

# Young Flattenings in the Schur module basis

Lennart J. Haas and Christian Ikenmeyer

April 2021

## Abstract

There are several isomorphic constructions for the irreducible polynomial representations of the general linear group in characteristic zero. The two most well-known versions are called Schur modules and Weyl modules. Steven Sam used a Weyl module implementation in 2009 for his Macaulay2 package PieriMaps. This implementation can be used to compute so-called Young flattenings of polynomials. Over the Schur module basis Oeding and Farnsworth describe a simple combinatorial procedure that is supposed to give the Young flattening, but their construction is not equivariant. In this paper we clarify this issue, present the full details of the theory of Young flattenings in the Schur module basis, and give a software implementation in this basis. Using Reuven Hodges' recently discovered Young tableau straightening algorithm in the Schur module basis as a subroutine, our implementation outperforms Sam's PieriMaps implementation by several orders of magnitude on many examples, in particular for powers of linear forms, which is the case of highest interest for proving border Waring rank lower bounds.

**Keywords:** Young flattening, representation theory, Pieri's rule, border Waring rank, complexity lower bounds

**AMS Subject Classification 2020:** 05E10, 68Q17

**ACM Subject Classification:** Mathematics of computing → Mathematical software

Computing methodologies → Symbolic and algebraic manipulation → Computer algebra systems

Theory of computation → Computational complexity and cryptography → Algebraic complexity theory

## 1 Motivation

Young flattenings of polynomials are equivariant linear maps from a space of homogeneous polynomials to a space of matrices, where the row and column space are irreducible representations of  $\mathrm{GL}_n := \mathrm{GL}(\mathbb{C}^n)$ . One is usually interested in finding lower bounds for the rank of the image of a Young flattening, as it can be used to obtain lower bounds on the border Waring rank of a polynomial, and more generally for any border  $X$ -rank for a  $\mathrm{GL}_n$ -variety  $X$ , i.e., given a point  $p$  to find a lower bound on the smallest  $i$  such that  $p$  lies on the  $i$ -th secant variety of  $X$ , see e.g. [Lan15]. One early example are Sylvester's *catalecticants* [Syl52]. Landsberg and Ottaviani [LO15] use Young flattenings in the tensor setting. The name *Young flattening* was introduced in the predecessor paper [LO13]. Young flattenings also appear in disguise in the area of algebraic complexity theory as *matrices of partial derivatives*, *shifted partial derivatives*, *evaluation dimension*, and *coefficient dimension* [NW95, GKKS14]. They can in principle be used to find computational complexity lower bounds in many algebraic computational models such as border determinantal complexity (see [LMR13]) and border continuantal complexity [BIZ18], which makes Young flattenings an interesting tool in the Geometric Complexity Theory approach by Mulmuley and Sohoni [MS02], [MS08], [BLMW11]. Limits of these methods (in the case of studying  $X$ -rank) have recently been proved in [EGOW18, GMOW19]. No such limits are known for using Young flattenings to study the orbit closure containment problems in geometric complexity theory. First results in this direction were obtained in [ELSW18], where limits to the method of shifted partial derivatives are shown. This was improved on in [GL19], where a setting was given in which Young flattenings give strictly more separation information than partial derivatives.

The *Waring rank* of a homogeneous degree  $d$  polynomial  $p \in \mathbf{S}^d V$  is defined as the smallest  $r$  such that  $p$  can be written as a sum of  $r$  many  $d$ -th powers of homogeneous linear forms (arbitrary linear combinations of  $d$ -th powers are usually allowed if the base field is not algebraically closed).

arXiv:2104.02363v1 [math.AG] 6 Apr 2021

For example  $(x - y)^3 + y^3 = x^3 - 3x^2y + 3xy^2$ , hence  $x^3 - 3x^2y + 3xy^2$  has Waring rank at most 2. The *border Waring rank* of  $p$  is the smallest  $r$  such that  $p$  can be approximated arbitrarily closely coefficient-wise by polynomials of Waring rank at most  $r$ . For example  $3\epsilon x^2y = \lim_{\epsilon \rightarrow 0} ((x + \epsilon y)^3 - x^3)$ , hence  $x^2y$  has border Waring rank at most 2.

If a Young diagram  $\lambda$  is contained in another Young diagram  $\mu$  such that the column lengths of both diagrams differ by at most 1 in each column, then we have a unique nonzero equivariant map between  $\mathbf{S}^d V \otimes \mathbf{S}^\lambda V \rightarrow \mathbf{S}^\mu V$ , where  $\mathbf{S}^\lambda V$ ,  $\mathbf{S}^\mu V$ , and  $\mathbf{S}^d V$  are irreducible polynomial  $\mathrm{GL}_n$ -representations, and  $d$  is the difference in the number of boxes of  $\mu$  and  $\lambda$ . This is called the *Pieri map*, and it induces a linear map  $\mathcal{F}_{\lambda, \mu} : \mathbf{S}^d V \rightarrow \mathrm{End}(\mathbf{S}^\lambda V, \mathbf{S}^\mu V)$ . Since border Waring rank is subadditive, a lower bound on the border Waring rank of  $p$  is obtained by rounding up to quotient of ranks

$$\left\lceil \frac{\mathrm{rank}(\mathcal{F}_{\lambda, \mu}(p))}{\mathrm{rank}(\mathcal{F}_{\lambda, \mu}(x^d))} \right\rceil, \quad (1.1)$$

where  $x$  is some variable that appears in  $p$ , and  $\mathrm{rank}(\cdot)$  is the usual rank of matrices.

There are several isomorphic constructions for the irreducible polynomial representations of the general linear group in characteristic zero. The two best known versions are called Schur modules and Weyl modules and they only differ in the order of the row-symmetrizer and the column-symmetrizer in their definition of the Young symmetrizer. This results in different bases for the irreducible representations. Sometimes results that are proved in one basis are reproved in the other basis, but the proofs look significantly different (see e.g. [BCI11] and [MM14]). In fact, so far some results are only provable in a natural way over one basis and not the other, see e.g. [Res20]. Based on an explicit paper by Olver over the Weyl module basis [Olv82] Steven Sam in 2009 implemented his Macaulay2 package PieriMaps [Sam08], which among other things can be used to compute the rank quotient (1.1), see Section 4 (A) below.

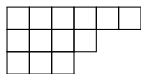
The papers [Far16] (in its Section 5<sup>1</sup>) and [Oed16] (only in version 1) describe the PieriMaps package as if it would be working in the Schur module basis and they assume that the Young flattenings have an extremely simple combinatorial description. However, this is wrong (see Section 5 below), which led to a revision of [Oed16].

In this paper we work out the details of Young flattenings in the Schur module basis: We closely mimic the arguments in [Olv82], but we take care of subtle sign issues that are not present in Olver's work over the Weyl module basis. We then make use of a recent fast algorithm (and implementation) by Reuven Hodges for Young tableau straightening in the Schur module basis [Hod17] to get a highly efficient Young flattening algorithm that outperforms Sam's PieriMaps implementation by several orders of magnitude in many examples. We obtain the most impressive speedup factor of 1000 for flattening the power of a linear form, which is the denominator of (1.1).

Our contribution is therefore twofold: We thoroughly clarify the theory of Young flattenings in the Schur module basis and we present a new and efficient implementation for Young flattenings that uses Hodges' state-of-the-art straightening algorithm over the Schur module basis.

## 2 Preliminaries

A *composition*  $\nu$  of a number  $d$  is a finite list of natural numbers adding up to  $d$ , i.e.,  $(3, 0, 2, 4)$  is a composition of 9. A *partition* is a nonincreasing composition, for example  $\lambda = (6, 4, 3)$  is a partition. We write  $\lambda \vdash d$  if  $\lambda$  is a partition of  $d$ . We write  $\lambda_i = 0$  if  $i$  is greater than the number of entries in  $\lambda$ . We define  $\ell(\lambda) := \min\{i \mid \lambda_i = 0\} - 1$ . We identify a partition with its *Young diagram*, which is a top-left justified array of boxes, i.e., the set of points  $\{(i, j) \mid 1 \leq i \text{ and } 1 \leq j \leq \lambda_i\}$ . For example, the Young diagram corresponding to  $(6, 4, 3)$  is



and we have  $(2, 4) \in \lambda$  and  $(4, 2) \notin \lambda$ . We see that  $\ell(\lambda)$  is the number of rows of the Young diagram corresponding to  $\lambda$ . We denote by  $|\lambda|$  the number of boxes in  $\lambda$ , i.e.,  $|\lambda| = \sum_i \lambda_i$ . We denote by  $\lambda^*$

<sup>1</sup> Although the description in the paper is wrong, the use of the software package is correct and gives the result claimed in the paper.

the Young diagram obtained by reflecting  $\lambda$  at the main diagonal, e.g.,  $(6, 4, 3)^* = (3, 3, 3, 2, 1, 1)$ . It follows that  $\lambda_i^*$  is the length of the  $i$ -th column of  $\lambda$ . We write  $\lambda \subseteq \mu$  if for all  $(i, j) \in \lambda$  we have  $(i, j) \in \mu$ . If  $\lambda \subseteq \mu$ , then we denote by  $\mu/\lambda$  the set of points that are in  $\mu$  but not in  $\lambda$ . We call  $\mu/\lambda$  a *horizontal strip* if it has at most 1 box in each column. In this situation we write  $\mu/\lambda \in \text{HS}$ .

A Young diagram  $\lambda$  whose entries are labeled with numbers is called a *Young tableau* of shape  $\lambda$ . For example,

$$\begin{array}{|c|c|c|c|c|} \hline 6 & 3 & 4 & 2 & 3 & 3 \\ \hline 3 & 3 & 1 & 2 & & \\ \hline 6 & 6 & 3 & & & \\ \hline \end{array}$$

is a Young tableau of shape  $(6, 4, 3)$ . A Young tableau is called *semistandard* if the entries strictly increase in each column from top to bottom and do not decrease in each row from left to right. For example,  $\begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 3 \\ \hline 2 & 3 & & \\ \hline \end{array}$  is a semistandard tableau. We denote by  $\mathfrak{S}_\lambda$  the symmetric group on the set  $\{(i, j) \mid (i, j) \in \lambda\}$  of positions in  $\lambda$ . The group  $\mathfrak{S}_\lambda$  acts on the set of all Young tableaux of shape  $\lambda$  by permuting the positions. We write  $\sigma T$  for the permuted Young tableau, where  $\lambda$  is the shape of  $T$  and  $\sigma \in \mathfrak{S}_\lambda$ . For a subset  $S \subseteq \lambda$  of positions we write  $\mathfrak{S}_S$  to denote the symmetric group that permutes only the positions in  $S$  among each other and fixes all other positions.

Let  $V^{\otimes \lambda} := V^{\otimes |\lambda|}$  be the  $|\lambda|$ -th tensor power of a vector space  $V$  and associate to every tensor factor  $V$  a position in  $\lambda$ . A rank 1 tensor  $v = v_1 \otimes v_2 \otimes \cdots \otimes v_{|\lambda|}$  can now be represented by a Young diagram in whose  $i$ -th box we write the vector  $v_i$ . If we fix a basis  $v_1, \dots, v_n$  of  $V$ , then a basis of  $V^{\otimes \lambda}$  is obtained by all ways of writing  $\{v_1, \dots, v_n\}$  into the boxes of  $\lambda$ , allowing repetitions. If the fixed basis is clear from the context, then we write  $i$  instead of  $v_i$  into the boxes and obtain a Young tableau. The basis vector corresponding to the Young tableau  $T$  is also denoted by  $T$  when no confusion can arise, so for example if  $x = v_1 + v_2$  we can use the multilinearity of the tensor product to write

$$\begin{array}{|c|c|} \hline x & v_2 \\ \hline v_1 & \\ \hline \end{array} = \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 1 & \\ \hline \end{array} + \begin{array}{|c|c|} \hline 2 & 2 \\ \hline 1 & \\ \hline \end{array}$$

## 2 (A) The Weyl module basis

Let  $\lambda$  be a Young diagram and let  $(i, j) \in \lambda$  such that  $(i + 1, j) \in \lambda$  (i.e., the box below  $(i, j)$  is still in  $\lambda$ ). Then we define  $B_{i,j} := \{(i, k) \mid j \leq k \leq \lambda_i\} \cup \{(i + 1, k) \mid 1 \leq k \leq j\}$ . Pictorially,  $B_{i,j}(\lambda)$  is the subset of the boxes of  $\lambda$  given by collecting all boxes on the following path: Start at position  $(i + 1, 1)$  and move from left to right along row  $i + 1$  to box  $(i + 1, j)$ , then switch the row to  $(i, j)$  and move along row  $i$  until reaching  $(i, \lambda_i)$ . For example,  $B_{1,3}((6, 4, 3))$  is given by the dotted boxes in the following diagram:

$$\begin{array}{|c|c|c|c|c|} \hline & & \bullet & \bullet & \bullet & \\ \hline \bullet & \bullet & \bullet & & & \\ \hline & & & & & \\ \hline \end{array}$$

**2.1 Definition** (Weyl module, [Wey03, 2.1.15]). *The Weyl module  $\mathbf{W}^\lambda(V)$  is defined as the quotient space  $V^{\otimes \lambda}/P^\lambda$ , where  $P^\lambda \subseteq V^{\otimes \lambda}$  is the linear subspace generated by the following two types of vectors:*

1. (Symmetric relation)  $T - \sigma T$ , if  $\sigma$  is a permutation that preserves the row indices of all positions of  $\lambda$  (in other words,  $\sigma$  permutes within the rows of  $\lambda$ ).
2. (Shuffle relation)  $\sum_{\sigma \in \mathfrak{S}_{B_{i,j}(\lambda)}} \sigma T$ , if  $i, j \in \mathbb{N}$  such that  $(i, j), (i + 1, j) \in \lambda$ .

The Weyl modules for Young diagrams  $\lambda$  with at most  $\dim V$  rows form a complete list of pairwise non-isomorphic irreducible polynomial representations of  $\text{GL}(V)$ . In this paper we will not work with Weyl modules, but with the isomorphic Schur modules, which are defined in the following section.

## 2 (B) The Schur module basis

Let  $T$  be a Young tableau of shape  $\lambda$ . Let  $1 \leq i < j \leq \lambda_1$  be two column indices. Let  $B$  and  $C$  be two equally large sets of boxes,  $B$  from column  $i$  and  $C$  from column  $j$ . An *exchange tableau* of  $T$

corresponding to  $B$  and  $C$  is defined as the tableau arising from  $T$  by exchanging the content of the boxes  $B$  with the content of the boxes  $C$  while preserving the vertical order of the entries in  $B$  and  $C$ . We denote this exchange tableau by  $E_C^B(T)$ . For a subset  $C$  of boxes from column  $j$ , we write  $E_C^i(T) := \bigcup_B E_C^B(T)$ , where  $B$  ranges over all cardinality  $|C|$  subsets of boxes in column  $i$ .

**2.2 Definition** (Schur module). *The Schur module  $\mathbf{S}^\lambda V$  is defined as the quotient space  $V^{\otimes \lambda}/Q^\lambda$  where  $Q^\lambda \subseteq V^{\otimes \lambda}$  is the linear subspace generated by the following vectors*

1. (Grassmann relation)  $T + T'$ , where  $T'$  is obtained from  $T$  by swapping two elements in the same column.
2. (Plücker relation)  $T - \sum_{T' \in E_C^i(T)} T'$  for any  $i$  and any subset  $C$  of a column  $j \neq i$  with  $\lambda_i^* \geq |C|$ .

For example, in  $\mathbf{S}^\lambda V$  we have  $\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} = 0$  and we have  $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = -\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$  by the Grassmann relation. The Plücker relation gives

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

via  $E_{\{(1,2)\}}^1$ . Note that  $\{(1,2)\}$  is the set containing the single box in row 1 and column 2.

## 2 (C) The Schur module via moving boxes between columns

If we only consider the Grassmann relation, then we call the quotient  $\mathbf{X}^\lambda V$ . More formally, let  $G^\lambda \subseteq V^{\otimes \lambda}$  be the linear subspace spanned by the  $T + T'$ , where  $T'$  is obtained from  $T$  by swapping two elements in the same column. The quotient  $V^{\otimes \lambda}/G^\lambda$  is denoted by  $\mathbf{X}^\lambda V$ . Clearly, in the language of skew-symmetric powers we have

$$\mathbf{X}^\nu V \simeq (\bigwedge^{\nu_1} V) \otimes \cdots \otimes (\bigwedge^{\nu_{\ell(\nu)}} V). \quad (2.3)$$

In terms of explicit bases, this isomorphism maps each column from top to bottom to a skew-symmetric tensor and vice versa: For example,  $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$  is mapped to  $x_1 \wedge x_3 \otimes x_2 \wedge x_4$ . We define

$$Y_d := \bigoplus_{\text{composition } \alpha \text{ of } d} \mathbf{X}^\alpha V$$

as an *outer* direct sum. Note that each  $\mathbf{X}^\alpha V \subseteq V^{\otimes |\alpha|}$ , but there is no obvious embedding of  $Y_d$  in  $V^{\otimes |\alpha|}$ . In fact, there is a natural isomorphism

$$Y_d \simeq \bigwedge^d (V^{\oplus d}) \quad (2.4)$$

that can be described explicitly in terms of basis vectors using first the isomorphism (2.3): A standard basis vector

$$x_{k_{1,1}} \wedge \cdots \wedge x_{k_{\nu_1,1}} \otimes x_{k_{1,2}} \wedge \cdots \wedge x_{k_{\nu_2,2}} \otimes \cdots \otimes x_{k_{1,d}} \wedge \cdots \wedge x_{k_{\nu_d,d}}$$

is mapped to

$$x_{k_{1,1},1} \wedge \cdots \wedge x_{k_{\nu_1,1},1} \wedge x_{k_{1,2},2} \wedge \cdots \wedge x_{k_{\nu_2,2},2} \wedge \cdots \wedge x_{k_{1,d},d} \wedge \cdots \wedge x_{k_{\nu_d,d},d}, \quad (2.5)$$

for example  $x_1 \wedge x_3 \otimes x_2 \wedge x_4$  is mapped to  $x_{1,1} \wedge x_{3,1} \wedge x_{2,2} \wedge x_{4,2}$ . We have an action of  $\mathbf{GL}(V^{\oplus d})$  on  $\bigwedge^d (V^{\oplus d})$ , which induces an action of  $\mathbf{GL}(V)$  via the group homomorphism  $g \mapsto \text{id}_d \otimes g$ , i.e., sending matrices  $g$  to block diagonal matrices that have  $d$  many copies of  $g$  on their main diagonal. With this action of  $\mathbf{GL}(V)$ , the isomorphism in (2.4) is an isomorphism of  $\mathbf{GL}(V)$ -representations.

For given  $i, j$ ,  $i \neq j$ , we consider the Lie algebra element  $\mathfrak{g}_{i,j} \in \mathfrak{gl}(V^{\oplus d})$  that is a block matrix as follows: the block matrix is zero everywhere but the block  $(i, j)$  is the identity matrix on  $V$ . Clearly applying  $\mathfrak{g}_{i,j}$  is a  $\mathbf{GL}(V)$ -equivariant map. Acting with  $\mathfrak{g}_{i,j}$  on the tensor from (2.5) gives

$$\sum_{1 \leq a \leq \nu_i} x_{k_{1,1},1} \wedge \cdots \wedge x_{k_{\nu_1,1},1} \wedge \cdots \wedge x_{k_{1,i},[i=a,j,i]} \wedge \cdots \wedge x_{k_{\nu_i,i},[i=a,j,i]} \wedge \cdots \wedge x_{k_{d,1},d} \wedge \cdots \wedge x_{k_{\nu_d,d},d}, \quad (2.6)$$

where

$$[i=a, j, i] := \begin{cases} j & \text{if } i = a \\ i & \text{otherwise.} \end{cases}$$

We denote by the map  $\sigma_{i,j}$  the application of  $\mathfrak{g}_{i,j}$ . Using the canonical isomorphisms (2.4) and (2.3) we can write this more explicitly: For  $x \in \Lambda^{\alpha_i}$ ,  $y \in \Lambda^{\alpha_j}$ ,  $z \in \Lambda^{\alpha_k}$  have

$$\begin{aligned} \sigma_{i,k}(x \otimes y \otimes z) &= \sum_{1 \leq a \leq \alpha_i} (-1)^{a+\alpha_i+\alpha_j} x_{-a} \otimes y \otimes (x_a \wedge z) \\ \sigma_{k,i}(x \otimes y \otimes z) &= \sum_{1 \leq c \leq \alpha_k} (-1)^{c+\alpha_j+1} (x \wedge z_c) \otimes y \otimes z_{-c} \end{aligned}$$

Here for a tensor  $x \in \bigwedge^d V$  we define  $x_{-k}$  as the tensor obtained by ‘‘removing the  $k$ -th tensor position’’ (this is only well-defined if a basis of  $V$  is fixed and an ordering of the basis vectors is fixed). For example, if  $x = e_3 \wedge e_1 \wedge e_4 \wedge e_3$ , then  $x_{-3} = e_3 \wedge e_1 \wedge e_3$ , and  $x_3 = e_4$ .

Note that if  $\lambda_j^* = 0$ , then  $\sigma_{j,j+1}(\mathbf{X}^{\lambda^*}) = \{0\}$ . Define  $\mathcal{I}^\lambda := \langle \sigma_{i,i+1}(Y_d) \cap \mathbf{X}^{\lambda^*} V \mid i \in \mathbb{N} \rangle$ . Note that if  $\lambda_j^* > 0$ , then  $\sigma_{j,j+1}(\mathbf{X}^{\lambda^*}) \subseteq \mathbf{X}^{\lambda^* - e_j + e_{j+1}}$ , hence

$$\mathcal{I}^\lambda = \langle \sigma_{i,i+1}(\mathbf{X}^{\lambda^* + e_i - e_{i+1}}) \mid \lambda_{i+1}^* > 0 \rangle. \quad (2.7)$$

**2.8 Proposition** ([Tow79, Cor. 1]).  $\mathbf{S}^\lambda V$  and  $\mathbf{X}^{\lambda^*} V / \mathcal{I}^\lambda$  are isomorphic representations of  $\mathrm{GL}(V)$ . The isomorphism maps each basis vector given by a semistandard tableau to a basis vector corresponding to the same semistandard tableau.

**2.9 Proposition** ([Tow77, Thm. 2.5]).  $\mathbf{S}^\lambda V$  and  $\mathbf{W}^\lambda V$  are isomorphic representations of  $\mathrm{GL}(V)$ .

The relation between the basis vectors corresponding to semistandard tableaux in Prop. 2.9 is more involved than the straightforward relationship in Prop. 2.8.

### 3 Young Flattenings in the Schur module basis

We give an explicit description of the construction of the so-called *Pieri inclusions* defined on the basis of Schur modules. Olver [Olv82] first described the corresponding construction based on Weyl modules and we closely mimic this construction while taking care of the subtle signs that are introduced when using the Schur module basis. To the best of our knowledge, this construction has never been explicitly described for Schur modules. This algorithm will directly give the construction for Young flattenings.

Pieri’s well-known formula states the following isomorphism of  $\mathrm{GL}(V)$ -representations:

$$\mathbf{S}^d V \otimes \mathbf{S}^\lambda V = \bigotimes_{\substack{\mu \vdash d + |\lambda| \\ \mu / \lambda \in \mathrm{HS}}} \mathbf{S}^\mu V.$$

The resulting  $\mathrm{GL}(V)$ -equivariant inclusions

$$\varphi_{\lambda,\mu} : \mathbf{S}^d V \otimes \mathbf{S}^\lambda V \rightarrow \mathbf{S}^\mu V$$

are called *Pieri inclusions* and are unique up to scale by Schur’s lemma. We define  $\varphi_{\lambda,\mu}$  by composing Pieri inclusions for  $d = 1$  as follows. For  $\mu / \lambda \in \mathrm{HS}$  let  $(\lambda = \lambda^{(0)}, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(d)} = \mu)$  be the sequence of partitions obtained by adding one box at a time to  $\lambda$  from left to right, so that after having added  $d$  boxes we arrive at  $\mu$ . Then define the  $\mathrm{GL}(V)$ -equivariant map  $\psi_{\lambda,\mu} : V^{\otimes d} \otimes \mathbf{S}^\lambda \rightarrow \mathbf{S}^\mu$  via

$$\psi_{\lambda,\mu} = \varphi_{\lambda^{(d-1)}, \lambda^{(d)}} \circ (\mathrm{id}_V \otimes \varphi_{\lambda^{(d-2)}, \lambda^{(d-1)}}) \circ \dots \circ (\mathrm{id}_{V^{\otimes (d-2)}} \otimes \varphi_{\lambda^{(1)}, \lambda^{(2)}}) \circ (\mathrm{id}_{V^{\otimes (d-1)}} \otimes \varphi_{\lambda^{(0)}, \lambda^{(1)}}). \quad (3.1)$$

By proving that the restriction of  $\psi_{\lambda,\mu}$  to  $\mathbf{S}^d V \otimes \mathbf{S}^\lambda V$  is nonzero (see Lemma 3.10), it immediately follows that this restriction equals  $\varphi_{\lambda,\mu}$  (up to a nonzero scalar). It remains to describe  $\varphi_{\lambda,\mu}$  for which  $\mu / \lambda$  has only a single box and then prove Lemma 3.10.

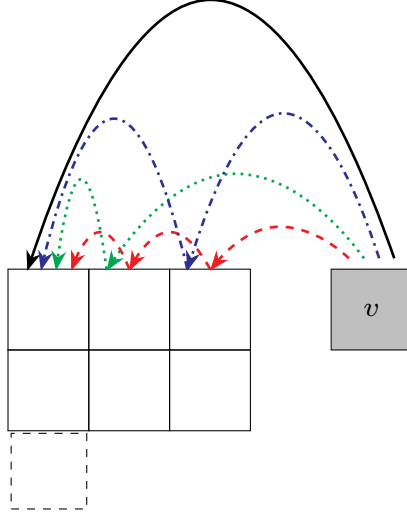


Fig. 1: The paths we need to traverse, when considering the Pieri inclusion from  $(3, 3)$  to  $(3, 3, 1)$ . Note that the boxes are moved in the direction of the arrows, but that the box movements are executed on each path starting with the *leftmost* arrow.

### 3 (A) The single box case

We assume that  $\mu/\lambda$  consists of a single box.

We will define the linear map  $\zeta_{\lambda, \mu} : V \otimes \mathbf{X}^{\lambda^*} V \rightarrow \mathbf{X}^{\mu^*} V$ , show its  $\mathrm{GL}(V)$ -equivariance (Lemma 3.4), and then prove that it is well-defined on the quotient space  $V \otimes \mathbf{S}^{\lambda} V$  if we interpret its image in the quotient space  $\mathbf{S}^{\mu} V$  (Theorem 3.5). In this way,  $\zeta_{\lambda, \mu}$  induces a map  $\varphi_{\lambda, \mu} : V \otimes \mathbf{S}^{\lambda} V \rightarrow \mathbf{S}^{\mu} V$ . By the uniqueness of the Pieri inclusions we have that this map equals  $\varphi_{\lambda, \mu}$  or is the zero map (nonzeroness is proved in Lemma 3.10). Note that  $V \otimes \mathbf{X}^{\lambda^*} V \simeq \mathbf{X}^{(\lambda^*, 1)} V$  by definition.

Let  $J = (J_1, \dots, J_p)$  be a finite sequence of positive integers. Define the linear map

$$\sigma_J : Y_d \rightarrow Y_d, \quad \sigma_J := \sigma_{J_1, J_2} \circ \sigma_{J_2, J_3} \circ \dots \circ \sigma_{J_{p-1}, J_p}. \quad (3.2)$$

Pictorially, this means that a box is shifted from column  $J_{p-1}$  to  $J_p$ , then from  $J_{p-2}$  to  $J_{p-1}$ , and so on. Clearly  $\sigma_J$  is  $\mathrm{GL}(V)$ -equivariant as the composition of  $\mathrm{GL}(V)$ -equivariant maps. Note that if  $v \in \mathbf{X}^{(\lambda^*, 1)} V$  and  $\sigma_J$  moves a box from an empty column to another column, then  $\sigma_J(v) = 0$ .

We consider the set of all strongly decreasing sequences of natural numbers from  $m$  to  $k$ , which we denote by

$$\mathcal{A}_k^m := \{J = (J_1, \dots, J_p) \mid p \in \mathbb{N}, k = J_p < J_{p-1} < \dots < J_1 = m\}.$$

For  $\iota \geq k$  let  $h_{k, \iota} := \lambda_k^* - \lambda_{\iota}^* + \iota - k + 1$  be the hook length in  $\lambda$  of the box in column  $k$  and row  $\iota$ . Let  $k$  be the column where  $\mu$  and  $\lambda$  differ. We define  $D_J(\lambda^*)$  as the product of all hook lengths in  $J$  with respect to column  $k$ , i.e.,

$$D_J(\lambda^*) := \prod_{q=2}^{|J|-1} h_{k, J_q}(\lambda^*) = \prod_{q=2}^{|J|-1} (\lambda_k^* - \lambda_{J_q}^* + J_q - k + 1). \quad (3.3)$$

Finally, we define  $\zeta_{\lambda}^{\mu} : Y_d \rightarrow Y_d$  by

$$\zeta_{\lambda}^{\mu} = \sum_{J \in \mathcal{A}_k^{\lambda^*+1}} \frac{\sigma_J}{D_J(\lambda^*)}.$$

The map  $\zeta_{\lambda}^{\mu}$  maps  $\mathbf{X}^{(\lambda^*, 1)} V$  into  $\mathbf{X}^{\mu^*} V$ .

**3.4 Lemma.**  $\zeta_\lambda^\mu : \mathbf{X}^{(\lambda^*, 1)} V \rightarrow \mathbf{X}^{\mu^*} V$  is  $\mathrm{GL}(V)$ -equivariant.

*Proof.*  $\zeta_\lambda^\mu$  is a linear combination of  $\mathrm{GL}(V)$ -equivariant maps.  $\square$

**3.5 Theorem.** Let  $\lambda, \mu$  be partitions with  $\mu$  being obtained from  $\lambda$  by appending a single box in row  $k$ . Then  $\zeta_\lambda^\mu(V \otimes \mathcal{I}_\lambda) \subseteq \mathcal{I}_\mu$  and hence  $\varphi_{\lambda, \mu}$  is well-defined.

Before proving Theorem 3.5 we first have to prove the following lemma.

We denote by  $[f, g] := fg - gf$  the commutator of two linear maps  $f$  and  $g$ .

**3.6 Lemma.** Let  $v \in \mathbf{X}^\alpha V$  for some column lengths  $\alpha$  with  $\alpha_i \neq 0$  and  $\alpha_k \neq 0$  and let  $i \neq j, k \neq l \in \mathbb{N}$ . Then,

$$[\sigma_{i,j}, \sigma_{k,l}](v) = \begin{cases} (\alpha_j - \alpha_i)v & \text{if } i = l, j = k, \\ \sigma_{k,j}(v) & \text{if } i = l, j \neq k, \\ -\sigma_{i,l}(v) & \text{if } i \neq l, j = k, \\ 0 & \text{otherwise.} \end{cases}$$

Note the similarity to Lemma 5.4 in [Olv82] with the exception that the sign in the first case is reversed. Moreover, [Olv82] ignores handling the special case when column lengths vanish. We handle these cases explicitly. If  $\lambda_j^* = 0, \lambda_i^* \neq 0$  and  $v \in \mathbf{X}^{\lambda^*}$ , then

$$\sigma_{j,i} \sigma_{i,j}(v) = \lambda_i^* v. \quad (3.7)$$

Moreover, if  $\lambda_j^* = 0, \lambda_i^* \neq 0, i \neq k$  and  $v \in \mathbf{X}^{\lambda^*}$ , then

$$\sigma_{j,k} \sigma_{i,j}(v) = \sigma_{i,k}(v). \quad (3.8)$$

*Proof of Lemma 3.6.* We focus on the key positions in the tensor.

- $i = l, j = k$ :  $x \in \wedge^{\alpha_i} V, y \in \wedge^{\alpha_j} V$ .

$$\begin{aligned} [\sigma_{i,j}, \sigma_{j,i}](x \otimes y) &= \sigma_{i,j} \left( \sum_{1 \leq b \leq \alpha_j} (-1)^{b+1} (x \wedge y_b) \otimes y_{-b} \right) - \sigma_{j,i} \left( \sum_{1 \leq a \leq \alpha_i} (-1)^{a+\alpha_i} x_{-a} \otimes (x_a \wedge y) \right) \\ &= \sum_{\substack{1 \leq a \leq \alpha_i+1 \\ 1 \leq b \leq \alpha_j}} (-1)^{a+b+\alpha_i} (x \wedge y_b)_{-a} \otimes ((x \wedge y_b)_a \wedge y_{-b}) \\ &\quad - \sum_{\substack{1 \leq a \leq \alpha_i \\ 1 \leq b \leq \alpha_j}} (-1)^{a+b+1+\alpha_i} (x_{-a} \wedge (x_a \wedge y)_b) \otimes (x_a \wedge y)_{-b} \\ &= \sum_{\substack{1 \leq a \leq \alpha_i \\ 1 \leq b \leq \alpha_j}} (-1)^{a+b+\alpha_i} (x_{-a} \wedge y_b) \otimes (x_a \wedge y_{-b}) \\ &\quad - \sum_{\substack{2 \leq b \leq \alpha_j+1 \\ 1 \leq a \leq \alpha_i}} (-1)^{a+b+\alpha_i+1} (x_{-a} \wedge y_{b-1}) \otimes (x_a \wedge y_{-(b-1)}) \\ &\quad + \sum_{1 \leq b \leq \alpha_j} (-1)^{b+1} x \otimes (y_b \wedge y_{-b}) \\ &\quad - \sum_{1 \leq a \leq \alpha_i} (-1)^{a+\alpha_i} (x_{-a} \wedge x_a) \otimes y \\ &= \sum_{1 \leq b \leq \alpha_j} x \otimes y - \sum_{1 \leq a \leq \alpha_i} x \otimes y = (\alpha_j - \alpha_i)(x \otimes y) \end{aligned}$$

- $i = l, j \neq k$ :  $x \in \wedge^{\alpha_i} V, y \in \wedge^{\alpha_j} V, z \in \wedge^{\alpha_K} V$ .

$$\begin{aligned}
[\sigma_{i,j}, \sigma_{k,i}](x \otimes y \otimes z) &= \sigma_{i,j} \left( \sum_{1 \leq c \leq \alpha_k} (-1)^{\alpha_j+c+1} (x \wedge z_c) \otimes y \otimes z_{-c} \right) \\
&\quad - \sigma_{k,i} \left( \sum_{1 \leq a \leq \alpha_i} (-1)^{\alpha_i+a} x_{-a} \otimes (x_a \wedge y) \otimes z \right) \\
&= \sum_{\substack{1 \leq c \leq \alpha_k \\ 1 \leq a \leq \alpha_i+1}} (-1)^{\alpha_i+\alpha_j+a+c} (x \wedge z_c)_{-a} \otimes ((x \wedge z_c)_a) \wedge y \otimes z_{-c} \\
&\quad - \sum_{\substack{1 \leq a \leq \alpha_i \\ 1 \leq c \leq \alpha_k}} (-1)^{a+c+\alpha_i+\alpha_j} (x_{-a} \wedge z_c) \otimes (x_a \wedge y) \otimes z_{-c} \\
&= \sum_{\substack{1 \leq a \leq \alpha_i \\ 1 \leq c \leq \alpha_k}} (-1)^{\alpha_i+\alpha_j+a+c} (x_{-a} \wedge z_c) \otimes (x_a \wedge y) \otimes z_{-c} \\
&\quad - \sum_{\substack{1 \leq a \leq \alpha_i \\ 1 \leq c \leq \alpha_k}} (-1)^{\alpha_i+\alpha_j+a+c} (x_{-a} \wedge z_c) \otimes (x_a \wedge y) \otimes z_{-c} \\
&\quad + \sum_{1 \leq c \leq \alpha_k} (-1)^{\alpha_j+c+1} x \otimes (z_c \wedge y) \otimes z_{-c} \\
&= \sum_{1 \leq c \leq \alpha_k} (-1)^{c+1} x \otimes (y \wedge z_c) \otimes z_{-c} \\
&= \sigma_{k,j}(x \otimes y \otimes z)
\end{aligned}$$

- $i \neq l, j = k$ : Equivalent to the case  $i = l, j \neq k$  but changing the order of elements in the commutator. Thus, the sign changes:

$$[\sigma_{i,j}, \sigma_{j,l}](x \otimes y \otimes z) = -[\sigma_{j,l}, \sigma_{i,j}](x \otimes y \otimes z) = -\sigma_{i,l}(x \otimes y \otimes z)$$

- $i \neq l, j \neq k$ : If all  $i, j, k, l$  are pairwise distinct, then both maps affect distinct columns and hence they commute. We first treat the case  $j = l \notin \{i, k\}, i \neq k$ .

$$\begin{aligned}
[\sigma_{i,k}, \sigma_{j,k}](x \otimes y \otimes z) &= \sigma_{i,k} \left( \sum_{1 \leq b \leq \alpha_j} (-1)^{\alpha_j+b} x \otimes y_{-b} \otimes (y_b \wedge z) \right) \\
&\quad - \sigma_{j,k} \left( \sum_{1 \leq a \leq \alpha_i} (-1)^{\alpha_i+\alpha_j+a} x_{-a} \otimes y \otimes (x_a \wedge z) \right) \\
&= \sum_{\substack{1 \leq a \leq \alpha_i \\ 1 \leq b \leq \alpha_j}} (-1)^{\alpha_j+b+a+\alpha_i+\alpha_j+1} x_{-a} \otimes y_{-b} \otimes (x_a \wedge y_b \wedge z) \\
&\quad - \sum_{\substack{1 \leq a \leq \alpha_i \\ 1 \leq b \leq \alpha_j}} (-1)^{\alpha_j+b+a+\alpha_i+\alpha_j} x_{-a} \otimes y_{-b} \otimes (y_b \wedge x_a \wedge z) \\
&= 0
\end{aligned}$$



We now treat the remaining case  $i = k \notin \{j, l\}$ ,  $j \neq l$ . For a basis vector  $x$  let  $x_{-\{a,b\}}$  denote the basis vector with positions  $a$  and  $b$  removed.

$$\begin{aligned}
[\sigma_{i,j}, \sigma_{i,l}](x \otimes y \otimes z) &= \sigma_{i,j} \left( \sum_{1 \leq a \leq \alpha_i} (-1)^{\alpha_j + a + \alpha_i} x_{-a} \otimes y \otimes (x_a \wedge z) \right) \\
&\quad - \sigma_{i,l} \left( \sum_{1 \leq a \leq \alpha_i} (-1)^{\alpha_i + a} x_{-a} \otimes (x_a \wedge y) \otimes z \right) \\
&= \sum_{\substack{1 \leq a \leq \alpha_i \\ 1 \leq a' \leq \alpha_i - 1}} (-1)^{a + a' + \alpha_j + 1} (x_{-a})_{-a'} \otimes ((x_{-a})_{a'} \wedge y) \otimes (x_a \wedge z) \\
&\quad - \sum_{\substack{1 \leq a \leq \alpha_i \\ 1 \leq a' \leq \alpha_i - 1}} (-1)^{a + a' + \alpha_j} (x_{-a})_{-a'} \otimes (x_a \wedge y) \otimes ((x_{-a})_{a'} \wedge z) \\
&= \sum_{\substack{1 \leq a, \tilde{a} \leq \alpha_i \\ a \neq \tilde{a}}} (-1)^{a + \tilde{a} + \alpha_j + 1 + [\tilde{a} > a]} x_{-\{a, \tilde{a}\}} \otimes (x_{\tilde{a}} \wedge y) \otimes (x_a \wedge z) \\
&\quad - \sum_{\substack{1 \leq a, \tilde{a} \leq \alpha_i \\ a \neq \tilde{a}}} (-1)^{a + \tilde{a} + \alpha_j + [\tilde{a} > a]} x_{-\{a, \tilde{a}\}} \otimes (x_a \wedge y) \otimes (x_{\tilde{a}} \wedge z) \\
&= 0
\end{aligned}$$

where  $[b > a]$  is 1 if  $b > a$  and 0 otherwise. Here we used the notation

$$\tilde{a} = \begin{cases} a' & \text{if } a' < a \\ a' + 1 & \text{if } a' \geq a \end{cases}.$$

Note that the second case happens exactly when  $\tilde{a} > a$ . □

We will make heavy use of the following identity: Let  $\sigma_1, \sigma_2, \sigma_3$  be linear maps, then

$$[\sigma_1 \circ \sigma_2, \sigma_3] = \sigma_1 \circ [\sigma_2, \sigma_3] + [\sigma_1, \sigma_3] \circ \sigma_2. \quad (3.9)$$

The rule can be interpreted as the Leibniz rule for  $ad_A(B) = [A, B]$ .

*Proof of Theorem 3.5.* Using (2.7) we see that it suffices to prove that if  $w \in \sigma_{i,i+1}(\mathbf{X}^{\lambda^* + e_i - e_{i+1}, 1})$ , then  $\zeta_{\lambda, \mu}(w) \in \sigma_{i,i+1}(\mathbf{X}^{\mu^* + e_i - e_{i+1}})$ . Let  $w = \sigma_{i,i+1}(v)$  with  $v \in \mathbf{X}^{\lambda^* + e_i - e_{i+1}}$ . We now show that  $\zeta_{\lambda, \mu}(\sigma_{i,i+1}(v)) = \sigma_{i,i+1}(\zeta_{\lambda, \mu}(v))$ , which finishes the proof.

We split the proof according to the different relations of  $i$  and  $k$ :

- $i < k$ : If  $i+1 < k$ , then clearly  $[\sigma_A, \sigma_{i,i+1}] = 0$  for every  $A \in \mathcal{A}_k^{\lambda_1+1}$ , because  $A \cap \{i, i+1\} = \emptyset$ . Consider the case  $i+1 = k$ . Every  $A \in \mathcal{A}_k^{\lambda_1+1}$  can be written as  $(B, i+1)$  with  $B \in \mathcal{A}_m^{\lambda_1+1}$  for some  $m > i+1$ .

$$[\sigma_A, \sigma_{i,i+1}] \stackrel{B \cap \{i, i+1\} = \emptyset}{=} \sigma_B [\sigma_{m, i+1}, \sigma_{i, i+1}] \stackrel{\text{Lemma 3.6}}{=} 0.$$

- $i = k$ : We divide the sequences in  $\mathcal{A}_k^{\lambda_1+1}$  (summed up over in  $\zeta_\lambda^\mu$ ) as follows: For every  $m > k+1$  and  $B \in \mathcal{A}_m^{\lambda_1+1}$ , let either  $A_2 = (B, k+1, k)$  or  $A_1 = (B, k)$ . In fact, the sequences come in pairs. Adding/removing the entry  $k+1$  maps the elements of the pairs to each other.

$$[\sigma_{A_1}, \sigma_{k, k+1}](v) \stackrel{B \cap \{k, k+1\} = \emptyset}{=} \sigma_B \circ [\sigma_{m, k}, \sigma_{k, k+1}](v) \stackrel{\text{Lemma 3.6}}{=} -\sigma_B \circ \sigma_{m, k+1}(v),$$

as well as (if  $\lambda_{k+1}^* > 1$ ):

$$\begin{aligned}
[\sigma_{A_2}, \sigma_{k, k+1}](v) &\stackrel{B \cap \{k, k+1\} = \emptyset}{=} \sigma_B \circ [\sigma_{m, k+1} \circ \sigma_{k+1, k}, \sigma_{k, k+1}](v) \\
&\stackrel{(3.9)}{=} \sigma_B \circ (\sigma_{m, k+1} [\sigma_{k+1, k}, \sigma_{k, k+1}] + [\sigma_{m, k+1}, \sigma_{k, k+1}] \circ \sigma_{k+1, k})(v) \\
&\stackrel{\text{Lemma 3.6}}{=} \sigma_B \circ (\sigma_{m, k+1} [\sigma_{k+1, k}, \sigma_{k, k+1}](v)) \\
&\stackrel{\text{Lemma 3.6, } v \in \mathbf{X}^{\lambda^* + e_i - e_{i+1}}}{=} ((\lambda_k^* + 1) - (\lambda_{k+1}^* - 1)) \sigma_B \circ \sigma_{m, k+1}(v) \\
&= (\lambda_k^* - \lambda_{k+1}^* + 2) \sigma_B \circ \sigma_{m, k+1}(v)
\end{aligned}$$

If  $\lambda_{k+1}^* = 1$ , then the same is true:

$$\begin{aligned}
[\sigma_{A_2}, \sigma_{k,k+1}](v) &= \sigma_B \circ (\sigma_{m,k+1}[\sigma_{k+1,k}, \sigma_{k,k+1}])(v) \\
&\stackrel{v \in \mathbf{X}^{\lambda^* + e_k - e_{k+1}}}{=} \sigma_B \circ \sigma_{m,k+1} \circ \sigma_{k+1,k} \circ \sigma_{k,k+1}(v) \\
&\stackrel{(3.7)}{=} (\lambda_k^* + 1)\sigma_B \circ \sigma_{m,k+1}(v) \\
&= (\lambda_k^* - \lambda_{k+1}^* + 2)\sigma_B \circ \sigma_{m,k+1}(v)
\end{aligned}$$

We take a weighted sum of two paired up sequences:

$$\begin{aligned}
[D_{A_1}^{-1}\sigma_{A_1} + D_{A_2}^{-1}\sigma_{A_2}, \sigma_{k,k+1}](v) &= (-D_{A_1}^{-1} + D_{A_2}^{-1}(\lambda_k^* - \lambda_{k+1}^* + 2))\sigma_B \circ \sigma_{m,k+1}(v) \\
&= D_{A_1}^{-1} \left( -1 + \frac{\lambda_k^* - \lambda_{k+1}^* + 2}{\lambda_k^* - \lambda_{k+1}^* + 2} \right) \sigma_B \circ \sigma_{m,k+1}(v) \\
&= 0
\end{aligned}$$

since  $D_{A_2} = D_{A_1}(\lambda_k^* - \lambda_{k+1}^* + 2)$ , because  $h_{k,k+1}(\lambda^*) = \lambda_k^* - \lambda_{k+1}^* + 2$  by (3.3). Since the weighted sum over two paired up sequences yields zero, the weighted sum over all sequences in  $\mathcal{A}_k^{\lambda_1+1}$  yields zero.

- $i > k$ : Again, we divide the sequences in  $\mathcal{A}_k^{\lambda_1+1}$ : For every  $B \in \mathcal{A}_k^{m_1}$  and  $C \in \mathcal{A}_{m_2}^{\lambda_1+1}$ , we have  $A_1 = (C, i, B)$ ,  $A_2 = (C, i+1, B)$ , and  $A_3 = (C, i+1, i, B)$  for  $B \in \mathcal{A}_k^{m_1}$  and  $C \in \mathcal{A}_{m_2}^{\lambda_1+1}$  for some  $m_1 < i$  and  $m_2 > i+1$ . This time the sequences come in quadruples of sequences that can be obtained from each other by adding/removing  $i$  and  $i+1$ . Clearly  $[\sigma_{A_0}, \sigma_{i,i+1}](v) = 0$ , because  $(B \cup C) \cap \{i, i+1\} = \emptyset$ . So these sequences contribute zero to the sum  $[\zeta_\lambda^\mu, \sigma_{i,i+1}]$ . We ignore these sequences and are left with triples of sequences instead of quadruples.

We have

$$\begin{aligned}
[\sigma_{A_1}, \sigma_{i,i+1}](v) &\stackrel{\{i,i+1\} \cap (B \cup C) = \emptyset}{=} \sigma_C \circ ([\sigma_{m_2,i} \circ \sigma_{i,m_1}, \sigma_{i,i+1}]) \circ \sigma_B(v) \\
&\stackrel{(3.9)}{=} \sigma_C \circ (\sigma_{m_2,i} \circ [\sigma_{i,m_1}, \sigma_{i,i+1}] + [\sigma_{m_2,i}, \sigma_{i,i+1}] \circ \sigma_{i,m_1}) \circ \sigma_B(v) \\
&\stackrel{\text{Lemma 3.6}}{=} -\sigma_C \circ \sigma_{m_2,i+1} \circ \sigma_{i,m_1} \circ \sigma_B(v)
\end{aligned}$$

If  $\lambda_{i+1}^* > 1$ :

$$\begin{aligned}
[\sigma_{A_2}, \sigma_{i,i+1}](v) &\stackrel{\{i,i+1\} \cap (B \cup C) = \emptyset}{=} \sigma_C \circ [\sigma_{m_2,i+1} \circ \sigma_{i+1,m_1}, \sigma_{i,i+1}] \circ \sigma_B(v) \\
&\stackrel{(3.9)}{=} \sigma_C \circ (\sigma_{m_2,i+1} \circ [\sigma_{i+1,m_1}, \sigma_{i,i+1}] + [\sigma_{m_2,i+1} \circ \sigma_{i,i+1}] \circ \sigma_{i+1,m_1}) \circ \sigma_B(v) \\
&\stackrel{\text{Lemma 3.6}}{=} \sigma_C \circ \sigma_{m_2,i+1} \circ [\sigma_{i+1,m_1}, \sigma_{i,i+1}] \circ \sigma_B(v) \\
&\stackrel{\text{Lemma 3.6, } \lambda_{i+1}^* > 1}{=} \sigma_C \circ \sigma_{m_2,i+1} \circ \sigma_{i,m_1} \circ \sigma_B(v)
\end{aligned}$$

We get the same result for the case  $\lambda_{i+1}^* = 1$ :

$$\begin{aligned}
[\sigma_{A_2}, \sigma_{i,i+1}](v) &= \sigma_C \circ \sigma_{m_2,i+1} \circ [\sigma_{i+1,m_1}, \sigma_{i,i+1}] \circ \sigma_B(v) \\
&= \sigma_C \circ \sigma_{m_2,i+1} \circ \sigma_{i+1,m_1} \circ \sigma_{i,i+1} \circ \sigma_B(v) \\
&\stackrel{(3.8)}{=} \sigma_C \circ \sigma_{m_2,i+1} \circ \sigma_{i,m_1} \circ \sigma_B(v)
\end{aligned}$$

If  $\lambda_{i+1}^* > 1$ :

$$\begin{aligned}
& [\sigma_{A_3}, \sigma_{i,i+1}](v) \\
\stackrel{\{i,i+1\} \cap (B \cup C) = \emptyset}{=} & \sigma_C \circ [\sigma_{m_2,i+1} \circ \sigma_{i+1,i} \circ \sigma_{i,m_1}, \sigma_{i,i+1}] \circ \sigma_B(v) \\
\stackrel{(3.9)}{=} & \sigma_C \circ (\sigma_{m_2,i+1} \circ \sigma_{i+1,i} \circ [\sigma_{i,m_1}, \sigma_{i,i+1}] + [\sigma_{m_2,i+1} \circ \sigma_{i+1,i}, \sigma_{i,i+1}] \circ \sigma_{i,m_1}) \circ \sigma_B(v) \\
\stackrel{\text{Lemma 3.6}}{=} & \sigma_C \circ ([\sigma_{m_2,i+1} \circ \sigma_{i+1,i}, \sigma_{i,i+1}] \circ \sigma_{i,m_1}) \circ \sigma_B(v) \\
\stackrel{(3.9)}{=} & \sigma_C \circ (\sigma_{m_2,i+1} \circ [\sigma_{i+1,i} \circ \sigma_{i,i+1}] + [\sigma_{m_2,i+1}, \sigma_{i,i+1}] \circ \sigma_{i+1,i}) \circ \sigma_{i,m_1} \circ \sigma_B(v) \\
\stackrel{\text{Lemma 3.6}}{=} & \sigma_C \circ (\sigma_{m_2,i+1} \circ [\sigma_{i+1,i} \circ \sigma_{i,i+1}]) \circ \sigma_{i,m_1} \circ \sigma_B(v) \\
\stackrel{\text{Lemma 3.6, } v \in \mathbf{X}^{\lambda^* + e_i - e_{i+1}}}{=} & (\lambda_i^* - (\lambda_{i+1}^* - 1)) \sigma_C \circ \sigma_{m_2,i+1} \circ \sigma_{i,m_1} \circ \sigma_B(v).
\end{aligned}$$

Note that in the last equation we used that  $(\sigma_{i,m_1} \circ \sigma_B)(v) \in \mathbf{X}^\nu V$  with  $\nu_i = \lambda_i^*$  and  $\nu_{i+1} = \lambda_{i+1}^* - 1$ , because  $v \in \mathbf{X}^{\lambda^* + e_i - e_{i+1}}$  and  $B \cap \{i, i+1\} = \emptyset$ .

We get the same result for the case  $\lambda_{i+1}^* = 1$ :

$$\begin{aligned}
& [\sigma_{A_3}, \sigma_{i,i+1}](v) \\
&= \sigma_C \circ (\sigma_{m_2,i+1} \circ [\sigma_{i+1,i} \circ \sigma_{i,i+1}]) \circ \sigma_{i,m_1} \circ \sigma_B(v) \\
\stackrel{(3.7)}{=} & \lambda_i^* \sigma_C \circ \sigma_{m_2,i+1} \circ \sigma_{i,m_1} \circ \sigma_B(v)
\end{aligned}$$

We take a weighted sum of a triple of grouped sequences:

$$[D_{A_1}^{-1} \sigma_{A_1} + D_{A_2}^{-1} \sigma_{A_2} - D_{A_3}^{-1} \sigma_{A_3}, \sigma_{i,i+1}](v) = 0,$$

which can be seen as follows:  $D_{A_3} = D_{A_1} h_{k,i+1}(\lambda^*)$  and  $D_{A_3} = D_{A_2} h_{k,i+1}(\lambda^*)$  implies

$$\begin{aligned}
& \frac{-1}{D_{A_1}} + \frac{1}{D_{A_2}} + \frac{\lambda_i^* - \lambda_{i+1}^* + 1}{D_{A_3}} = \\
& \frac{1}{D_{A_3}} (-h_{k,i+1}(\lambda^*) + h_{k_i}(\lambda^*) + \lambda_i^* - \lambda_{i+1}^* + 1) = -\lambda_i^* + \lambda_{i+1}^* - 1 + \lambda_i^* - \lambda_{i+1}^* + 1 = 0. \quad \square
\end{aligned}$$

### 3 (B) Nonzeroness

Next, we show that the map  $\varphi_\lambda^\mu$  is nonzero. Let  $Z^\lambda$  denote the Young tableau of shape  $\lambda$  in which each box in row  $i$  has entry  $i+1$ , with the exception that columns with  $\dim V$  many boxes just have entries  $1, 2, \dots, \dim V$  from top to bottom. Recall the definition of  $\psi_\lambda^\mu$  from (3.1).

**3.10 Lemma.** *Let  $\lambda, \mu$  be partitions with  $\lambda \subseteq \mu$ ,  $\mu/\lambda \in \text{HS}$ . Then the restriction of  $\psi_\lambda^\mu$  to  $\mathbf{S}^d V \otimes \mathbf{S}^\lambda V$  is nonzero. More precisely,  $\psi_\lambda^\mu(v_1^{\otimes(|\mu| - |\lambda|)} \otimes Z^\lambda)$  is nonzero.*

*Proof.* Let  $k$  denote the smallest column index in which  $\lambda$  and  $\mu$  differ. Since in all columns to the right of column  $k$  we only have numbers that appear in column  $k$ , the only transition sequence in  $A_k^{\lambda_1+1}$  which does not vanish is  $J = (k, \lambda_1 + 1)$ . Hence the image of  $\varphi_\lambda^{\lambda'}$  is a single tableau which is either of shape  $\mu$ , or for the first column  $k'$  in which  $\mu$  and  $\lambda'$  differ we have that all columns right of column  $k'$  only have entries that occur in column  $k'$ . Therefore, again there is only one transition sequence. We continue this and end up with a tableau that differs from  $Z^\lambda$  by having additional entries 1 in each column where  $\mu$  and  $\lambda$  differ. This tableau does not have a repeated entry in any column, so straightening this tableau does not result in the zero vector [Hod20], which finishes the proof.  $\square$

In the proof of Lemma 3.10 we used the exact order of maps in (3.1). The following small argument shows that this order does not matter.

**3.11 Claim.** *If the boxes in (3.1) are added in any other order, then we get the same map up to a nonzero scalar.*

*Proof.* First, we can see that the proof in Lemma 3.10 can be adapted to show nonzeroness for different orders. Indeed, if we add the boxes in a different order, then more transition sequences have to be considered, but the only relevant transition sequences all end up with the same tableau and they all give a positive contribution to the end result, so nothing cancels out.

Since  $\varphi_{\lambda,\mu}$  maps  $\mathbf{S}^d V \otimes \mathbf{S}^\lambda V$  to  $\mathbf{S}^\mu V$  and the right-hand side is irreducible and the left-hand side contains a single copy of  $\mathbf{S}^\mu V$ , Schur's lemma implies that all such maps are the same up to scale.  $\square$

## 4 Software

The Pieri inclusion

$$\varphi_{\lambda,\mu} : \mathbf{S}^d V \otimes \mathbf{S}^\lambda V \rightarrow \mathbf{S}^\mu V$$

induces a linear map

$$\mathcal{F}_{\lambda,\mu} : \mathbf{S}^d V \rightarrow \text{End}(\mathbf{S}^\lambda V, \mathbf{S}^\mu V) \quad (4.1)$$

For a homogeneous polynomial  $p$  we are interested in the rank of the image  $\mathcal{F}_{\lambda,\mu}(p)$ . Analogously for Weyl modules.

In this section we provide a small example for computing Young flattenings in the basis of Schur and Weyl modules, using our implementation and Sam's Macaulay2 implementation. In both cases we will be working over  $V = \mathbb{C}^3$  with a basis  $\{a, b, c\}$ . We will consider the shapes  $\lambda = (2, 1, 1)$  and  $\mu = (5, 2, 1)$  and are interested in the rank of the respective flattening of the following polynomial

$$p = a^3 + bc^2 \in \mathbb{Q}[a, b, c]_{=3} \cong \mathbf{S}^{(3)} V.$$

In other words we search for

$$\text{rank}(\mathcal{F}_{(5,2,1),(2,1)}(a^3 + bc^2)).$$

First, let us take a look at Macaulay2.

### 4 (A) An example of PieriMaps

To the best of our knowledge, the first use of Macaulay2 to compute the rank of Young flattenings together with code examples was given by Oeding [Oed16].

The following command loads the `PieriMaps` package:

```
loadPackage "PieriMaps"
```

We can define a polynomial ring over the rational numbers and define the polynomial  $p$  as follows:

```
R = QQ[a,b,c]
p = a^3 + b c^2
```

The function `pieri` available in `PieriMaps` computes the polarization map

$$\mathcal{P}_{\mu,\lambda}^{\mathbf{W}} : \mathbf{W}^\mu V \rightarrow \mathbf{W}^{(d)} V \otimes \mathbf{W}^\lambda V,$$

which was also explicitly described by Olver [Olv82]. It is the dual of the Pieri inclusion.

The function `pieri` takes three arguments: the dominating tableau  $\mu$ ; a list of  $d$  row indices  $r = (r_1, \dots, r_d)$ , such that  $\lambda$  can be obtained from  $\mu$  by deleting the last box in row  $r_1$ , then the last box in row  $r_2$ , and so on; and the underlying polynomial ring  $\mathbb{Q}[a, b, c]$ .

The following code computes a matrix representing  $\mathcal{P}_{(5,2,1),(4,1)}^{\mathbf{W}}$ . Note that to go from  $(5, 2, 1)$  to  $(4, 1)$  we have to remove a box in row 1, then from row 2, and finally from row 3. The following command corresponds to this operation:

```
MX = pieri({5,2,1}, {1,2,3}, R)
```

which outputs a  $24 \times 24$  matrix of homogeneous degree 3 polynomials in  $a, b, c$ . We differentiate the entries by  $p$  and compute the rank

```
rank(diff(p, MX))
```

which gives 18. It follows that  $\text{rank}(\mathcal{P}_{\lambda,\mu}^{\mathbf{W}}(p)) = 18$ .

## 4 (B) Using our software

Our implementation is available as ancillary files to this paper. The README file contains detailed installation instructions. For the rank computation we rely on the linear algebra implementation of Macaulay2, so we assume that Macaulay2 is installed.

After the installation, our tool is called from the command line with 4 parameters:

1. The number of variables  $n$ ,
2. the partition  $\mu$ ,
3. the list of row indices  $(r_1, \dots, r_d)$  from which boxes are to be removed to obtain  $\lambda$ ,
4. and the polynomial  $p$ .

The example from the previous sections is calculated via

```
./flattening 3 [5,2,1] [1,2,3] a^3+b*c^2
```

which also outputs 18.

## 4 (C) Running time comparison

We compared our implementation to the `PieriMaps` package. Our implementation does not have multi-processor support. The computations were run on a laptop, quad-core i5-6200U CPU with 2.30GHz with 8GB of memory. On this fairly weak machine, the larger examples from [Oed19] crash `PieriMaps`. Our software constructs flattening matrices for each monomial and adds them up, so for a fairer comparison we used random dense polynomials  $p$  that were generated with the following `Sagemath` code (adjust the degree and the number of variables appropriately):

```
var('a,b,c,d,e')
R=ZZ[a,b,c,d,e]
str(sum([R.random_element(degree=4,terms=Infinity)
        .homogenize(var=randint(0,len(R.gens())-1)) for _ in range(10)]).replace(" ",""))
```

We chose six quite different examples in our comparison.

| $n$ | $\mu$         | $r$             | $p$                       | <code>PieriMaps</code> | Our software |
|-----|---------------|-----------------|---------------------------|------------------------|--------------|
| 5   | [4,4,4,4]     | [4,4,4,4]       | random                    | 1m 18s                 | 6s           |
| 5   | [5,3,1]       | [1,2,3]         | random                    | 5m 19s                 | 2s           |
| 5   | [7,5,4,3,2]   | [1,1,2,3,4,5,5] | $x_0^2 x_1^2 x_2 x_3 x_4$ | 59m 22s                | 1m           |
| 5   | [7,5,4,3,2]   | [1,1,2,3,4,5,5] | $x^7$                     | 59m 47s                | 1s           |
| 4   | [7,5,3,1]     | [1,2,3,4]       | random                    | 2h 35m 3s              | 9s           |
| 6   | [7,5,4,3,2,1] | [1,1,2,3,4,5,6] | $x^7$                     | >11h <sup>2</sup>      | 16s          |

We observe that our software is much faster than the `PieriMaps` implementation. The most extreme boost (an improvement factor of over 1000) is obtained when flattening the important case  $x^d$ . This rank is used in the denominator of (1.1). The use of the Schur basis allows us to stop a computation path for a tableau as soon as the tableau has a double entry in a column. This speeds up the computation. But we think that most of the speed-up comes from the fact that we circumvent the construction of the parameterized flattening matrix and construct the matrix directly.

## 5 The oversimplification in the literature

In [Far16] (Section 5) and [Oed16] (Def. 3.2) the Young flattening is described with an simple procedure that we call the *box-filling flattening*. We present it here and give a small counterexample to its equivariance.

<sup>2</sup> We manually terminated the computation after 11 hours.

We define the box-filling flattening  $\mathcal{F}_{\lambda,\mu}^{fill} : \mathbf{S}^{(d)}V \rightarrow \text{Hom}(\mathbf{S}^\lambda V, \mathbf{S}^\mu V)$  for monomials and use linear continuation. For  $\alpha \in \mathbb{N}^n$  we use the notation  $x^\alpha := \prod_{i=1}^n x_i^{\alpha_i}$ . Let  $x^\alpha \in \mathbf{S}^{(d)}V$ . Let  $T$  be a semistandard Young tableau of shape  $\lambda$  with entries from  $\{1, \dots, \dim V\}$ . These  $T$  with the Grassmann-Plücker relations form a basis of  $\mathbf{S}^\lambda V$ . Define

$$\mathcal{F}_{\lambda,\mu}^{fill}(x^\alpha)(T) = \sum_S S,$$

where the sum is over all  $S$  of shape  $\mu$  that can be obtained from  $T$  by adding boxes to  $T$ : exactly  $\alpha_i$  boxes with number  $i$ . The resulting  $S$  may not be semistandard, but can be expressed over the basis of semistandard tableaux via straightening.

Let  $\lambda = (2)$ ,  $\mu = (2, 1)$ ,  $V = \mathbb{C}^2$ ,  $g = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \text{GL}_2$ .

The matrices corresponding to the linear maps  $\mathcal{F}_{(2),(2,1)}^{fill}(x_1)$  and  $\mathcal{F}_{(2),(2,1)}^{fill}(x_2)$  are

$$\begin{array}{c|ccc} \mathcal{F}_{(2),(2,1)}^{fill}(x_1) & \boxed{11} & \boxed{12} & \boxed{22} \\ \hline \boxed{\begin{array}{c} 1 \\ 2 \end{array}} \boxed{1} & 0 & 0 & 0 \\ \boxed{\begin{array}{c} 1 \\ 2 \end{array}} \boxed{2} & 0 & 0 & -1 \end{array} \qquad \begin{array}{c|ccc} \mathcal{F}_{(2),(2,1)}^{fill}(x_2) & \boxed{11} & \boxed{12} & \boxed{22} \\ \hline \boxed{\begin{array}{c} 1 \\ 2 \end{array}} \boxed{1} & 1 & 0 & 0 \\ \boxed{\begin{array}{c} 1 \\ 2 \end{array}} \boxed{2} & 0 & 1 & 0 \end{array}$$

but applying  $g$  to the matrix on the left-hand side (this means swapping column 1 and 3, swapping row 1 and 2, and inverting the sign) does not yield the matrix on the right-hand side, but instead yields

$$\begin{array}{c|ccc} g\mathcal{F}_{(2),(2,1)}^{fill}(x_1) & \boxed{11} & \boxed{12} & \boxed{22} \\ \hline \boxed{\begin{array}{c} 1 \\ 2 \end{array}} \boxed{1} & 1 & 0 & 0 \\ \boxed{\begin{array}{c} 1 \\ 2 \end{array}} \boxed{2} & 0 & 0 & 0 \end{array}$$

Using the correct maps that we obtained in this paper, we obtain the following matrices instead.

$$\begin{array}{c|ccc} \mathcal{F}_{(2),(2,1)}(x_1) & \boxed{11} & \boxed{12} & \boxed{22} \\ \hline \boxed{\begin{array}{c} 1 \\ 2 \end{array}} \boxed{1} & 0 & \frac{1}{2} & 0 \\ \boxed{\begin{array}{c} 1 \\ 2 \end{array}} \boxed{2} & 0 & 0 & 1 \end{array} \qquad \begin{array}{c|ccc} \mathcal{F}_{(2),(2,1)}(x_2) & \boxed{11} & \boxed{12} & \boxed{22} \\ \hline \boxed{\begin{array}{c} 1 \\ 2 \end{array}} \boxed{1} & -1 & 0 & 0 \\ \boxed{\begin{array}{c} 1 \\ 2 \end{array}} \boxed{2} & 0 & -\frac{1}{2} & 0 \end{array}$$

## Acknowledgments

We thank Reuven Hodges for important discussions. The authors were supported by the DFG grant IK 116/2-1.

## References

- [BCI11] Peter Bürgisser, Matthias Christandl, and Christian Ikenmeyer. Even partitions in plethysms. *Journal of Algebra*, 328(1):322 – 329, 2011.
- [BIZ18] Karl Bringmann, Christian Ikenmeyer, and Jeroen Zuiddam. On algebraic branching programs of small width. *J. ACM*, 65(5), August 2018.
- [BLMW11] Peter Bürgisser, J. M. Landsberg, Laurent Manivel, and Jerzy Weyman. An overview of mathematical issues arising in the geometric complexity theory approach to  $\text{VP} \neq \text{VNP}$ . *SIAM J. Comput.*, 40(4):1179–1209, 2011.

- [EGOW18] Klim Efremenko, Ankit Garg, Rafael Oliveira, and Avi Wigderson. Barriers for Rank Methods in Arithmetic Complexity. In Anna R. Karlin, editor, *9th Innovations in Theoretical Computer Science Conference (ITCS 2018)*, volume 94 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 1:1–1:19, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- [ELSW18] Klim Efremenko, Joseph M. Landsberg, Hal Schenck, and Jerzy Weyman. The method of shifted partial derivatives cannot separate the permanent from the determinant. *Math. Comput.*, 87(312):2037–2045, 2018.
- [Far16] Cameron Farnsworth. Koszul–Young flattenings and symmetric border rank of the determinant. *Journal of Algebra*, 447:664 – 676, 2016.
- [GKKS14] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. *J. ACM*, 61(6):33:1–33:16, 2014.
- [GL19] Fulvio Gesmundo and Joseph M. Landsberg. Explicit polynomial sequences with maximal spaces of partial derivatives and a question of K. Mulmuley. *Theory of Computing*, 15(3):1–24, 2019.
- [GMOW19] A. Garg, V. Makam, R. Oliveira, and A. Wigderson. More barriers for rank methods, via a “numeric to symbolic” transfer. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 824–844, 2019.
- [Hod17] Reuven Hodges. A closed non-iterative formula for straightening fillings of Young diagrams. *arXiv:1710.05214*, 2017.
- [Hod20] Reuven Hodges. A non-iterative formula for straightening fillings of Young diagrams. manuscript, based on arXiv:1710.05214, 2020.
- [Lan12] Joseph M. Landsberg. *Tensors : Geometry and Applications* . Providence, R.I. : American Mathematical Society, 2012.
- [Lan15] J. M. Landsberg. Geometric complexity theory: an introduction for geometers. *Annali dell’ Università di Ferrara*, 61(1):65–117, May 2015.
- [LMR13] J. M. Landsberg, Laurent Manivel, and Nicolas Ressayre. Hypersurfaces with degenerate duals and the geometric complexity theory program. *Commentarii Mathematici Helvetici*, 88(2):469–484, 2013.
- [LO13] J. M. Landsberg and Giorgio Ottaviani. Equations for secant varieties of veronese and other varieties. *Annali di Matematica Pura ed Applicata*, 192(4):569–606, Aug 2013.
- [LO15] Joseph M. Landsberg and Giorgio Ottaviani. New lower bounds for the border rank of matrix multiplication. *Theory of Computing*, 11(11):285–298, 2015.
- [MM14] Laurent Manivel and Mateusz Michałek. Effective constructions in plethysms and weintraub’s conjecture. *Algebras and Representation Theory*, 17(2):433–443, Apr 2014.
- [MS02] Ketan D. Mulmuley and Milind Sohoni. Geometric complexity theory I: An approach to the P vs. NP and related problems. *SIAM J. Comput.*, 31(2):496–526, 2002.
- [MS08] Ketan D. Mulmuley and Milind Sohoni. Geometric complexity theory II: Towards explicit obstructions for embeddings among class varieties. *SIAM J. Comput.*, 38:1175–1206, July 2008.
- [NW95] Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. In *Proceedings of the 36th Annual Symposium on Foundations of Computer Science, FOCS ’95*, pages 16–25, Washington, DC, USA, 1995. IEEE Computer Society.

- [Oed16] Luke Oeding. Border ranks of monomials. *arXiv:1608.02530v1*, August 2016.
- [Oed19] Luke Oeding. Border ranks of monomials. *arXiv:1608.02530v3*, January 2019.
- [Olv82] Peter J. Olver. *Differential Hyperforms I*. Preprint. University of Minnesota, 1982. available at [http://www-users.math.umn.edu/~olver/a\\_/hyper.pdf](http://www-users.math.umn.edu/~olver/a_/hyper.pdf).
- [Res20] Nicolas Ressayre. Vanishing symmetric kronecker coefficients. *Beiträge zur Algebra und Geometrie / Contributions to Algebra and Geometry*, 61(2):231–246, Jun 2020.
- [Sam08] Steven V. Sam. PieriMaps: A *Macaulay2* package. Version 1.0. A *Macaulay2* package available at <https://github.com/Macaulay2/M2/tree/master/M2/Macaulay2/packages>, 2008.
- [Str83] V. Strassen. Rank and optimal computation of generic tensors. *Linear Algebra and its Applications*, 52-53:645 – 685, 1983.
- [Syl52] James J. Sylvester. *On the Principles of the Calculus of Forms*. Cambridge and Dublin Mathematical Journal, 1852.
- [Tow77] Jacob Towber. Two new functors from modules to algebras. *Journal of Algebra*, 47:80–104, 1977.
- [Tow79] Jacob Towber. Young symmetry, the flag manifold, and representations of  $gl(n)$ . *Journal of Algebra*, 61(2):414 – 462, 1979.
- [Wey03] Jerzy M. Weyman. *Cohomology of Vector Bundles and Syzygies*. Cambridge Tracts in Mathematics. Cambridge University Press, 2003.