely that the family $(m_\lambda)_{\lambda \text{ is an } N\text{-partition}}$ is K -linearly independent. The proof of part (a) is thus complete.

(c) This is obvious: a sum of two homogeneous polynomials of degree n is again homogeneous of degree n , and the same applies to scaling.

(d) This is analogous to part (a) but now, instead of the equality (5), we need to show that each $f \in \mathcal{S}_n$ satisfies

$$f = \sum_{\substack{\lambda \text{ is an } N\text{-partition;} \\ |\lambda|=n}} \left([x^\lambda] f \right) m_\lambda. \quad (8)$$

But this equality can be easily derived from (5): Let $\pi_n : \mathcal{P} \rightarrow \mathcal{P}$ be the map that sends each polynomial $f \in \mathcal{P}$ to its degree- n homogeneous component (i.e., removes all monomials of degrees other than n from f). This is a K -linear map that fixes each monomial x^a satisfying $|a| = n$ and annihilates all other monomials.

Now, let $f \in \mathcal{S}_n$. Then, f is symmetric, so that $f \in \mathcal{S}$, and therefore (5) applies. But f is furthermore homogeneous of degree n , so that $\pi_n(f) = f$. On

the other hand, since π_n is a K -linear map, we have

$$\begin{aligned}
 & \pi_n \left(\sum_{\lambda \text{ is an } N\text{-partition}} \left([x^\lambda] f \right) m_\lambda \right) \\
 &= \sum_{\lambda \text{ is an } N\text{-partition}} \left([x^\lambda] f \right) \underbrace{\pi_n(m_\lambda)}_{\substack{= \begin{cases} m_\lambda, & \text{if } |\lambda| = n; \\ 0, & \text{if } |\lambda| \neq n \end{cases} \\ \text{(since } m_\lambda \text{ is homogeneous of degree } |\lambda|)}} \\
 &= \sum_{\lambda \text{ is an } N\text{-partition}} \left([x^\lambda] f \right) \begin{cases} m_\lambda, & \text{if } |\lambda| = n; \\ 0, & \text{if } |\lambda| \neq n \end{cases} \\
 &= \sum_{\substack{\lambda \text{ is an } N\text{-partition}; \\ |\lambda|=n}} \left([x^\lambda] f \right) m_\lambda
 \end{aligned}$$

(here, we have removed all the addends with $|\lambda| \neq n$ from our sum, since all these addends are 0). Using (5), we can rewrite this as

$$\pi_n(f) = \sum_{\substack{\lambda \text{ is an } N\text{-partition}; \\ |\lambda|=n}} \left([x^\lambda] f \right) m_\lambda.$$

In view of $\pi_n(f) = f$, this equality is equivalent to (8). Thus, we have proved (8). The proof of part **(d)** is therefore complete. $\square$

Example 1.3.11. Let $N = 3$. With the usual shorthands x, y, z for the indeterminates x_1, x_2, x_3 , we have

$$\begin{aligned}
 (x+y)(x+z)(y+z) &= \underbrace{x^2y + xy^2 + x^2z + xz^2 + y^2z + yz^2}_{=m_{(2,1,0)}} + 2 \underbrace{xyz}_{=m_{(1,1,1)}} \\
 &= \underbrace{m_{(2,1,0)}}_{=m_{(2,1)}} + 2m_{(1,1,1)}.
 \end{aligned}$$

This can be generalized to higher N 's: expanding $\prod_{i < j} (x_i + x_j)$ in the basis $(m_\lambda)_{\lambda \text{ is an } N\text{-partition}}$ gives a linear combination of m_λ 's whose coefficients count certain directed graphs (see a homework exercise).

1.4. Functoriality

We have been (and will be) working over a fixed commutative ring K . In particular, the polynomial ring $\mathcal{P}$ and the symmetric polynomial ring $\mathcal{S}$ were defined over K (so that the coefficients of polynomials are elements of K). We can also

denote them by $\mathcal{P}_K$ and $\mathcal{S}_K$ to stress their dependence on K . Then we can compare $\mathcal{P}_K$'s and $\mathcal{S}_K$'s for different K 's. The following proposition is pretty obvious, but surprisingly useful:

Proposition 1.4.1. Let K and L be two commutative rings.

- (a) If K is a subring of L , then $\mathcal{P}_K$ becomes a subring of $\mathcal{P}_L$, and $\mathcal{S}_K$ becomes a subring of $\mathcal{S}_L$; moreover, we have $\mathcal{S}_K = \mathcal{S}_L \cap \mathcal{P}_K$ (that is, a symmetric polynomial over K is the same as a symmetric polynomial over L whose coefficients belong to K).
- (b) If $f : K \rightarrow L$ is a ring morphism, then f induces ring morphisms $\mathcal{P}_K \rightarrow \mathcal{P}_L$ (by applying f to each coefficient) and $\mathcal{S}_K \rightarrow \mathcal{S}_L$ (by restricting the morphism $\mathcal{P}_K \rightarrow \mathcal{P}_L$ to $\mathcal{S}_K$).

Proof. LTTR. (All of this is essentially trivial.) □

Example 1.4.2.

- (a) For $K = \mathbb{Z}$ and $L = \mathbb{Q}$, Proposition 1.4.1 (a) is saying that polynomials over $\mathbb{Z}$ are polynomials over $\mathbb{Q}$, and that the symmetric polynomials over $\mathbb{Z}$ are precisely the symmetric polynomials over $\mathbb{Q}$ whose coefficients belong to $\mathbb{Z}$.
- (b) An example of a ring morphism that is not just an inclusion is the projection $\pi : \mathbb{Z} \rightarrow \mathbb{Z}/3$. Thus, Proposition 1.4.1 (b) shows that it induces a ring morphism $\mathcal{P}_{\mathbb{Z}} \rightarrow \mathcal{P}_{\mathbb{Z}/3}$ and a ring morphism $\mathcal{S}_{\mathbb{Z}} \rightarrow \mathcal{S}_{\mathbb{Z}/3}$. These morphisms are known as “reducing all coefficients modulo 3”.

Proposition 1.4.1 is particularly useful because some base rings like $\mathbb{Z}$ and $\mathbb{Q}$ are particularly easy to work with, and then the proposition lets you transfer (some kinds of) knowledge about $\mathcal{S}_{\mathbb{Z}}$ and $\mathcal{S}_{\mathbb{Q}}$ to the more general setting of $\mathcal{S}_K$ for arbitrary K . (Recall that for any ring K , there is a ring morphism $\mathbb{Z} \rightarrow K$.) Much of the theory of symmetric polynomials/functions can be built over $\mathbb{Z}$ or $\mathbb{Q}$, or occasionally over some larger rings such as the polynomial ring $\mathbb{Q}[t]$ or the rational function field $\mathbb{Q}(t)$.

1.5. Schur polynomials

We now come to a subtler and far more interesting family of symmetric polynomials.

1.5.1. Alternants

Here is one way to generate symmetric polynomials:

Example 1.5.1. Let $N = 3$ again. The Vandermonde determinant formula [21s, Theorem 6.4.31 (a)] yields

$$\det \begin{pmatrix} x^2 & x & 1 \\ y^2 & y & 1 \\ z^2 & z & 1 \end{pmatrix} = (x - y)(x - z)(y - z).$$

But what can we say about similar determinants where the powers are “raised”, such as

$$\det \begin{pmatrix} x^5 & x^3 & 1 \\ y^5 & y^3 & 1 \\ z^5 & z^3 & 1 \end{pmatrix} ?$$

We can compute this determinant and factor it, thus obtaining

$$\det \begin{pmatrix} x^5 & x^3 & 1 \\ y^5 & y^3 & 1 \\ z^5 & z^3 & 1 \end{pmatrix} = (x - y)(x - z)(y - z)q, \quad (9)$$

where

$$q = x^2y^3 + x^3y^2 + x^2z^3 + x^3z^2 + y^2z^3 + y^3z^2 \\ + xyz^3 + xy^3z + x^3yz + 2xy^2z^2 + 2x^2yz^2 + 2x^2y^2z.$$

The presence of the three factors $y - z$, $x - z$ and $x - y$ here is not surprising: The determinant in (9) vanishes when we put $y = z$, so it is divisible (as a polynomial over $\mathbb{Z}$) by $y - z$; and similar arguments lead to it being divisible by $x - z$ and $x - y$. Since a polynomial ring over $\mathbb{Z}$ is always a unique factorization domain, this shows that the determinant is divisible by $(x - y)(x - z)(y - z)$. The remaining factor q is easily seen to be a symmetric polynomial in x, y, z (this is easiest to see using Lemma 1.2.9: if we swap x with y , then both the determinant in (9) and the Vandermonde product $(x - y)(x - z)(y - z)$ get multiplied by -1 , so that the quotient q stays unchanged; likewise for any other swap of two variables). But what is more surprising is that this quotient q has nonnegative integer coefficients (it is $m_{(3,2,0)} + m_{(3,1,1)} + 2m_{(2,2,1)}$). Is this a fluke or does it generalize?

Let us try to generalize this:

Definition 1.5.2 (alternants and ρ).

- (a) We let ρ denote the N -tuple $(N-1, N-2, \dots, 1, 0) \in \mathbb{N}^N$. Thus, its j -th entry is $N-j$ for each $j \in [N]$.
- (b) For any N -tuple $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N) \in \mathbb{N}^N$, we define the α -**alternant**

$$a_\alpha := \det \left(\left(x_i^{\alpha_j} \right)_{i,j \in [N]} \right) = \det \begin{pmatrix} x_1^{\alpha_1} & x_1^{\alpha_2} & \cdots & x_1^{\alpha_N} \\ x_2^{\alpha_1} & x_2^{\alpha_2} & \cdots & x_2^{\alpha_N} \\ \vdots & \vdots & \ddots & \vdots \\ x_N^{\alpha_1} & x_N^{\alpha_2} & \cdots & x_N^{\alpha_N} \end{pmatrix} \in \mathcal{P}.$$

For example, $\det \begin{pmatrix} x^5 & x^3 & 1 \\ y^5 & y^3 & 1 \\ z^5 & z^3 & 1 \end{pmatrix}$ is $a_{(5,3,0)}$. Note that

$$\begin{aligned} a_\rho &= \det \left(\left(x_i^{\rho_j} \right)_{i,j \in [N]} \right) && \text{(writing } \rho \text{ as } \rho = (\rho_1, \rho_2, \dots, \rho_N)) \\ &= \det \left(\left(x_i^{N-j} \right)_{i,j \in [N]} \right) && \text{(since } \rho_j = N-j \text{ for all } j \in [N]) \\ &= \prod_{i < j} (x_i - x_j) && \left(\text{where } \prod_{i < j} \text{ means } \prod_{1 \leq i < j \leq N} \right) \end{aligned}$$

(by the Vandermonde determinant formula [21s, Theorem 6.4.31 (a)]). Now we suspect:

Conjecture 1.5.3. For every $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N) \in \mathbb{N}^N$ satisfying $\alpha_1 > \alpha_2 > \cdots > \alpha_N$, the alternant a_α can be expressed as a_ρ times a symmetric polynomial with nonnegative integer coefficients.

Note that the assumption $\alpha_1 > \alpha_2 > \cdots > \alpha_N$ is not much of a restriction. Indeed, if two of the N entries $\alpha_1, \alpha_2, \dots, \alpha_N$ are equal, then a_α is easily seen to be 0 (since a determinant with two equal columns is 0); in all remaining cases, we can sort the N entries $\alpha_1, \alpha_2, \dots, \alpha_N$ in strictly decreasing order by repeatedly swapping them, and each swap multiplies a_α by -1 . Thus, if we did not require $\alpha_1 > \alpha_2 > \cdots > \alpha_N$ in Conjecture 1.5.3, then we would have to replace “nonnegative” by “nonnegative or nonpositive, depending on the sign of the permutation that sorts the N numbers $\alpha_1, \alpha_2, \dots, \alpha_N$ in decreasing order”.

The “nonnegative” part of Conjecture 1.5.3 is far from easy. Try it! It was first proved by Carl Kostka in an 1882 paper [Kostka82] (relying on an older paper of his [Kostka77])¹ by finding a combinatorial interpretation of the quotient

¹I find it peculiar that Kostka produced very few papers apart from these; he essentially founded a field of mathematics and left.

a_α/a_ρ . This is precisely what we shall do, though we will instead prove it using a modern approach by Stembridge (2002, [Stembr02]). The quotients a_α/a_ρ in Conjecture 1.5.3 are known as the **Schur polynomials**, and are perhaps the most interesting of the classical bases of $\mathcal{S}$.

Example 1.5.4. If $\alpha = \rho$, then the quotient a_α/a_ρ is simply 1. Boring. The simplest case beyond that is the case when

$$\alpha = (N, N-2, N-3, \dots, 0).$$

In this case, we have

$$\begin{aligned} a_\alpha &= \det \begin{pmatrix} x_1^N & x_1^{N-2} & x_1^{N-3} & \dots & x_1^0 \\ x_2^N & x_2^{N-2} & x_2^{N-3} & \dots & x_2^0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_N^N & x_N^{N-2} & x_N^{N-3} & \dots & x_N^0 \end{pmatrix} \\ &\quad \left(\begin{array}{l} \text{this is the Vandermonde determinant except that} \\ \text{its first column has exponent } N \text{ instead of } N-1 \end{array} \right) \\ &= \prod_{i < j} (x_i - x_j) \cdot \underbrace{(\text{some homogeneous symmetric polynomial of degree 1})}_{\text{i.e., a scalar multiple of } x_1 + x_2 + \dots + x_N} \\ &\quad \left(\begin{array}{l} \text{using an argument similar to the one in} \\ \text{Example 1.5.1 and using degree reasons} \\ \text{to nail down the degree – exercise!} \end{array} \right) \\ &= \text{a scalar multiple of } \prod_{i < j} (x_i - x_j) \cdot (x_1 + x_2 + \dots + x_N). \end{aligned}$$

By comparing coefficients of the monomial $x_1^N x_2^{N-2} x_3^{N-3} \dots x_N^0$ on both sides, we can easily see that the scalar on the right-hand side must be 1, and thus we obtain

$$a_\alpha = \prod_{i < j} (x_i - x_j) \cdot (x_1 + x_2 + \dots + x_N)$$

in this case.

What is the next interesting case?

References

- [21s] Darij Grinberg, *An Introduction to Algebraic Combinatorics (Math 531, Winter 2024 lecture notes)*, 22 July 2026.
<http://www.cip.ifi.lmu.de/~grinberg/t/21s/lecs.pdf>
- [Kostka77] Carl Kostka, *Ueber Borchardts Function*, *Journal für die reine und angewandte Mathematik* **82** (1877), pp. 212–229.

- [Kostka82] Carl Kostka, *Ueber den Zusammenhang zwischen einigen Formen von symmetrischen Functionen*, Journal für die reine und angewandte Mathematik **93** (1882), pp. 89–123.
- [Stembr02] John R. Stembridge, *A Concise Proof of the Littlewood-Richardson Rule*, Electronic Journal of Combinatorics **9** (2002), #N5
-