Symmetric ideals, Specht polynomials and solutions to symmetric systems of equations

Preprint · December 2019

Cordian Riener
UiT The Arctic University of Norway

SYMMETRIC IDEALS, SPECHT POLYNOMIALS AND SOLUTIONS TO SYMMETRIC SYSTEMS OF EQUATIONS

PHILIPPE MOUSTROU, CORDIAN RIENER, AND HUGUES VERDURE

Abstract. An ideal of polynomials is symmetric if it is closed under permutations of variables. We relate general symmetric ideals to the so called Specht ideals generated by all Specht polynomials of a given shape. We show a connection between the leading monomials of polynomials in the ideal and the Specht polynomials contained in the ideal. This provides applications in several contexts. Most notably, this connection gives information about the solutions of the corresponding set of equations. From another perspective, it restricts the isotypic decomposition of the ideal viewed as a representation of the symmetric group.

1. Introduction

Let S_n denote the symmetric group on n-elements, and let $\mathbb{K}$ be a field. Then S_n acts naturally on an n-dimensional space $\mathbb{K}^n$ by permuting coordinates. This linear action then gives rise to an action on the corresponding polynomial ring by permuting coordinates. In this article we consider ideals $I \subset \mathbb{K}[X_1, \ldots X_n]$ which are stable under this action. The study of such ideals appears quite naturally in different contexts (see for example [21, 6, 10, 4]).

Our interest in such ideals stems from algorithmic purposes: the symmetry on a set of equations often can be used to simplify its resolution. In this flavour, a fundamental result by Timofte (see [23, 16]) yields that every symmetric variety defined by polynomials of degree d is non-empty over $\mathbb{R}$ if and only if it contains a real point with at most d distinct coordinates. Here we aim at generalising this aspect of Timofte's result in various ways. We are able to show that - under the assumption that the number of variables is sufficiently large - the variety corresponding to the symmetric ideal will not contain any point with strictly more than d distinct coordinates. Moreover, our result yields that the set of possible configurations of these d distinct coordinates can be further restricted by the shape of the monomials of highest degree amongst the generators of I, see Section 5.1. The arguments put forward to establish this result are purely algebraic and work with no requirements on $\mathbb{K}$. In fact, this result will follow from a study of Specht polynomials contained in symmetric ideals. More precisely, we assign a partition of n to these monomials and we relate these partitions to specific Specht polynomials which belong to the ideal, see Section 4.

In characteristic 0, this property also has an application to the structure of the decomposition of I in terms of S_n-representations. The action of S_n on I turns I and $\mathbb{K}[X_1, \ldots, X_N]/I$ into $\mathbb{K}[S_n]$ modules. Over a field of characteristic zero this modules can be decomposed into irreducible $\mathbb{K}[S_n]$ modules, which are usually called Specht modules and we show in Section 5.2 that the possible Specht modules appearing in this decomposition are also very restricted (see [1] for a result

---

Key words and phrases. Symmetric group, Specht polynomials, polynomial equations.

This work has been supported by European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement 813211 (POEMA) and the Tromsø Research foundation grant agreement 17matteCR. .

in a similar spirit in the real setting). One application of such a decomposition concerns sums of squares representations of positive symmetric polynomials modulo symmetric ideals. The understanding of the irreducible representations in I gives a control on the complexity of sums of squares decomposition in this setup, see Section 5.3.

This article is structured as follows. Section 2 collects necessary standard notations and definitions. Then, Section 3 focuses on varieties defined by Specht polynomials and their properties. This is used in Section 4 to describe the Specht polynomials contained in symmetric ideals. Finally, Section 5 is devoted to applications.

2. PRELIMINARY NOTATIONS AND DEFINITIONS

Partitions and Young tableaux. For any natural number n, one can consider its *partitions*:

Definition 1. *Let $n \in \mathbb{N}$. A partition λ of n, (denoted $\lambda \vdash n$) is a sequence $\lambda = (\lambda_1, \cdots, \lambda_l)$ of positive natural numbers ordered such that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_l \geqslant 0$ with the property $\lambda_1 + \ldots + \lambda_l = n$. The length of λ is*

$$\operatorname{len}(\lambda) = \max\{i : \lambda_i \neq 0\}.$$

We will allow ourselves to identify partitions that only differ by 0 terms.

Definition 2. *For a given partition λ, its dual partition $\lambda^\perp$ is defined by*

$$(\lambda^\perp)_i = |\{j, \ \lambda_j \geqslant i\}|.$$

Partitions are very well known and are closely related to Young tableaux (see *e.g.* [18]).

Definition 3. *Given $\lambda \vdash n$, a* Young tableau *of shape $\lambda \vdash n$, or a λ-*tableau *consists of* $\operatorname{len}(\lambda)$ *rows, with λ_i entries in the i-th row. Each entry is an element in $\{1, \ldots, n\}$, and each of these numbers occurs exactly once.*

Definition 4. *Let $n \in \mathbb{N}$. Let $\lambda = (\lambda_1, \cdots, \lambda_l) \vdash n$ and $\mu = (\mu_1, \cdots, \mu_m) \vdash n$ be two partitions. We say that λ dominates μ if*

$$\sum_{j=1}^{i} \lambda_j \geqslant \sum_{j=1}^{i} \mu_j \quad \textit{holds for all } 1 \leqslant i \leqslant \min\{\operatorname{len}(\lambda), \operatorname{len}(\mu)\}.$$

We will write $\lambda \trianglerighteq \mu$ in this case.

Equipped with the dominance order, the set of all partitions of a given $n \in \mathbb{N}$ is a partially ordered set. We will further use $\lambda \ntrianglerighteq \nu$ to denote the case in which λ does not dominate ν. Note that the order is only partial and hence this does not entail that ν dominates λ, since they also might be not comparable. Furthermore, note that $(\lambda^\perp)_1 = \operatorname{len}(\lambda)$, $(\lambda^\perp)^\perp = \lambda$, and

$$\lambda \trianglerighteq \mu \Leftrightarrow \mu^\perp \trianglerighteq \lambda^\perp.$$

Orbit types and partitions. The action of S_n on $\mathbb{K}^n$ naturally decomposes the space into orbits.

Definition 5. *Let $x \in \mathbb{K}^n$. The orbit type of x is the conjugacy class of the associated stabilizer subgroup* $\operatorname{Stab}(x) \subseteq S_n$*. Since, up to conjugation,*

$$\operatorname{Stab}(x) = S_{\ell_1} \times S_{\ell_2} \times \cdots \times S_{\ell_k}$$

with $\ell_1 \geqslant \ell_2 \geqslant \ldots \geqslant \ell_k$, the orbit type of x is characterized by $\Lambda(x) := (\ell_1, \ell_2, \cdots, \ell_k)$. Then, for a given $\lambda \vdash n$ we can define

$$H_\lambda := \{x \in \mathbb{K}^n : \Lambda(x) = \lambda\}.$$

Remark 1. *Note that we have*

$$\mathbb{K}^n = \biguplus_{\lambda \vdash n} H_\lambda .$$

Polynomials, varieties, and symmetric ideals. Let $\mathbb{K}$ be a field, and consider the polynomial ring in n variables $\mathbb{K}[X_1, \cdots, X_n]$. For any $P \in \mathbb{K}[X_1, \cdots, X_n]$, we denote by $\mathrm{Mon}(P)$ the set of monomials appearing in P, and by P_h its homogenous component of degree h.

Given a monomial

$$m = \prod_{j=1}^{n} X_j^{k_j},$$

its *support* is

$$\mathrm{Supp}(m) = \{j, \quad k_j \neq 0\}$$

and its *weight* $\mathrm{wt}(m)$ is the cardinality of its support. By taking the union over all the monomials in $\mathrm{Mon}(P)$, we generalize these notions to P to define $\mathrm{Supp}(P)$ and $\mathrm{wt}(P)$.

To every ideal in $\mathbb{K}[X_1, \ldots, X_n]$ one can associate a variety:

Definition 6. *Let I be an ideal in $\mathbb{K}[X_1, \ldots, X_n]$. The* variety $V(I)$ *associated with I is the subset of $\mathbb{K}^n$ made by the common zeros of all the polynomials in I, namely*

$$V(I) = \{x \in \mathbb{K}^n, \quad P(x) = 0 \text{ for every } P \in I\}.$$

The action of S_n on $\mathbb{K}_n$ induces an action on $\mathbb{K}[X_1, \ldots, X_n]$ by permuting the variables. We denote by σP the image of a polynomial P under the action of a permutation σ.

Definition 7. *An ideal $I \subset \mathbb{K}[X_1, \ldots, X_n]$ is called a* symmetric ideal *if for every P in I and for every $\sigma \in S_n$, σP belongs to I.*

Note that if I is a symmetric ideal, then the variety $V(I)$ is closed by the action of S_n on the coordinates.

Specht Polynomials. The so-called *Specht polynomials* will play a central role in our proofs. Those polynomials were originally designed by Specht [20] to construct the different irreducible representations of S_n.

Definition 8. *Let $n \in \mathbb{N}$.*

1. *For a set $S = \{i_1, \ldots, i_r\} \subset \{1, \ldots, n\}$, we define the* Vandermonde determinant $\Delta(S)$ *of the variables X_i, for $i \in S$:*

$$\Delta(S) \coloneqq \prod_{1 \leqslant j < k \leqslant r} (X_{i_j} - X_{i_k}).$$

2. *Let $\lambda \vdash n$ and T be a λ-tableau. Then the* Specht polynomial *associated with T is the polynomial*

$$\mathrm{sp}_T \coloneqq \prod_c \Delta(T_{\cdot,c}),$$

where c runs through the columns of T, and $T_{\cdot,c}$ denotes the entries in the cth column. We will say that sp_T is a Specht polynomial of shape λ.

Example 1. *The Specht polynomial associated with the Young tableau*

$$\begin{array}{|c|c|c|c|} \hline 4 & 2 & 6 & 1 \\ \hline 8 & 7 & 5 \\ \cline{1-3} 3 \\ \cline{1-1} \end{array}$$

is

$$(X_4 - X_8)(X_4 - X_3)(X_8 - X_3)(X_2 - X_7)(X_6 - X_5).$$

3. Zeros of Specht polynomials

Throughout the paper, we will be interested in the ideal generated by all the Specht polynomials of a given shape, as well as in the corresponding variety.

Definition 9. *Let $\mathbb{K}$ be a field, n be an integer, and $\mu \vdash n$.*

- *The* μ-Specht ideal, *denoted by I_μ^{sp}, is the ideal of $\mathbb{K}[X_1,\dots,X_n]$ generated by all the Specht polynomials of shape μ.*
- *We denote by V_μ the set of common zeros of all Specht polynomials of shape μ, that is*
$$V_\mu := V(I_\mu^{\mathrm{sp}}) = \bigcap_{\mathrm{sh}(T)=\mu} V(\mathrm{sp}_T).$$

We aim at describing more precisely those varieties. Particular cases of these varieties have been studied for example in [25, 24, 7]. Let us start with a few remarks. Let T be a Young tableau, of shape $\mu \vdash n$. A point $x = (x_1,\dots,x_n)$ is a zero of sp_T if and only if there exists a column of T containing two indexes $i \neq j$ such that $x_i = x_j$. Following this easy observation, we get a characterization of V_μ that will be useful later: a point $x \in \mathbb{K}^n$ *does not* belong to V_μ if and only if one can fill in a Young tableau of shape μ with the coordinates of x in such a way that in every column, all the values are distinct.

Now we want to prove that V_λ contains all the V_μ such that $\mu \trianglerighteq \lambda$. This is a consequence of the corresponding inclusion of ideals:

Theorem 1. *Let $\lambda \vdash n$. Then for any $\mu \vdash n$ such that $\mu \trianglerighteq \lambda$,*
$$I_\lambda^{\mathrm{sp}} \subset I_\mu^{\mathrm{sp}}.$$
In particular
$$V_\mu \subset V_\lambda.$$

Remark 2. *Note that the inclusion of ideals is a little bit stronger statement than the inclusion of varieties, since it is not known whether Specht ideals are radical, see* [25].

Proof. Let $\lambda = (\lambda_1,\dots,\lambda_t)$. We only need to consider the particular case where μ is of the form
$$\mu = (\lambda_1,\dots,\lambda_{i-1},\lambda_i+1,\lambda_{i+1},\dots,\lambda_{j-1},\lambda_j-1,\lambda_{j+1},\dots,\lambda_t),$$
where $\lambda_{i-1} > \lambda_i$ and $\lambda_j > \lambda_{j+1}$, in order to ensure that μ is a partition. Indeed, it is known that we can go from λ to any $\mu \trianglerighteq \lambda$ by a finite number of such steps, see e.g. [3, Prop. 2.3].

Let T be a Young tableau of shape λ. We need to show that sp_T belongs to the ideal generated by all the sp_U with $\mathrm{sh}(U) = \mu$. If U is a Young Tableau of shape μ, then its columns have the same number of elements than T, except for two of them. Let U_1 and U_2 be these columns, and let T_1 and T_2 be the corresponding columns in T. If $a = |T_1|$ and $b = |T_2|$, then $|U_1| = a-1$, $|U_2| = b+1$ and $a-b \geqslant 2$. By restricting our attention to these two columns, it is enough to prove, up to permutation, that the polynomial
$$P = \Delta(\{1,\dots,a\})\Delta(\{a+1,\dots,a+b\})$$
is in the ideal I of $\mathbb{K}[X_1,\dots,X_{a+b}]$ generated by all the polynomials of the form
$$\Delta(S_1)\Delta(S_2),\ \text{with } |S_1| = a-1, |S_2| = b+1,\ \text{and } S_1 \cup S_2 = \{1,\dots,a+b\}.$$
Consider the polynomial
$$Q = \Delta(\{2,\dots,a\})\Delta(\{1,a+1,\dots,a+b\})X_1^{a-b-1},$$

and the polynomial

$$R = \sum_{i=1}^{a} \epsilon((1,i))(1,i)Q,$$

where ϵ denotes the signature. By construction, R is in the ideal I. We need to prove that for any pair $1 \leqslant \alpha < \beta \leqslant a$ and $a+1 \leqslant \alpha < \beta \leqslant a+b$,

$$R_{\alpha,\beta} = R(X_1 \dots, X_{\beta-1}, X_\alpha, X_{\beta+1}, \dots, X_n) = 0.$$

The latter case is obvious by definition of Q and R. Suppose $1 \leqslant \alpha < \beta \leqslant a$. Then for every $i \neq \alpha, \beta$ we have $((1,i)Q)_{\alpha,\beta} = 0$, so that

$$R_{\alpha,\beta} = \epsilon((1,\alpha))((1,\alpha)Q)_{\alpha,\beta} + \epsilon((1,\beta))((1,\beta)Q)_{\alpha,\beta}.$$

Now, we have do distinguish two cases, namely $\alpha = 1$ and $\alpha \neq 1$. In the first case, we have

$$R_{1,\beta} = Q_{1,\beta} - ((1,\beta)Q)_{1,\beta} = 0$$

while in the second case, the signatures are both negative, but the polynomial differ from each other by interchanging the variables X_1 and X_α in places α and β, and by definition of Q, these polynomials are opposite of each other, that is,

$$R_{\alpha,\beta} = 0.$$

This implies that $P = \Delta(\{1, \dots, a\})\Delta(\{a+1, \dots, a+b\})$ divides R. Since the degree of R satisfies

$$\begin{aligned} \deg(R) \leqslant \deg(Q) &= \frac{(a-1)(a-2)}{2} + \frac{(b+1)b}{2} + a - b - 1 \\ &= \frac{a(a-1)}{2} + \frac{b(b-1)}{2} \\ &= \deg(P), \end{aligned}$$

it implies that

$$R = cP$$

with $c \in \mathbb{K}$, and we need to check that c is not zero. The degree of Q in the variable X_1 is $b + a - b - 1 = a - 1$ while its degree in the variable X_i for $2 \leqslant i \leqslant a$ is $a - 2$. Thus, the degree of R in the variable X_1 is exactly $a - 1$ and the corresponding coefficient is $\Delta(\{2, \dots, a\})\Delta(\{a+1, \dots, a+b\})$, which is also the coefficient of X_1^{a-b} in P. Hence $R = P$.

□

Finally, we get a characterization of V_μ in terms of orbit types:

Theorem 2. *The set of common zeros of Specht polynomials associated with Young tableaux of given shape μ can be characterised as*

$$V_\mu = \left(\bigcup_{\lambda \trianglelefteq \mu} H_\lambda \right)^c = \bigcup_{\nu \ntrianglelefteq \mu} H_\nu$$

Proof. We want to show $V_\mu^c = \bigcup_{\lambda \trianglelefteq \mu} H_\lambda$.

Take x not in V_μ. This means that there exists a Young tableau T of shape μ such that $\mathrm{sp}_T(x) \neq 0$. Let $\lambda = \Lambda(x) = (\ell_1, \dots, \ell_r)$. Let $u_1, \dots, u_r$ be the distinct coordinates of x, where for each $1 \leqslant i \leqslant rr$, u_i appears l_i times. Since $\mathrm{sp}_T(x) \neq 0$, we can fill T with $u_1, \dots, u_r$ in such a way that for every $1 \leqslant i \leqslant r$, u_i appears at most once per column, and we may assume without loss of generality that in every column, if u_j is below u_i, then $i < j$. As a consequence, for every $1 \leqslant k \leqslant r$, the u_i for $i \leqslant k$ are contained in the first k rows of T. This means that

$$\ell_1 + \dots + \ell_k \leqslant \mu_1 + \dots + \mu_k,$$

which implies $\lambda \trianglelefteq \mu$.

Conversely, let $x \in \mathbb{K}^n$, let $\lambda = \Lambda(x) = (\ell_1, \dots, \ell_r)$, and let $u_1, \dots, u_r$ be the distinct coordinates of x, where for each $1 \leqslant i \leqslant r$, u_i appears l_i times. Let T be any Young tableau of shape λ such that the indexes in the ith row are the indexes j such that $x_j = u_i$. Then $\mathrm{sp}_T(x) \neq 0$, and hence x is outside V_λ. As a consequence, if $\lambda \trianglelefteq \mu$, then according to Theorem 1, $V_\mu \subset V_\lambda$, so that $x \in V_\mu^c$.

□

4. Specht polynomials in symmetric ideals

In this section we show that if a symmetric ideal contains polynomials with sparse leading component, then this ideal will contain many Specht polynomials. Let us be more precise: First, to every monomial we associate a partition:

Definition 10. *Let be a monomial of weight l and degree d in $\mathbb{K}[X_1, \dots, X_n]$. The partial degrees of m induce a partition of d of length l, say $(\lambda_1, \dots, \lambda_l)$.*

If moreover we assume that $l + d \leqslant n$, we can define a partition $\mu(m)$ of n by

$$\mu(m) = (\lambda_1 + 1, \lambda_2 + 1, \dots, \lambda_l + 1, \underbrace{1, \dots, 1}_{n-d-l}).$$

Example 2. *Let $n = 12$, and*

$$m = X_2 X_4^4 X_5^2.$$

Then

$$\mu(m) = (5, 3, 2, 1, 1).$$

Then we will show:

Theorem 3. *Let $I \subset \mathbb{K}[X_1, \dots, X_n]$ be a symmetric ideal. Assume that there exists $P \in I$ of degree d, such that $d + \mathrm{wt}(P_d) \leqslant n$. Then, for every monomial $m \in \mathrm{Mon}(P_d)$, the ideal I contains every sp_T for which $\mathrm{sh}(T) \trianglelefteq \mu(m)^\perp$.*

According to Theorem 1, it is enough to prove that I contains the Specht polynomials of shape $\mu(m)^\perp$. Hence we only need to prove:

Theorem 4. *Let $I \subset \mathbb{K}[X_1, \dots, X_n]$ be a symmetric ideal. Assume that there exists $P \in I$ of degree d, such that $d + \mathrm{wt}(P_d) \leqslant n$. Then, for every monomial $m \in \mathrm{Mon}(P_d)$, the ideal I contains every sp_T for which $\mathrm{sh}(T) = \mu(m)^\perp$.*

In the following proof, we will assume that the characteristic of $\mathbb{K}$ is zero. This allows for more conceptual proof. We provide a general proof in the Appendix.

Proof. Let $\mathbb{K}$ be a field of characteristic 0. Since the ideal I is symmetric, we may assume that

$$m = X_1^{k_1} X_2^{k_2} \cdots X_l^{k_l}.$$

and that $\mathrm{Supp}(P_d) = \{1, \dots, \mathrm{wt}(P_d)\}$. Its associated partition is

$$\mu := \mu(m) = (k_1 + 1, k_2 + 1, \dots, k_l + 1, \underbrace{1, \dots, 1}_{n-d-l}).$$

The statement says that the ideal I contains any Specht polynomial of the form

$$\Delta_1 \Delta_2 \cdots \Delta_l,$$

where the Δ_i are Vandermonde polynomials in disjoint sets of variables, each of size $k_i + 1 = \mu_i$. Thanks to its symmetry, it is enough to show that I contains one such polynomial.

Our strategy consists in using $X_i^{k_i}$ to build a Vandermonde polynomial involving X_i and k_i variables that do not appear in P_d. Our assumption on P_d guarantees that there are enough free variables to do so.

More precisely, we can take $I_1, \dots, I_l$, disjoint subsets of $\{1, \dots, n\}$ such that for any $1 \leqslant i \leqslant l$, there are k_i elements in I_i, and none of them appears in P_d. Let, for $1 \leqslant i \leqslant l$,

$$J_i = \{i\} \cup I_i.$$

We will show that there exist polynomials $R_\sigma \in \mathbb{K}[X_1, \dots, X_n]$, for $\sigma \in S_n$ such that:

$$\Delta(J_1) \cdots \Delta(J_l) = \sum_{\sigma \in S_n} R_\sigma \sigma P.$$

Here, applying the strategy used to prove Theorem 1 we give explicit polynomials R_σ when the characteristic of $\mathbb{K}$ is 0. In the general case, we can give a recursive construction of these polynomials; we postpone this construction to Appendix A. Consider the polynomials

$$Q = \Delta(I_1) \cdots \Delta(I_l) P$$

and

$$R = \sum_{\sigma \in S_{J_1} \times \cdots \times S_{J_l}} \epsilon(\sigma) \sigma Q,$$

where $S_{J_1} \times \cdots \times S_{J_l}$ is seen as a subgroup of S_n. By construction, for every $\rho \in S_{J_1} \times \cdots \times S_{J_l}$,

$$\rho R = \epsilon(\rho) R$$

so that $\Delta(J_1) \cdots \Delta(J_l)$ divides R. Furthermore, since

$$\begin{aligned} \deg(Q) &= d + \sum_{i=1}^{l} \frac{k_i(k_i - 1)}{2} \\ &= \sum_{i=1}^{l} \frac{k_i(k_i - 1)}{2} + k_i \\ &= \sum_{i=1}^{l} \frac{k_i(k_i + 1)}{2} \\ &= \deg(\Delta(J_1) \cdots \Delta(J_l)), \end{aligned}$$

we get

$$R = c \Delta(J_1) \cdots \Delta(J_l)$$

with $c \in \mathbb{K}$. In order to check that c is not 0, we look at the coefficent of R corresponding with $m = X_1^{k_1} \cdots X_l^{k_l}$, seen as an element of $(\mathbb{K}[X_{l+1}, \dots, X_n])[X_1, \dots, X_l]$.

In Q, this coefficient is $\Delta(I_1) \cdots \Delta(I_l)$. If, for $1 \leqslant i \leqslant l$, the permutation $\sigma \in S_{J_1} \times \cdots \times S_{J_l}$ does not let i invariant, the assumption on P_d ensures that the coefficient of m in σQ will be 0. Therefore, the coefficient of R corresponding with m is

$$\sum_{\sigma \in S_{I_1} \times \cdots \times S_{I_l}} \epsilon(\sigma) \sigma \Delta(I_1) \cdots \Delta(I_l) = k_1! \cdots k_l! \Delta(I_1) \cdots \Delta(I_l)$$

and hence $R = k_1! \cdots k_l! \Delta(J_1) \cdots \Delta(J_l)$.

□

5. APPLICATIONS

5.1. **Computing points in symmetric varieties.** Let n be an integer and I be a symmetric ideal in $\mathbb{K}[X_1, \dots, X_n]$. What can we say about the variety $V(I)$? For instance, can we algorithmically decide if $V(I)$ is empty in an efficient manner making use of the structure of I?

Over any real closed field $\mathbb{R}$ the so-called *half-degree principle* [23] can be used to simplify the algorithmical task of root finding. This statement says ([16, Corollary 1.3]):

Theorem 5. *Let $\mathbb{K}$ be a real closed field, and let P be a symmetric polynomial in $\mathbb{K}[X_1,\dots,X_n]$ of degree d, and let $k=\max(2,\lfloor\frac{d}{2}\rfloor)$. Then there exists $x\in\mathbb{K}^n$ such that $P(x)=0$ if and only if there exists $y\in\mathbb{K}^n$ with at most k distinct coordinates such that $P(y)=0$.*

This implies the following result on symmetric ideals:

Corollary 1. *Let $\mathbb{K}$ be a real closed field, and let I be a symmetric ideal of $\mathbb{K}[X_1,\dots,X_n]$, generated by $P_1,\dots,P_l$. Let $d=\max(\deg(P_1),\dots,\deg(P_l))$. Then $V(I)$ is non empty if and only if it contains a point $x\in\mathbb{K}^n$ with at most d distinct coordinates.*

Proof. Over a real closed field the variety $V(I)$ is exactly the variety defined by

$$Q=\sum_{i=1}^{l}\sum_{\sigma\in S_n}\sigma(P_i)^2$$

and we can apply Theorem 5. □

The algorithmic implications of this result are the following. Consider $\mu\vdash n$ of length d. For every polynomial $P\in\mathbb{K}[X_1,\dots,X_n]$ we consider

$$P^{\mu}:=P(\underbrace{Z_1,\dots,Z_1}_{\mu_1},\underbrace{Z_2,\dots,Z_2}_{\mu_2},\dots,\underbrace{Z_d,\dots,Z_d}_{\mu_d})$$

and denote by $I^{\mu}\subset\mathbb{K}[Z_1,\dots,Z_d]$ the resulting ideal in d variables. In other words, consider the topological closure $\overline{H}_{\mu}$ of H_{μ} and the map

$$\Phi_{\mu}:V(I^{\mu})\longrightarrow(V(I)\cap\overline{H}_{\mu})/S_n$$

which associates to a point $x=(x_1,\dots,x_d)\in V(I^{\mu})$ the S_n-orbit of the point

$$x=(\underbrace{x_1,\dots,x_1}_{\mu_1},\underbrace{x_2,\dots,x_2}_{\mu_2},\dots,\underbrace{x_d,\dots,x_d}_{\mu_d}).$$

This map is clearly surjective and thus Corollary 1 says precisely that $V(I)$ is empty if and only if $V(I^{\mu})$ is empty for every μ partition of n with $\operatorname{len}(\mu)\leqslant d$. Since the number of d-partitions of n is bounded by $(n+1)^d$ the original problem in n variables reduces to a polynomial number of problems in d variables.

Our result yields a stronger version of this principle, under additional assumption on the support of the polynomials: On the one hand our results are valid for any field, and on the other hand, not only our varieties contain points with few distinct coordinates, but they contain *only* points with few distinct coordinates. More precisely, Theorem 3 gives, in this context:

Theorem 6. *Let $I\subset\mathbb{K}[X_1,\dots X_n]$ be a symmetric ideal. Assume that there exists $P\in I$ of degree d such that $\operatorname{wt}(P_d)+d\leqslant n$ and let $m\in Mon(P_d)$. Then*

$$V(I)\cap H_{\lambda}=\varnothing\ \text{for all}\ \lambda\trianglelefteq\mu(m)^{\perp}.$$

In other words,

$$V(I)/S_n=\bigcup_{\nu\ntrianglelefteq\mu(m)^{\perp}}\Phi_{\nu}(V(I^{\nu})).$$

Hence, if one is able to compute the points in the variety $V(I^{\nu})$, one gets all the points of $V(I)$. Also note that the length of the partitions ν is at most d, this comes from the following observation:

Lemma 1. *Let n be an integer, and m be a monomial of degree d, with $d+\operatorname{wt}(m)\leqslant n$. Then for every λ partition of n such that $\operatorname{len}(\lambda)>d$,*

$$\mu(m)^{\perp}\trianglerighteq\lambda.$$

Proof. The proof consists in two steps. First we prove that

$$\mu(m)^{\perp} \trianglerighteq (n-d, \underbrace{1,\dots,1}_{d}).$$

Indeed, if $m = X_1^{k_1} \cdots X_l^{k_l}$,

$$\mu(m) = (k_1+1, k_2+1, \dots, k_l+1, \underbrace{1,\dots,1}_{n-d-l})$$

has length $n-d$, so that

$$\mu(m)^{\perp} = (n-d, \dots) \trianglerighteq (n-d, \underbrace{1,\dots,1}_{d}).$$

Second, if $\operatorname{len}(\lambda) > d$, then

$$(n-d, \underbrace{1,\dots,1}_{d}) \trianglerighteq \lambda.$$

Indeed, if not, there exists j such that

$$\sum_{i=1}^{j} \lambda_i > (n-d) + (j-1).$$

In this case, since $\operatorname{len}(\lambda) \geqslant d+1$,

$$\begin{aligned} n = \sum_{i=1}^{\operatorname{len}(\lambda)} \lambda_i &\geqslant \sum_{i=1}^{j} \lambda_i + \operatorname{len}(\lambda) - j \\ &> n - d + j - 1 + \operatorname{len}(\lambda) - j \\ &> n. \end{aligned}$$

□

Remark 3. *The above Lemma shows that we can always ensure that every point of the variety $V(I)$ has at most d distinct points. If the monomial m is X_i^d then $\mu(m)^{\perp} = (n-d, 1, \dots, 1)$ and we need to consider all d-partitions of n. However, depending on the actual monomials of highest degree in the generators of I a finer analysis can even reduce the number of partitions that need to be considered. Hence we get a stronger version of the degree principle.*

One further natural consequence is that our variety is contained in a union of d-dimensional subspaces, hence:

Corollary 2. *Let $I \subset \mathbb{K}[X_1, \dots, X_n]$ be a symmetric ideal. Assume that there exists $P \in I$ of degree d, such that $d + \operatorname{wt}(P_d) \leqslant n$. Then the dimension of the variety $V(I)$ is at most d.*

In a more general setup, Nagel and Römer [13] study sequences of symmetric ideals. They show in particular that the dimension of the ideals they study is a linear function in n. In our more restricted framework, we thus obtain a stabilization of the dimension of such sequences.

5.2. **Isotypic components of symmetric ideals.** The action of the symmetric group S_n on $\mathbb{K}[X_1, \dots, X_n]$ is linear, giving the polynomial ring the structure of a $\mathbb{K}[S_n]$-module. If we assume that the characteristic of $\mathbb{K}$ is 0, then every $\mathbb{K}[S_n]$-module can be decomposed as a direct sum of irreducible submodules. It is well known (see *e.g.* [18]) that the irreducible $\mathbb{K}[S_n]$-modules are in correspondence with the partitions of n. These modules are called Specht modules, denoted by S^{λ}. It follows that every $\mathbb{K}[S_n]$-module U can be uniquely written as

$$U \simeq \bigoplus_{\lambda \vdash n} U_{\lambda},$$

where for every partition λ of n, U_λ is a direct sum of irreducible submodules isomorphic to S^λ. The submodule U_λ is called the λ-*isotypic component* of U.

Now let $I \subset \mathbb{K}[X_1, \ldots, X_n]$ be a symmetric ideal. It is also a $\mathbb{K}[S_n]$-module, and we have, for every partition λ of n,

$$I_\lambda = \mathbb{K}[X_1, \ldots, X_n]_\lambda \cap I.$$

Let λ be a partition of n, then the linear subspace W_λ of $\mathbb{K}[X_1, \ldots, X_n]$ generated by all the Specht polynomials of shape λ is an irreducible submodule of $\mathbb{K}[X_1, \ldots, X_n]_\lambda$ isomorphic to S^λ. For any other irreducible submodule $\widetilde{W}_\lambda$ in $\mathbb{K}[X_1, \ldots, X_n]_\lambda$, we have an isomorphism φ between W_λ and $\widetilde{W}_\lambda$. Let T be a Young tableau of shape λ. Since φ respects the action of S_n, for any τ transposition of two elements in a same column of T,

$$\tau\varphi(\mathrm{sp}_T) = -\varphi(\mathrm{sp}_T),$$

so that $\varphi(\mathrm{sp}_T)$ has to be divisible by sp_T. It follows that $\mathbb{K}[X_1, \ldots, X_n]_\lambda$ is included in the Specht ideal I_λ^{sp}, and therefore Theorem 3 gives:

Theorem 7. *Let $\mathbb{K}$ be a field of characteristic 0 and $I \subset \mathbb{K}[X_1, \ldots, X_n]$ be a symmetric ideal. Assume that there exists $P \in I$ of degree d, such that $d + \mathrm{wt}(P_d) \leqslant n$. Then for every $m \in \mathrm{Mon}(P_d)$, for every $\lambda \trianglelefteq \mu(m)^\perp$, the ideal I contains the isotypic component $\mathbb{K}[X_1, \ldots, X_n]_\lambda$. In other words*

$$I_\lambda = \mathbb{K}[X_1, \ldots, X_n]_\lambda,$$

or equivalently

$$(\mathbb{K}[X_1, \ldots, X_n]/I)_\lambda = \{0\}.$$

Given polynomials $P_1, \ldots, P_l \in \mathbb{K}[X_1, \ldots, X_{n_0}]$, they naturally induce a symmetric ideal in $\mathbb{K}[X_1, \ldots, X_n]$ for any $n \geqslant n_0$. We get an increasing sequence of symmetric ideals $(I)_{n \geqslant n_0}$, and one can study stabilization properties in terms of representations (see for instance [19, 5]).

We remark that when n is large enough, the condition on the support of the leading component of P is automatically fulfilled and we can immediately deduce:

Theorem 8. *Let $\mathbb{K}$ be of characteristic 0 and n_0 be an integer. Given $Q_1, \ldots, Q_l$ in $\mathbb{K}[X_1, \ldots, X_{n_0}]$, consider for any $n \geqslant n_0$ the ideal*

$$I_n = \langle \sigma(Q_i), \sigma \in S_n, 1 \leqslant i \leqslant l \rangle.$$

Then if n is large enough, for every $1 \leqslant i \leqslant l$, every monomial m of Q_i of maximal degree, and any λ partition of n such that $\lambda \trianglelefteq \mu(m)^\perp$,

$$(\mathbb{K}[X_1, \ldots, X_n]/I_n)_\lambda = \{0\}.$$

5.3. **Symmetric sums of squares on symmetric varieties.** As a final application we consider sums of squares of real polynomials. Let $P \in \mathbb{K}[X_1, \ldots, X_n]$. Then P is called a sum of squares, if there exists polynomials $P_1, \ldots, P_k$ with

$$P = P_1^2 + \ldots + P_k^2.$$

Sum of squares are the cornerstone in the so called moment approach to polynomial optimization [12]: In general, it can be decided by semidefinite programming if a given polynomial affords a decomposition as a sum of squares. The case of symmetric sums of squares has received some interest by different authors [2, 8, 15, 17, 11].

In [2], Blekherman and the second author described how to characterize symmetric sums of squares through representation theory. More precisely, they use the theory of higher Specht polynomials [22] to construct, for every $\lambda \vdash n$, a square matrix polynomial Q^λ of size $s_\lambda = \dim(S^\lambda)$, whose entries are symmetric polynomials. Furthermore, these entries are products and sums of elements in $\mathbb{R}[X_1, \ldots, X_n]_\lambda$.

So even though they are symmetric, they belong to the ideal generated by the Specht polynomials of shape λ. These matrices can be used to show that every symmetric polynomial P that is a sum of squares can be written in the form

$$P = \sum_{\lambda \vdash n} \mathrm{Tr}(P^{\lambda} \cdot Q^{\lambda}),$$

where each $P^{\lambda} \in \mathbb{R}[X_1, \dots, X_n]^{s_\lambda \times s_\lambda}$ is a sum of symmetric squares matrix polynomial, i.e.

$$P^{\lambda} = L^t L$$

for some matrix L whose entries are symmetric polynomials.

Since the λ-Specht ideal contains all the coefficients of Q^{λ} we can apply Theorem 7 to obtain the following result on representations of a symmetric polynomial modulo a symmetric ideal.

Theorem 9. *Let $P \in \mathbb{R}[X_1, \dots, X_n]^{S_n}$ be a symmetric sum of squares polynomial and I be a symmetric ideal in $\mathbb{R}[X_1, \dots, X_n]$. Further, we assume that there exists $F \in I$ of degree d, such that $d + \mathrm{wt}(F_d) \leqslant n$. Then P can be written as*

$$P = \sum_{\nu \ntrianglelefteq \mu(m)^{\perp}} \mathrm{Tr}(P^{\lambda} \cdot Q^{\lambda}) \mod I,$$

where again each $P^{\lambda} = L^t L$ for some matrix polynomial L whose entries are symmetric polynomials.

A case of special interest is the case when I is the gradient ideal $I_{\mathrm{grad}}(P)$ of a given polynomial P of even degree $2d$. Sturmfels and Nie [14] showed that a polynomial that is positive on its gradient variety $V(I_{\mathrm{grad}})$ can always be written as a sum of squares modulo its gradient ideal. When P is a symmetric polynomial our results can be applied to reduce the problem size. Building on the ideas of considering the gradient ideal, Jibetean and Laurent [9] considered the following perturbation of a polynomial:

$$P_{\varepsilon} := \varepsilon \cdot (X_1^{2d+2} + \ldots + X_n^{2d+2}) + P.$$

Since the perturbation term is positive definite and of higher degree, a global minimum of P_{ϵ} converges to a global minimum of P, when $\varepsilon \to 0$. Furthermore, in this case the quotient $\mathbb{R}[X_1, \dots, X_n]/I_{\mathrm{grad}}(P_{\varepsilon})$ is generated by the polynomials $2dX_i^{2d-1} + \frac{\partial P}{\partial X_i}$, for $1 \leqslant i \leqslant n$. Thus the quotient is a finite dimensional vector space and contained in the residues of all polynomials of degree at most $2d - 1$. This implies in particular, that both the number and the sizes of the matrices in Theorem 9 can be restricted further (see [2, Proposition 4.1] and are in fact independent of n whenever $n > 2d + 1$. In order to use Theorem 9 practically to decide if a polynomial is a sum of squares of polynomials, one uses semi-definite programming. The complexity of such a program is mainly determined by the size of the matrices used to define it. Therefore our discussion actually shows that checking if a symmetric polynomial $P_{\varepsilon} \geqslant 0$ for an $\varepsilon > 0$ can be done in a complexity which does not depend on n.

REFERENCES

[1] S. Basu and C. Riener. On the isotypic decomposition of cohomology modules of symmetric semi-algebraic sets: polynomial bounds on multiplicities. *International Mathematics Research Notices*, to appear. 1

[2] G. Blekherman and C. Riener. Symmetric nonnegative forms and sums of squares. *arXiv preprint arXiv:1205.3102*, 2012. 10, 11

[3] T. Brylawski. The lattice of integer partitions. *Discrete mathematics*, 6(3):201–219, 1973. 4

[4] L. Bus and A. Karasoulou. Resultant of an equivariant polynomial system with respect to the symmetric group. *Journal of Symbolic Computation*, 76:142 – 157, 2016. 1

[5] T. Church, J. S. Ellenberg, B. Farb, et al. Fi-modules and stability for representations of symmetric groups. *Duke Mathematical Journal*, 164(9):1833–1910, 2015. 10

[6] J.-C. Faugère and J. Svartz. Solving polynomial systems globally invariant under an action of the symmetric group and application to the equilibria of n vortices in the plane. In *Proceedings of the 37th International Symposium on Symbolic and Algebraic Computation*, pages 170–178. ACM, 2012. 1

[7] R. Fröberg and B. Shapiro. On vandermonde varieties. *Math. Scand*, 119(1):7391, 2016. 4

[8] C. Goel, S. Kuhlmann, and B. Reznick. On the choi–lam analogue of hilbert's 1888 theorem for symmetric forms. *Linear Algebra and its Applications*, 496:114–120, 2016. 10

[9] D. Jibetean and M. Laurent. Semidefinite approximations for global unconstrained polynomial optimization. *SIAM Journal on Optimization*, 16(2):490–514, 2005. Pagination: 25. 11

[10] R. Krone. Equivariant gröbner bases of symmetric toric ideals. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, pages 311–318. ACM, 2016. 1

[11] A. Kurpisz, S. Leppänen, and M. Mastrolilli. Sum-of-squares hierarchy lower bounds for symmetric formulations. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 362–374. Springer, 2016. 10

[12] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on optimization*, 11(3):796–817, 2001. 10

[13] U. Nagel and T. Römer. Equivariant hilbert series in non-noetherian polynomial rings. *Journal of Algebra*, 486:204–245, 2017. 9

[14] J. Nie, J. Demmel, and B. Sturmfels. Minimizing polynomials via sum of squares over the gradient ideal. *Mathematical Programming*, 106(3):587–606, May 2006. 11

[15] A. Raymond, M. Singh, and R. R. Thomas. Symmetry in turán sums of squares polynomials from flag algebras. *Algebraic Combinatorics*, 1(2):249–274, 2018. 10

[16] C. Riener. On the degree and half-degree principle for symmetric polynomials. *Journal of Pure and Applied Algebra*, 216(4):850–856, 2012. 1, 7

[17] C. Riener, T. Theobald, L. J. Andrén, and J. B. Lasserre. Exploiting symmetries in sdp-relaxations for polynomial optimization. *Mathematics of Operations Research*, 38(1):122–141, 2013. 10

[18] B. E. Sagan. *The symmetric group: representations, combinatorial algorithms, and symmetric functions*, volume 203. Springer Science & Business Media, 2013. 2, 9

[19] S. V. Sam and A. Snowden. Stability patterns in representation theory. In *Forum of Mathematics, Sigma*, volume 3. Cambridge University Press, 2015. 10

[20] W. Specht. Die irreduziblen darstellungen der symmetrischen gruppe. *Mathematische Zeitschrift*, 39(1):696–711, 1935. 3

[21] S. Steidel. Gröbner bases of symmetric ideals. *Journal of Symbolic Computation*, 54:72–86, 2013. 1

[22] T. Terasoma, H. Yamada, et al. Higher specht polynomials for the symmetric group. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, 69(2):41–44, 1993. 10

[23] V. Timofte. On the positivity of symmetric polynomial functions.: Part i: General results. *Journal of Mathematical Analysis and Applications*, 284(1):174–190, 2003. 1, 7

[24] J. Watanabe and K. Yanagawa. Vandermonde determinantal ideals, 2019. 4

[25] K. Yanagawa. When is a specht ideal cohen-macaulay?, 2019. 4

Appendix A. Proof of Theorem 4 in general characteristic

Recall that we want to show that there exist polynomials $R_\sigma \in \mathbb{K}[X_1, \dots, X_n]$, for $\sigma \in S_n$ such that:

$$\Delta(J_1)\cdots\Delta(J_l) = \sum_{\sigma \in S_n} R_\sigma \sigma P.$$

In order to avoid an overload of notation we rename the variables in the following way:

$$Z_{i,s} \text{ for } 1 \leqslant i \leqslant \operatorname{len}(\mu) \text{ and } 1 \leqslant s \leqslant \mu_i,$$

where we identify $Z_{i,1}$ with X_i for $1 \leqslant i \leqslant \operatorname{len}(\mu)$. In this way whenever $s > 1$ then $Z_{i,s}$ does not appear in P_d. In this setting, the monomial m is written as

$$m = Z_{1,1}^{k_1} Z_{2,1}^{k_2} \cdots Z_{l,1}^{k_l}.$$

We will show the existence of polynomials R_σ, only involving the variables $Z_{i,s}$ for $1 \leqslant i \leqslant l$, such that

$$\Delta(J_1)\cdots\Delta(J_l) = \prod_{i=1}^{l} \prod_{1\leqslant s<t\leqslant \mu_i} (Z_{i,s} - Z_{i,t}) = \sum_{\sigma\in S_\mu} R_\sigma \sigma P,$$

where

$$S_\mu = S_{\mu_1} \times S_{\mu_2} \times \cdots \times S_{\mu_l} \times \underbrace{S_1 \times \cdots \times S_1}_{n-d-l}$$

and each factor acts on the corresponding subset of variables, namely:

$$\sigma Z_{i,s} = Z_{i,\sigma_i(s)}.$$

Let us show this by induction on the degree d of P. If $d = 0$, there is nothing to prove.

Now assume $d > 1$. Up to a rescaling, we may assume that

$$P = m + S,$$

where S is a polynomial of degree at most d such that $m \notin \mathrm{Mon}(S)$ and S_d does not contain any variable $Z_{i,s}$ with $s > 1$.

Let τ be the transposition exchanging $Z_{l,1}$ and Z_{l,μ_l}. Then

$$P - \tau P = (Z_{l,1} - Z_{l,\mu_l})Q$$

where Q is a polynomial of degree $d-1$. Because Z_{l,μ_l} does not appear in S_d, we can write

$$Q = m' + S'$$

with $m' = \left(\prod_{i=1}^{l-1} Z_{i,1}^{k_i}\right) Z_{l,1}^{k_l-1}$ and S' is a polynomial of degree at most $d-1$. Since the only new variable appearing in Q_d is Z_{l,μ_l}, we have

$$d - 1 + \mathrm{wt}(Q_d) \leqslant d + \mathrm{wt}(P_d) \leqslant n,$$

in such a way that we can apply the induction hypothesis on Q. This provides polynomials R'_ρ, only involving the variables $Z_{i,s}$ for $1 \leqslant i \leqslant l$, except Z_{l,μ_l}, such that

$$\left(\prod_{i=1}^{l-1} \prod_{1\leqslant s<t\leqslant \mu_i} (Z_{i,s} - Z_{i,t})\right) \prod_{1\leqslant s<t\leqslant \mu_l-1} (Z_{l,s} - Z_{l,t}) = \sum_{\rho\in S'} R'_\rho \rho Q,$$

where

$$S' = S_{\mu_1} \times S_{\mu_2} \times \cdots \times (S_{\mu_l-1} \times S_1) \times \underbrace{S_1 \times \cdots \times S_1}_{n-d-l}$$

can be seen as a subgroup of S_μ.

Since any $\rho \in S'$ leaves the product $\prod_{s=1}^{\mu_l-1}(Z_{l,s} - Z_{l,\mu_l})$ unchanged, we have

$$\begin{aligned}
\prod_{i=1}^{l} \prod_{1\leqslant s<t\leqslant \mu_i} (Z_{i,s} - Z_{i,t}) &= \prod_{s=1}^{\mu_l-1} (Z_{l,s} - Z_{l,\mu_l}) \sum_{\rho\in S'} R'_\rho \rho Q \\
&= \sum_{\rho\in S'} R'_\rho \rho \left(\left(\prod_{s=1}^{\mu_l-1} (Z_{l,s} - Z_{l,\mu_l})\right) Q\right) \\
&= \sum_{\rho\in S'} R'_\rho \prod_{s=2}^{\mu_l-1} (Z_{l,\rho_l(s)} - Z_{l,\mu_l}) \rho (P - \tau P).
\end{aligned}$$

Because $\rho\tau \in S_\mu$, we can rewrite this expression as $\sum_{\sigma \in S_\mu} R_\sigma \sigma P$ with the desired properties.

Department of Mathematics and Statistics, UiT - the Arctic University of Norway, 9037 Tromsø, Norway
E-mail address: `philippe.moustrou@uit.no`

Department of Mathematics and Statistics, UiT - the Arctic University of Norway, 9037 Tromsø, Norway
E-mail address: `cordian.riener@uit.no`

Department of Mathematics and Statistics, UiT - the Arctic University of Norway, 9037 Tromsø, Norway
E-mail address: `hugues.verdure@uit.no`