Algorithms for sparse polynomial systems : Gröbner bases and resultants

Matias Rafael Bender

▶ **To cite this version:**

Matias Rafael Bender. Algorithms for sparse polynomial systems : Gröbner bases and resultants. Symbolic Computation [cs.SC]. Sorbonne Université, 2019. English. NNT : 2019SORUS029 . tel-02935894

HAL Id: tel-02935894
https://theses.hal.science/tel-02935894v1

Submitted on 10 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT DE
SORBONNE UNIVERSITÉ

Spécialité

Informatique

École doctorale Informatique, Télécommunications et Électronique (Paris)

Présentée par

Matías Rafael BENDER

Pour obtenir le grade de

DOCTEUR de SORBONNE UNIVERSITÉ

Sujet de la thèse :

Algorithms for sparse polynomial systems : Gröbner bases and resultants

Contents

Chapter 1

Introduction

This thesis is about different aspects of effective algorithms that exploit structure and sparsity to solve systems of polynomial equations.

Solving polynomial systems is one of the oldest and most important problems in computational mathematics and has many applications in several domains of science and engineering [Stu02, SW05]. It is an intrinsically hard problem with complexity at least single exponential in the number of variables [HM93]. However, in most of the cases, the polynomial systems coming from applications have some kind of structure. For example, several problems in biology [PS05, SFR14], chemistry [Mor09, Ch. 9], computational geometry [BT06, Can88], control theory [GHJZ14], cryptography [KS99, FJ03], game theory [Jud98], optimization [Las00], statistics and learning [CS02, DSS09], and verification and theorem proving [DMB08] involve polynomial systems that are structured. In this thesis we focus on exploiting the structure related to the sparsity of the supports of the polynomials; that is, we exploit the fact that the polynomials only have a few monomials with non-zero coefficients. Our objective is to solve the systems faster than the worst case estimates that assume that all the terms are present. We say that a sparse system is *unmixed* if all its polynomials have the same Newton polytope, and *mixed* otherwise. Most of the work on solving sparse systems concern the unmixed case, with the exceptions of mixed sparse resultants [ER94, Emi96], geometric resolution algorithms [HJS13] and homotopy methods [VVC94, HS95]. In this thesis, we develop algorithms for *mixed* systems. We use two prominent tools in nonlinear algebra: *sparse resultants* and *Gröbner bases*. We work on each theory independently, but we also combine them to introduce new algorithms: we take advantage of the algebraic properties of the systems associated to a non-vanishing resultant to improve the complexity of computing their Gröbner bases; for example, we exploit the exactness of some strands of the associated Koszul complex to deduce an early stopping criterion for our Gröbner bases algorithms and to avoid every redundant computation (reductions to zero).

In addition, we introduce quasi-optimal algorithms to decompose binary forms; this is the simplest case of symmetric tensor decomposition, also known as, polynomial Waring's problem.

1.1 Previous work

In the sequel we overview some of the most important results on *sparse resultants*, *Gröbner bases*, *solving sparse polynomial systems*, and *tensor decomposition*.

Sparse resultant. One of the main questions in (computational) algebraic geometry is to decide efficiently when an overdetermined polynomial system has a solution over a projective variety. The resultant provides an answer to this question. It is a multihomogeneous polynomial in the coefficients of the polynomials of the system which vanishes if and only if the system has a solution. We can also use it to solve square systems. In the case of sparse polynomials, there exists an analogous concept called the *sparse resultant* [GKZ08]. This object generalizes the resultant to the context of projective toric varieties [CLS11, Ch. 2]. The sparse resultant is one of the few tools we can use to solve systems taking into account the sparsity of the polynomials. Hence, its efficient computation is fundamental in computational algebraic geometry.

In particular, we are interested in the computation of the multiprojective resultant, as it is defined in [Rém01, DKS13, DS15], which corresponds to sparse systems consisting of multihomogeneous polynomials. In general, we compute the resultant of a system as a quotient of determinants of matrices whose elements are polynomials in the coefficients of the input polynomials [Mac02, Jou91, DD01, D'A02, KS97, CK00, CK04]; thus the best we can hope for are linear polynomials. A classical example of such a matrix is the Macaulay matrix, which represents a map of the form $(g_0, \ldots, g_n) \mapsto \sum_i g_i f_i$. In this case, we say that we have a *Sylvester-type* formula. Other classical formulas include *Bézout-* and *Dixon-type*; nevertheless, the elements of the corresponding matrices are not linear anymore. We refer to [EM99b] and references therein for details.

When we can compute the resultant as the determinant of a matrix we say that we have a *determinantal formula*. Such a formula does not exists in general and it is an open question to decide when it exists. When we consider *unmixed* multihomogeneous systems, these formulas are well studied; see, for example, [SZ94, WZ94, CK00, DE03]. However, in the case of *mixed* multihomogeneous systems, there are very few results. We know determinantal formulas for scaled multihomogeneous systems [EM12], that is when the supports are scaled copies of one of them, for bivariate tensor-product polynomial systems [BMT17], and some special kind of determinantal systems, for example [Atk72, Ch. 8]. One way to obtain such formulas is using the Weyman complex [Wey94]. For an introduction to this complex we refer to [Wey03, Sec. 9.2] and [GKZ08, Sec. 2.5.C, Sec. 3.4.E].

Gröbner bases of structured systems. The introduction of the first algorithm to compute Gröbner bases in 1965 [Buc06] established them as a central tool in nonlinear algebra. Their applications span most of the spectrum of mathematics and engineering [BW98]. Computing Gröbner bases is an intrinsically hard problem. For many "interesting" cases related to applications, the complexity of the algorithms to compute them is single exponential in the number of variables, but there are instances where the complexity is double exponential; it is an EXPSPACE complete problem [May97], see also [GG13, Sec. 21.7]. There are many practically efficient algorithms, see [CLO15, Ch. 10] and references therein, for which, under genericity assumptions, we can deduce precise complexity estimates [BFS15]. However, the polynomial systems coming from applications have some kind of structure. One of the main challenges in Gröbner basis theory is to improve the complexity and the practical performance of the related algorithms by exploiting the structure. This problem was studied intensively in the last few years. Different examples of the systems for which there are specific algorithms to compute their Gröbner bases include bilinear systems [FSEDS11], critical point systems [FEDS12], systems invariant under the action of a symmetric [FS12] and commutative groups [FS13], quasi-homogeneous systems [FSEDV13], determinantal systems [FSEDS13], Boolean systems, [BFSS13], chordal structured sys-

tems [CP16], and sparse unmixed systems [FSS14]. Up to this thesis, there were no algorithms that could exploit the structure of *sparse mixed systems*.

An approach to exploit the sparsity of the polynomials is to compute Gröbner bases over semigroup algebras [Stu93, FSS14]. Semigroup algebras are related to toric varieties. An affine toric variety is the spectrum of a semigroup algebra [CLS11, Thm. 1.1.17]. Hence, the variety defined by the polynomials over a semigroup algebra is a subvariety of a toric variety. This variety is different from the one defined by the polynomials over the original polynomial algebra, but they are related and in many applications the difference is irrelevant; see, for example, [EM99a]. We refer to [CLS11] for an introduction to toric varieties and to [Stu96] for their relation with Gröbner basis.

Following [Stu93], Faugère et al. considered *sparse unmixed systems* and introduced an algorithm to compute Gröbner bases over the semigroup algebra generated by their Newton polytope [FSS14]. By embedding the systems in a semigroup algebra, they can predict the structure of regular sparse unmixed systems and exploit it algorithmically. Their algorithm is a variant of the `Matrix-F5` algorithm [Fau02, BFS15]. They homogenize the polynomials and compute a Gröbner basis, degree-by-degree, by performing Gaussian elimination on various Macaulay matrices [Laz83]. They use the `F5` criterion [Fau02] to avoid redundant computations, that is, to skip rows reducing to zero after performing Gaussian elimination. Once they have computed the Gröbner basis for the homogenized system, they recover a Gröbner basis of the original system by dehomogenizing the computed basis. The efficiency of this approach relies on an incremental degree-by-degree construction which, under regularity assumptions, skips all the rows reducing to zero. One of the properties that they exploit in this work is that, for normal Newton polytopes [CLS11, Def. 2.2.9], the homogenization of a *generic* unmixed system forms a regular sequence over the corresponding semigroup algebra [CLS11, Def. 9.2.9]. Unfortunately, this property is no longer true for *mixed systems*. So, for mixed systems, this algorithm fails to predict all rows reducing to zero during Gaussian elimination.

Solving sparse polynomial systems. There are different ways of solving sparse polynomial systems by exploiting their sparsity. These approaches include homotopy methods, for example [VVC94, HS95], chordal elimination [CP16], triangular decomposition [MB18], geometric resolutions [HJS13], and various other techniques [Roj99, Mas16, TMVB18].

Among the symbolic approaches related to toric geometry, the main tool to solve sparse systems is the *sparse resultant* [GKZ08]. We can use it to solve zero-dimensional square polynomial systems $(f_1, \ldots, f_n)$. For example we can hide a variable or use the technique of the u-resultant; we refer to [CLO06, Chp. 3 & Chp. 7.6] for a general introduction. When a *Sylvester-type* formula is available, through the resultant matrix, we obtain the matrix of the multiplication map of a polynomial f_0 in $\mathbb{K}[\boldsymbol{x}]/\langle f_1, \ldots, f_n\rangle$. Then, we solve the system $(f_1, \ldots, f_n)$ via the *eigenvalue* [Laz81] and *eigenvector criteria* [AS88, ER94]. These criteria relate the solutions of the system to the eigenvalues and eigenvectors of the matrix corresponding to the multiplication map of f_0: the eigenvalues correspond to the evaluations of f_0 at the zeros of the system, and, under some assumptions, from the eigenvectors we can recover the coordinates of the zeros. Canny and Emiris [CE93] and Sturmfels [Stu94] showed that we can compute the sparse resultant as the determinant of a square Macaulay matrix (Sylvester-type formula) whose rows are related to mixed subdivisions of some polytopes. Under genericity assumptions, we can use this matrix to solve square sparse systems by performing eigenvalue and eigenvector computations [ER94, Emi96]. Recently, Massri [Mas16] dropped the genericity assumptions by considering a bigger

matrix. In all these cases, the arithmetic complexity of the approaches depends on the number of integer points in the Minkowski sum of the Newton polytopes of the input polynomials.

Alternatively, we can use Gröbner bases over semigroup algebras to solve sparse systems [FSS14]. Unfortunately, even in the *unmixed case*, there are no complexity results for this approach, as the bounds in [FSS14] miss some assumptions to hold, see Sec. 8.3.1.

Multihomogeneous systems form an important subclass of mixed sparse systems as they are ubiquitous in applications, for example in kinematics [WS11] and in cryptography [KS99]. Their properties are well understood, for example, the degree (number of solutions) of the system [vdW78], the arithmetic Nullstellensätze [DKS13], and the (multigraded) Castelnuovo-Mumford regularity [HVT04, ACG05, SVT06, BC17]. We can solve these systems using general purpose algorithms based on resultants [ER94, Emi96] and in some cases benefit from the existence of determinantal formulas [SZ94, WZ94]. We can also use a straight-line program representation of the multihomogeneous resultant [JS07] or homotopy methods [MS87, HJSS02, HR17, SEDS17]. For regular unmixed bilinear systems, we can compute Gröbner bases without performing reductions to zero [FSEDS11].

Binary forms decomposition. Symmetric tensor decomposition is an important problem with applications in several areas, for example signal processing, statistics, data analysis and computational neuroscience [Com14]. It is equivalent to Waring's problem for homogeneous polynomials; that is, to write a homogeneous polynomial in n variables of degree D as a sum of D-th powers of linear forms, using the minimal number of summands. This minimal number is called the *rank* of the polynomial/tensor. Under this formulation, the problem finds its roots in invariant theory where the decompositions are related to canonical forms.

In this thesis, we focus on a particular case of this problem corresponding to decomposing binary forms, that is, to decompose symmetric tensors of dimension 2 and order D. In terms of homogeneous polynomials, we consider a binary form

$$f(x,y) := \sum\nolimits_{i=0}^{D} \tbinom{D}{i} a_i x^i y^{D-i}, \tag{1.1}$$

where $a_i \in \mathbb{K} \subset \mathbb{C}$ and $\mathbb{K}$ is some field of characteristic zero. We want to compute a decomposition

$$f(x,y) = \sum\nolimits_{j=1}^{r} (\alpha_j x + \beta_j y)^D, \tag{1.2}$$

where $\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_r \in \overline{\mathbb{K}}$ (the algebraic closure of $\mathbb{K}$) and r is minimal. We say that a decomposition *unique* if, for all the decompositions, the set of points $\{(\alpha_j, \beta_j) : 1 \leq j \leq r\} \subset \mathbb{P}^1(\mathbb{K})$ is unique, where $\mathbb{P}^1(\mathbb{K})$ is the projective space of $\overline{\mathbb{K}}$ [Rez13a].

Starting from Sylvester in the 19th century [Syl04a, Syl04b], the decomposition of binary forms (Eq. (1.2)) has been largely studied for $\mathbb{K} = \mathbb{C}$. Sylvester described the necessary and sufficient conditions for a decomposition to exist. He related the decompositions to the kernel of Hankel matrices. For a modern approach of this topic, we refer to [KR84, Kun90, Rez13a, IK99].

From the algorithmic point of view, Sylvester's work leads to an algorithm to decompose binary forms, see [CM96, Sec. 3.4.3]. In the case where the binary form is of odd degree, then we can compute the decompositions using Berlekamp-Massey algorithm [Dü89]. When the decomposition is unique, the Catalecticant algorithm, which also works for symmetric tensors of bigger dimension [IK99, OO13],

improves Sylvester's work. For an arbitrary binary form, [Hel92] presented a randomized algorithm based on Padé approximants and continued fractions, in which he also characterized the different possible decompositions. Unfortunately, all these algorithms have complexity at least quadratic in the degree of the binary form.

Besides the problem of computing the decomposition(s) many authors considered the subproblems of computing the rank and deciding where there exists a unique decomposition, e.g., [Syl04a, Syl04b, Hel92, CS11, BGI11]. For example, [Syl04a, Syl04b] considered generic binary forms, that is binary forms with coefficients belonging to a dense algebraic open subset of $\overline{\mathbb{K}}^{D+1}$ [CM96, Sec. 3], and proved that when the degree is $2k$ or $2k+1$, the rank is $k+1$ and that the minimal decomposition is unique only when the degree is odd. In the non-generic case, [Hel92, CS11, IK99], among others, proved that the rank is related to the kernel of a Hankel matrix and that the decomposition of a binary form of degree $2k$ or $2k-1$ and rank r, is unique if and only if $r \leq k$. With respect to the rank, different authors, for example, [CS11, CGLM08, BGI11], proposed algorithms to compute its value. They showed that the rank of a tensor can have only two values t or $D-t+2$, where t is the rank of a Hankel matrix. Even though the authors do not provide complexity estimates, using recent superfast algorithms for Hankel matrices [Pan01], we can deduce a nearly-optimal arithmetic complexity bound for the approach of [CS11].

For the general problem of symmetric tensor decomposition, Sylvester's work was successfully extended to cases in which the decomposition is unique [BCMT10, OO13]. For example, [BCMT10] reduces the problem to find the generators of a linear recursive multidimensional sequence. There are several ways to do so as Gröbner-basis-based methods like the Berlekamp-Massey-Sakata algorithm [CLO06, Ch. 10] and `Scalar-FGLM` [BBF17], and border-basis-based methods [Mou17].

Besides tensor decomposition, there are other related decompositions for binary forms and univariate polynomial that we do not consider, for example, [Gun87, Rez96, IK99, Rez13b, GKL03, GR10, GMKP17].

1.2 Contributions

The contributions of this thesis are along the directions of the previous work that we just introduced, that is, *sparse resultants*, *Gröbner bases*, *solving sparse polynomial systems*, and *tensor decomposition*.

Sparse resultant. In Chapter 6, we consider mixed multilinear polynomial systems. On the one hand, this is the simplest case of mixed multihomogeneous systems where no determinantal formula was known. On the other hand, multilinear polynomial systems are common in applications, for example in the *Multiparameter Eigenvalue Problem* related to mathematical physics [Atk72, Vol88].

In the first part of the chapter, Section 6.1, we study determinantal formulas for the multiprojective resultant of *mixed multilinear polynomial systems*. Following [WZ94], we use the Weyman complex, see Def. 3.3.3, to introduce determinantal formulas for two kinds of mixed multilinear systems. If $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_A$ and $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_B$ are $(A+B)$ different blocks of variables, then we consider the following mixed multilinear systems $(f_1, \ldots, f_n)$:

1. **Star multilinear systems**: For each polynomial f_k there is $1 \leq j_k \leq B$ such that

$$f_k \in \mathbb{K}[\boldsymbol{X}_1]_1 \otimes \cdots \otimes \mathbb{K}[\boldsymbol{X}_A]_1 \otimes \mathbb{K}[\boldsymbol{Y}_{j_k}]_1.$$

2. **Bipartite bilinear systems**: For each polynomial f_k there are $1 \leq i_k \leq A$ and $1 \leq j_k \leq B$ such that
$$f_k \in \mathbb{K}[\boldsymbol{X}_{i_k}]_1 \otimes \mathbb{K}[\boldsymbol{Y}_{j_k}]_1.$$

We add an additional polynomial f_0, linear or multilinear, and show that the resultant of the new system is the determinant of a *Koszul resultant matrix* (related to the maps in the Koszul complex, Prop. 3.3.13). As the size of the matrix, that is, the degree of the resultant, depends on the multidegree of f_0, we derive determinantal formulas for different choices of f_0. We relate the size of the matrices to the number of solutions of $(f_1, \ldots, f_n)$ (Sections 6.1.2 and 6.1.3). For example, in Lem. 6.1.12 we prove that, if we consider a square *star multilinear system* $(f_1, \ldots, f_n)$ having Υ solutions and we introduce a multilinear $f_0 \in \mathbb{K}[X_1]_1 \otimes \cdots \otimes \mathbb{K}[X_A]_1$, then the size of the matrix associated to the determinantal formula of $(f_0, f_1, \ldots, f_n)$ is
$$\Upsilon \cdot \left(\sum_{i=1}^{A} \#X_i - A + 1 \right).$$
Moreover, Υ is bigger than $(\sum_i \#X_i - A + 1)$ and so, the size of the formula is *polynomial in the number of solutions*.

In the second part of the chapter, Section 6.3, we exploit the structure of these *Koszul resultant matrices* to solve square *star multilinear systems* and *bipartite bilinear systems*. For that, we generalize the classical eigenvalue criterion for multiplication maps [Laz81], see Prop. 5.2.3. Our extension applies to a general class of matrices (Def. 6.2.1), including the Sylvester and Koszul resultant matrices. Moreover, it relies only on the degree and structure of the associated formula. We prove that if the matrix $\left[\begin{smallmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{smallmatrix}\right]$ corresponds to a determinantal formula for $(f_0, f_1, \ldots, f_n)$ such that the diagonal of the matrix $M_{2,2}$ corresponds to the coefficient of the monomial $\boldsymbol{x}^{\boldsymbol{\theta}}$ in f_0, this coefficient only appears in this diagonal, and the system $(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1, \ldots, f_n)$ has no solutions, then the matrix $M_{1,1}$ is invertible and eigenvalues of the Schur complement of $M_{2,2}$, $M_{2,2}^c := M_{2,2} - M_{2,1}\, M_{1,1}^{-1}\, M_{1,2}$, correspond to the evaluation of $\frac{f_0}{\boldsymbol{X}^{\boldsymbol{\theta}}}$ at every solution of $(f_1, \ldots, f_n)$. That is,
$$\lambda \text{ is an eigenvalue of } M_{2,2}^c \iff \exists\, \boldsymbol{\alpha} \text{ such that } \begin{cases} f_1(\boldsymbol{\alpha}) = \cdots = f_n(\boldsymbol{\alpha}) = 0 \text{ and} \\ \lambda = \frac{f_0}{\boldsymbol{X}^{\boldsymbol{\theta}}}(\boldsymbol{\alpha}) \end{cases}.$$

In the third part of the chapter, Section 6.3, we extend the classical eigenvector criterion [AS88], see Prop. 5.2.4, to the case of *Koszul resultant matrices* for *2-bilinear systems*. These systems correspond to the star multilinear systems where $A = 1$ and $B = 2$.

Finally, in Section 6.4, we merge all the tools that we introduce in the chapter to propose a new algorithm to solve the *Multiparameter Eigenvalue Problem*. The complexity of our approach is *polynomial in the number of solutions*.

Gröbner basis and mixed sparse polynomial systems In Chapter 7, we study extensions of the algorithms to compute Gröbner bases for unmixed sparse systems [FSS14] to the mixed case. We present algorithms which, under regularity assumptions, perform no reductions to zero.

- Our first extension (Section 7.3) changes the *degree-by-degree* strategy (Sec. 4.4) of the algorithm proposed in [FSS14] by a *polynomial-by-polynomial* strategy, that is, we first compute a Gröbner

basis for the ideal generated by the first i polynomials and then we extend this basis to a Gröbner basis for the ideal generated by first $(i+1)$ polynomials. In the language of signature-based algorithms, we change the module monomial order of `F5` from $<_{\text{d-pot}}$ (degree-by-degree) to $<_{\text{pot}}$ (polynomial-by-polynomial), see [EF17]. This strategy avoids every reduction to zero when the original (non-homogenized) system is a regular sequence. Unfortunately, this approach requires to compute using a $\mathrm{GRevLex}$ monomial orderings which, as we show in Ex. 7.2.1, we can not define for semigroup algebras. Hence, we introduce a novel Gröbner-like basis, that we call *scant Gröbner basis*. A *scant Gröbner basis* is a basis for an ideal over a semigroup algebra with similar properties to the usual Gröbner basis. Their main advantage is that they allow us to define $\mathrm{GRevLex}$-like orderings over semigroup algebras with many of the expected properties, see Sec. 4.3. To define *scant Gröbner basis*, we need to work with non-monomial-orderings and redefine the monomial division relation. Hence, a *scant Gröbner basis* is not Gröbner basis over a semigroup algebra. We introduce an algorithm to compute *scant Gröbner basis* for *unmixed sparse systems* which, under regularity assumptions, performs no reductions to zero.

- Our second extension (Section 4.6) computes a Gröbner basis over a multigraded semigroup algebra; the multigrading is related to the different polytopes of the sparse polynomials. Even though we embed the system in the multigraded semigroup algebra, the straightforward homogenization of the input polynomials never results in a regular sequence. Therefore, the existing criteria do not avoid all the trivial (expected) reductions to zero. Hence, to avoid all the trivial reductions to zero, we extend the classical `F5` criterion (Prop. 4.4.6) by using the exactness of the strands of the *Koszul complex*. We introduce the concept of *(sparse) regularity*, related to the exactness of these strands, to guarantee that all the reductions to zero are trivial. We present the first algorithm that computes Gröbner bases over these multigraded semigroup algebras which, under (sparse) regularity assumptions, performs no reductions to zero.

We emphasize that besides the similarity in their name, scant Gröbner bases and Gröbner bases over semigroup algebras are completely different objects. In particular, scant Gröbner bases *are not* Gröbner basis over semigroup algebras. Moreover, the corresponding algorithms follow different computational strategies: to compute scant Gröbner bases, we proceed polynomial-by-polynomial, while to compute Gröbner bases over semigroup algebras, degree-by-degree. The assumptions for the algorithms to perform no reductions to zero are also different, but both of them are satisfied under (sparse) regularity assumptions. Last but not least, we can use both objects to solve sparse systems, but we do not have complexity bounds when we use scant Gröbner bases.

Solving sparse polynomial systems We can use both *scant Gröbner basis* and *Gröbner basis over semigroup algebras* to compute normal forms and so, to solve sparse zero-dimensional system. In Chapter 8, we introduce complexity bounds for solving sparse polynomial systems using our algorithm to compute *Gröbner basis over semigroup algebras*. We do not discuss how to solve sparse polynomial systems using *scant Gröbner basis*; nevertheless, we can deduce straightforwardly an algorithm from the variant of the `FGLM` algorithm (Alg. 5) proposed in [FSS16, Sec. 4.2]. Unfortunately, we have not bounds for the arithmetic complexity of this approach, let alone bounds depending on the Newton polytopes.

We build on [ER94, Emi96, Mas16] and, under some assumptions (Ass. 8.1.1), we propose an algorithm to solve zero-dimensional square systems. Because we work with toric varieties, we only compute

the solutions lying in $(\mathbb{C} \setminus \{0\})^n$. The arithmetic complexity of our algorithm is polynomial in the number of integer points in the Minkowski sum of the Newton polytopes. Our strategy is to reuse part of our algorithm to compute *Gröbner bases over semigroup algebras* (Sec. 7.3) to compute multiplication maps and, via `FGLM` [FGLM93], recover a Gröbner basis over the standard polynomial algebra $\mathbb{K}[\boldsymbol{x}]$. As we compute the solutions over $(\mathbb{C} \setminus \{0\})^n$, we do not recover a Gröbner basis for the original ideal, but for its saturation with respect to the product of all the variables. Our algorithm relies on ideas from resultant theory to avoid the computation of a Gröbner basis. Instead, we compute a part of the Gröbner basis which suffices to solve the systems. We compute with a matrix that has the same size as the one in resultant-based approaches [ER94, Emi96]. Hence, the complexity of our algorithm is similar to the one of the resultant-based approaches; however, we do not compute a mixed subdivision and we rely on weaker assumptions which, in addition, are geometric. Moreover, in general, Gröbner-type algorithms can be extended without any modification to the overdetermined case.

We introduce an algorithm to solve *mixed sparse systems* (Sec. 8.1) and two variants for two specific subfamilies: *mixed multihomogeneous systems* (Sec. 8.2) and *unmixed systems* (Sec. 8.3). The three algorithms are similar, but they are not the same. In what follows we present a rough description of the algorithms, and we refer the reader to the respective sections for details. We consider $\mathrm{NP}(f_1), \ldots, \mathrm{NP}(f_n)$ as the Newton polytopes of the input polynomials $f_1, \ldots, f_n \in \mathbb{K}[\boldsymbol{x}]$ and $\mathrm{MV}(\mathrm{NP}(f_1), \ldots, \mathrm{NP}(f_n))$ as their mixed volume.

1. Let $f_1, \ldots, f_n \in \mathbb{K}[\boldsymbol{x}]$ be a (sparse) regular polynomial system with a finite number of solutions over $(\mathbb{C}^*)^n$ that equals the BKK bound, that is, $\mathrm{MV}(\mathrm{NP}(f_1), \ldots, \mathrm{NP}(f_n))$ (Thm. 2.12.19).

2. Embed the polynomials in a multigraded semigroup algebra $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ related to the Newton polytopes $\mathrm{NP}(f_1), \ldots, \mathrm{NP}(f_n)$, see Def. 7.1.1.

3. For each variable x_i:

   - Use our algorithm to compute Gröbner bases over semigroup algebras (Alg. 10) to construct a square Macaulay matrix related to $(f_1, \ldots, f_n, x_i)$ at a degree $\boldsymbol{d}$, where $\boldsymbol{d}$ is, roughly speaking, a bound on the Castelnuovo-Mumford regularity of the system.
   - Split the matrix in four parts and compute a Schur complement. The latter is the multiplication map of x_i in $\mathbb{K}[\boldsymbol{x}^{\pm 1}]/\langle f_1, \ldots, f_n\rangle$.

4. Use the multiplication maps and `FGLM` to get a Gröbner basis for $\langle f_1, \ldots, f_n\rangle : \langle \prod_i x_i\rangle^\infty$ with respect to any monomial order; in particular, with respect to a lexicographical monomial ordering.

5. Use the Gröbner basis in lexicographical ordering to solve the system.

Let $P(\boldsymbol{d})$ be the number of monomials in $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$, that is, the number of monomials of multidegree $\boldsymbol{d}$ in the semigroup algebra $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$. The complexity of our approach depends on $P(\boldsymbol{d})$ and $\mathrm{MV}(\mathrm{NP}(f_1), \ldots, \mathrm{NP}(f_n))$ (Thm. 8.1.14): we can solve a sparse system in

$$O(2^{n+1}\, P(\boldsymbol{d})^{2.373} + n\, \mathrm{MV}(\mathrm{NP}(f_1), \ldots, \mathrm{NP}(f_n))^3) \text{ arithmetic operations.}$$

We instantiate the previous complexity bound according to the three families of polynomial systems that we consider.

- In Sec. 8.1, we consider *mixed sparse systems*. We construct an algebra $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ graded by each polytope. The complexity of solving this sparse system (Thm. 8.1.14) is polynomial in the the number of integer points in the Minkowski sum of the Newton polytopes $\mathrm{NP}(f_1), \ldots, \mathrm{NP}(f_n)$ and the n-simplex Δ_n, that is,

$$P(\boldsymbol{d}) = \#\Big((\Delta_n + \sum_{j=0}^{n} \mathrm{NP}(f_j)) \cap \mathbb{Z}^n\Big).$$

  This bound agrees with the one obtained using sparse resultant methods [ER94, Emi96].

- In Sec. 8.2, we consider *mixed multihomogeneous systems*. Using bounds for the multigraded Castelnuovo-Mumford regularity, we obtain a better complexity bound for our algorithm. We consider q blocks of variables $\boldsymbol{x_1}, \ldots, \boldsymbol{x_q}$, the standard $\mathbb{Z}^q$-graded algebra $\mathbb{K}[S^h_{\boldsymbol{\Delta}}] = \mathbb{K}[\boldsymbol{x_1}] \otimes \cdots \otimes \mathbb{K}[\boldsymbol{x_q}]$, and multihomogeneous polynomials $(f_1, \ldots, f_n)$ of multidegrees $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n \in \mathbb{N}^q$, respectively. Then, the complexity of our approach (Thm. 8.2.8) depends on

$$P(\boldsymbol{d}) = \#\{\boldsymbol{X}^\alpha \in \mathbb{K}[\boldsymbol{x_1}, \ldots, \boldsymbol{x_q}]_{\boldsymbol{d}}\}, \text{ where } \boldsymbol{d} = \sum_{i=1}^{n} \boldsymbol{d}_i - (\#\boldsymbol{x_1}, \ldots, \#\boldsymbol{x_q}) + (1, \ldots, 1),$$

  that is, on the number of monomials with multidegree $\sum_{i=1}^{n} \boldsymbol{d}_i - (\#\boldsymbol{x_1}, \ldots, \#\boldsymbol{x_q}) + (1, \ldots, 1)$. Our bound generalizes the classical *Macaulay bound*, which governs the complexity of solving zero-dimensional square homogeneous systems via Gröbner bases [Laz83]. The Macaulay bound corresponds to the case $q = 1$.

- In Sec. 8.3, we consider *unmixed sparse systems*. We consider the $\mathbb{Z}$-graded algebra $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ associated to the polytope Δ and a homogeneous regular sequence $f_1, \ldots, f_n$ such that $f_i \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{d_i}$. We consider r as the smallest integer such that $r \cdot \Delta$ contains an integer interior point and we fix $d := (\sum_{i \geq 1} d_i) - r + 1$. Then, the complexity of our algorithm (Thm. 8.3.5) is polynomial in the number of integer points in the d homothety of Δ, that is,

$$P(d) = \#\Big(d\,\Delta \cap \mathbb{Z}^n\Big).$$

  This complexity agrees with the claimed complexity of the algorithm in [FSS14]. Nevertheless, as we show in Sec. 8.3.1, their claimed complexity bound is not always correct for their algorithm, but it is for ours.

Decomposition of binary forms In chapter 9, we introduce a new *superfast* algorithm that improves the complexity of previous approaches to decompose binary forms. For that, we use results from *structured linear algebra*. It achieves a *softly linear*, and so *quasi-optimal*, arithmetic complexity bound. To the best of our knowledge, the previously known algorithms have at least quadratic complexity bounds. Our algorithm computes a symbolic decomposition in $O(\mathtt{M}(D)\log(D))$ arithmetic operations (Thm. 9.4.19), where $\mathtt{M}(D)$ is the complexity of multiplying two polynomials of degree D. It is deterministic when the decomposition is unique. When the decomposition is not unique, our algorithm is

randomized. We present a Monte Carlo version of it and we show how to modify it to a Las Vegas one, within the same complexity.

From the symbolic decomposition, we approximate the terms of the decomposition with an error of $2^{-\varepsilon}$ in $O\big(D\log^2(D)\big(\log^2(D)+\log(\varepsilon)\big)\big)$ arithmetic operations (Thm. 9.4.20). Moreover, we bound the algebraic degree of the problem by $\min(rank, D-rank+1)$ (Thm. 9.4.15). We show that this bound can be tight. When the input polynomial has integer coefficients, our algorithm performs, up to poly-logarithmic factors, $\widetilde{\mathcal{O}}_B(D\ell + D^4 + D^3\tau)$ bit operations (Thm. 9.4.21), where τ is the maximum bitsize of the coefficients and $2^{-\ell}$ is the relative error of the terms in the decomposition.

1.3 Conclusions and perspectives

Besides the technical contributions, our work introduces and exploits novel directions to study the underlying algebraic objects and the corresponding applications.

- *We introduce determinantal formulas for mixed multilinear systems.* Determinantal formulas do not exist in general, and so they are hard to find. Using the Weyman complex, we construct Koszul-type determinantal formulas for some families of mixed multilinear systems. These families are related to particular kinds of (hyper)graphs: star and bipartite. Roughly speaking, we construct these (hyper)graphs by considering their nodes as the blocks of variables and the edges as the different supports of the polynomials. We have some partial understanding about the relation between the combinatorics of the (hyper)graphs and the existence of determinantal formulas coming from the Weyman complex. We believe that we can generalize this relation.

- *We extend the classical eigenvalue and eigenvector criteria.* We introduce a new linear algebra approach to solve polynomial systems relying on Koszul-type determinantal formulas. Our eigenvalue criterion is independent of the kind of the formula, for example, Sylvester, Koszul, etc. It relies only on the degree and structure of the formula. In contrast, we present our eigenvector criterion to solve only 2-bilinear systems. Nevertheless, our experiments indicate that we can extend this criterion to general Koszul-type formulas. Our results motivate the following three questions:

  - Besides Sylvester- and Koszul-type formulas, we can use the Weyman complex to derive other determinantal formulas of bigger degree as Bézout-type and hybrid formulas; see for example [DE03]. Up to now, there are not linear algebra approaches to solve the systems using such formulas. It worth to explore if we can extend our approach to solve polynomial systems using Koszul-type formulas to this context.
  - When Sylvester-type formulas are available, we can use the resultant matrices to solve polynomial systems by computing multiplication maps [Laz77, AS88, ER94]. As a by-product of our eigenvalue criterion, we show how to construct a matrix which is similar to a multiplication map, that is, there is a change of coordinates for our matrix that transforms it into a multiplication map. Unfortunately, we do not know how to construct such a change of coordinates without solving the system first. This is why, our eigenvector criterion studies ad-hoc the structure a Koszul-type formula. It worth to explore if there is a universal approach to recover directly multiplication maps from the Koszul resultant formulas.

  - Sylvester-type formulas have been studied intensively in structured linear algebra to improve the complexity of solving polynomial systems, see for example [MP00, EP02], and in elimination theory [Vil18]. As Koszul-type formulas are closely related to Sylvester-type formulas, it worth to explore how we can exploit their structure to improve the arithmetic complexity of solving mixed sparse systems.

- *We propose a new algorithm to solve the Multiparameter Eigenvalue Problem.* It is a symbolic algorithm. It worth to study its practical efficiency and to compare it to the state-of-the-art symbolic and numerical methods, for example MATLAB's library `MultiParEig` [PMH18].

- *We introduce the first effective algorithm to compute Gröbner bases over semigroup algebras associated to mixed polynomial systems.* We generalize the work of [FSS14] to the mixed case and propose a `MatrixF5`-like algorithm which, under (sparse) regularity assumptions, performs no reduction to zero. As the classical `MatrixF5` algorithm [Fau02], and also as [FSS14], we depend on a degree bound to stop our computations. In the thesis, we deduce these degree bounds from the regularity assumptions. An open question, both in [FSS14] and in our work, is to propose a stopping criterion not depending on regularity assumptions. We believe that an extension of Buchberger's S-polynomial criterion is possible. Moreover, we expect that we can adapt our algorithm to follow a critical pairs approach as other variants of `F5`, see [EF17].

- *We relate the solving techniques using Sylvester-type formulas in resultant theory with Gröbner bases computations.* The simplest, but not necessarily the most efficient as there are more compact formulas [WZ94, SZ94], way to compute the resultant is to use a Sylvester-type formula and compute it as the determinant of a Macaulay matrix [CLO06, Chp. 3.4]. Using this matrix we extract multiplication maps and solve polynomial systems. In the standard polynomial algebra, such matrices are at the heart of some linear algebra algorithms to compute Gröbner bases, as `Matrix-F5` [Fau02, BFS15], because they correspond to the biggest matrix that appears during Gröbner basis computations for regular zero-dimensional systems [Laz83]. However, such a relation was not known for the sparse case. We bring out this relation and we build on it algorithmically.

- *We generalize the* `F5` *criterion to depend on Koszul complexes instead of regular sequences.* The exactness of the Koszul complex is closely related to regular sequences [Eis04, Ch. 17] and, geometrically, to complete intersections. Roughly speaking, when we consider generic square systems in the coordinate ring of a "nice" projective variety, then the variety that the system defines is closely related to a complete intersection. In this case, the Koszul complex of the system might not be exact in general, but only in some "low" degrees. Hence, even if the system is not a regular sequence, by focusing on the degrees at which the strands of the Koszul complex are exact, we can still predict the algebraic structure of the system and perform efficient computations. Using this property, we extend the classical `F5` criterion that applies only to regular sequences. Moreover, additional information on the exactness of the strands of the Koszul complex and the multigraded Castelnuovo-Mumford regularity [MS04, BC17] results in better degree and complexity bounds; similarly to the case of the multihomogeneous systems, see Sec. 8.2, or unmixed systems, see Sec. 8.3.

- *Working with Gröbner bases over semigroup algebras introduces new computational challenges.*

Graded reverse lexicographical monomial orderings ($\mathrm{GRevLex}$) are ubiquitous in (standard) Gröbner bases theory. Computing Gröbner bases with respect to $\mathrm{GRevLex}$ gives us a lot of algebraic and geometric information about the ideal. Moreover, in general, it is faster to compute a Gröbner bases with respect to $\mathrm{GRevLex}$ than with respect to any other monomial ordering; see for example [BM92] and references there in. In addition, in generic coordinates, the complexity of computing Gröbner bases with respect to $\mathrm{GRevLex}$ for a homogeneous ideal is governed by its Castelnuovo-Mumford (CM) regularity [BS87a]. More precisely, the maximal degree of a polynomial in a reduced Gröbner basis is the CM regularity. Roughly speaking, the CM regularity is the last degree on which our ideal can behave in a "strange way". It is related to the vanishing of the local cohomology and to the minimal free resolution of the ideal (Betti numbers) [Cha03]. In the context of resultants, the size of a Sylvester-type formula is related to the CM regularity of a generic overdetermined system; see for example [Bus06] and references therein. This relation explains why the complexities of solving a generic system using resultants or Gröbner bases are similar.

For most of the semigroup algebras, we cannot define $\mathrm{GRevLex}$-like monomial orderings, see Ex. 7.2.1. Hence, it is not clear that we can find a monomial ordering such that the complexity of computing Gröbner bases over the semigroup algebra is governed by the CM regularity of the ideal, see Sec. 8.3.1. Over the standard polynomial algebra, the difference between the maximal degree in a Gröbner basis and the CM regularity can be huge; see discussion about the Mayr & Meyer example [MM82] and the work of Bayer & Stillman [BS87a] in [GG13, Sec. 21.7]. This, together with the fact that it is not straightforward how to perform a generic linear change of coordinates over a semigroup algebra, suggest that computing Gröbner bases over semigroup algebras could be extremely hard, even harder than in the standard case. In particular, solving sparse polynomial systems using Gröbner bases over semigroup algebras might be harder than using sparse resultants. Another approach is needed to solve zero-dimensional systems; we present such an approach.

- *We solve sparse systems by truncating our computations of Gröbner bases.* The classical approach for solving zero-dimensional systems using Gröbner bases involves the computation of an intermediate Gröbner basis, usually with respect to $\mathrm{GRevLex}$, that we use to deduce the multiplication maps. By using `FGLM`, we obtain the lexicographical Gröbner basis of the ideal. If the intermediate Gröbner basis is computed with respect to $\mathrm{GRevLex}$ and the input system "behaves well" when we homogenize it, this strategy is in some sense optimal because it is related to the Castelnuovo-Mumford regularity of the homogenized ideal [Cha03, Cor. 3], see Sec. 4.5.

  However, over semigroup algebras, it might not be always possible to relate the complexity of the intermediate Gröbner basis computation to the Castelnuovo-Mumford regularity of the ideal. We overcome this obstacle by truncating the computation of the intermediate Gröbner basis over semigroup algebras in such a way that the complexity is given by Castelnuovo-Mumford regularity of the ideal. This way, our strategy has a similar complexity to the sparse resultant methods.

- *We do not know how the complexity bounds to solve sparse polynomial systems and the number of solutions of the systems relate.* An open challenge when we solve polynomial systems is to have an algorithm whose complexity is polynomial in the number of solutions of the system. For techniques involving (classical) resultants or Gröbner bases, we know such an algorithm only under

regularity assumptions, see for example [BFS15, Thm. 2]. In the sparse setting, the complexity bounds that we present in Chapter 8 and the ones for solving polynomial systems using the sparse resultant [Emi96, Ch. 9] depend on the number of integer points in the Minkowski sum of the input Newton polytopes, meanwhile the number of solutions of the input system is the mixed volume of these polytopes. In some cases, the difference between these two numbers is exponentially big and so the bounds of the algorithms are not polynomial in the number of solutions, see [Emi96, Ch. 8]. It worth to explore for which families of Newton polytopes we can get a polynomial bound for the number of integer points in the Minkowski sum of these polytopes with respect to their mixed volume. For these families, our algorithm is *polynomial in the number of solutions*.

- *We introduce scant Gröbner bases.* Before introducing our algorithm to compute Gröbner bases over semigroup algebras, our Gröbner-based approach to deal with sparse systems was to compute scant Gröbner bases. They generalize the classical definition of Gröbner basis in such a way that we can consider $\mathrm{GRevLex}$-like orderings over semigroup algebras. Our motivation was to use them to solve mixed sparse systems. Unfortunately, as we do not have any complexity result for our algorithm, in this thesis we could not study the complexity of solving sparse systems using this technique. The reason we present scant Gröbner bases in the thesis is because the techniques that we use to develop them lead to a general framework to construct Gröbner-like bases, that is, objects that behave like Gröbner bases in some sense, but they are not restricted to monomial orderings. Our key idea is to relate these objects to Gröbner bases over a standard polynomial algebra. We suspect that these techniques could be useful to develop a Gröbner-like basis such that, given a positive dimensional ideal, we can compute this basis in a complexity related to the Castelnuovo-Mumford regularity of the ideal.

- *We decompose binary forms in quasi-linear time.* We present the first algorithm to decompose this forms in *quasi-optimal* arithmetical complexity. Our approach is restricted to Sylvester's algorithm, so it is not straightforward to extend it to decompose binary forms over their base field, for example, over the reals. We expect that our approach to study the algebraic degree of the problem could be useful for this task.

1.4 Organization of the thesis

The thesis is divided in two parts: *Preliminaries* and *Contributions*. The former part contains the background needed in the latter part, and it contains no original contributions.

Preliminaries

- **Chapter 2:** We introduce the *commutative algebra* and *algebraic geometry* background of the thesis.

- **Chapter 3:** We introduce the theory of *resultants* with a particular emphasis on the *projective* and *multiprojective* case. To compute the multiprojective resultant we introduce the *Weyman complex*.

- **Chapter 4:** We introduce some computational aspects of *Gröbner bases theory* that we extend later in the thesis.

- **Chapter 5:** We introduce *polynomial system solving* techniques involving Gröbner bases and resultants.

Contributions

- **Chapter 6:** This chapter contains our contributions to *resultant theory*.
  - **Section 6.1:** This section contains *determinantal formulas* for the multiprojective resultant of particular families of *mixed multilinear systems*.
  - **Section 6.2:** This section contains our generalization of the *eigenvalue criterion*.
  - **Section 6.3:** This section contains our generalization of the *eigenvector criterion* for *Koszul-type resultant formulas* of *2-bilinear systems*, that is, bilinear systems in three blocks of variables with two different supports.
  - **Section 6.4:** In this section, we use our formulas to solve the *Multiparameter Eigenvalue Problem*.
- **Chapter 7:** This chapter contains our contributions in *Gröbner bases theory*. We consider *mixed sparse systems* and introduce novel algorithms to compute basis of these systems which, under regularity assumptions, perform no reductions to zero.
  - **Section 7.2:** In this section, we introduce *scant Gröbner bases* and an algorithm to compute them.
  - **Section 7.3:** In this section, we introduce an algorithm to compute *Gröbner bases* over semigroup algebras.
- **Chapter 8:** In this chapter we use our algorithm to compute *Gröbner bases* over semigroup algebras to solve *mixed sparse systems*. We present complexity results for three families of systems.
  - **Section 8.1:** In this section, we consider *mixed sparse systems*.
  - **Section 8.2:** In this section, we consider *mixed multihomogeneous systems*.
  - **Section 8.3:** In this section, we consider *unmixed sparse systems*.
- **Chapter 9:** In this chapter we introduce a new algorithm to *decompose binary forms* and study its *arithmetic, Boolean and algebraic complexity*.

1.5 Publications

The contributions of Chapter 6 were a joint work with Jean-Charles Faugère, Angelos Mantzaflaris, and Elias Tsigaridas. Our results were presented in two papers:

- The contents of Sections 6.2 and 6.3 appeared in

  Matías R. Bender, Jean-Charles Faugère, Angelos Mantzaflaris, and Elias Tsigaridas. Bilinear Systems with Two Supports: Koszul Resultant Matrices, Eigenvalues, and Eigenvectors. In *Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*, ISSAC '18, pages 63–70, New York, NY, USA, 2018. ACM

- The contents of Sections 6.1 and 6.4 will appear in

  Matías R. Bender, Jean-Charles Faugère, Angelos Mantzaflaris, and Elias Tsigaridas. Determinantal formulas for families of mixed multilinear systems. In preparation, 2019

The contributions of Chapters 7 and 8 were a joint work with Jean-Charles Faugère and Elias Tsigaridas. Our results were presented in two papers:

- The contents of Sections 7.2 and 8.2 appeared in

  Matías R. Bender, Jean-Charles Faugère, and Elias Tsigaridas. Towards Mixed Gröbner Basis Algorithms: The Multihomogeneous and Sparse Case. In *Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*, ISSAC '18, pages 71–78, New York, NY, USA, 2018. ACM

  We warn the reader that what we call *scant Gröbner bases* in this thesis is called *sparse Gröbner bases* in this paper. We changed the name to prevent further confusion between this object and *Gröbner bases over semigroup algebras*.

- The contents of Sections 7.3, 8.1 and 8.3 will appear in

  Matías Bender, Jean-Charles Faugère, and Elias Tsigaridas. Gröbner Basis over Semigroup Algebras: Algorithms and Applications for Sparse Polynomial Systems. *arXiv:1902.00208 [cs]*, February 2019. Submitted

The contributions of Chapter 9 were a joint work with Jean-Charles Faugère, Ludovic Perret, and Elias Tsigaridas. Our results were presented in two papers:

- Matías R. Bender, Jean-Charles Faugère, Ludovic Perret, and Elias Tsigaridas. A Superfast Randomized Algorithm to Decompose Binary Forms. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '16, pages 79–86, New York, NY, USA, 2016. ACM

  I was awarded *ACM SIGSAM's Distinguished Student Author Award, ISSAC 2016* for this paper at the *International Symposium on Symbolic and Algebraic Computation*, ISSAC '16.

- Matías R. Bender, Jean-Charles Faugère, Ludovic Perret, and Elias Tsigaridas. A nearly optimal algorithm to decompose binary forms. Submitted, 2018

  This paper is an extended version of the conference paper above.

Part I

Preliminaries

Chapter 2

Commutative algebra and algebraic geometry

In this thesis we focus on solving polynomials systems. The branch of mathematics that study these systems is Commutative Algebra. We can also think the solutions of the systems as geometric objects given by points, curves, etc. This approach corresponds to what we know as Algebraic Geometry. Hence, we introduce some classical results from both worlds, that we will need during the thesis. Part of the results can be found in the introductory chapters of [Eis04] or [Har77]. For the rest of them, in each section we give precise references of chapters and books that we used to elaborate this material.

2.1 Notation

In this section we introduce some of the notation we use throughout the thesis. Let $\mathbb{K}$ be an algebraically closed field of characteristic zero. Given a $\mathbb{K}$-vector space V, let $\dim_{\mathbb{K}}(V)$ be its dimension V. Consider the set of variables $x_1, \ldots, x_n$, where $n \in \mathbb{N}$. Let $\mathbb{K}[\boldsymbol{x}]$ be the polynomial algebra over $\mathbb{K}$ generated by the variables $x_1, \ldots, x_n$. Consider the variable x_0, to which we will refer as the *homogenization variable*, and the $\mathbb{K}$-algebra $\mathbb{K}[\boldsymbol{x}][x_0] := \mathbb{K}[x_0, x_1, \ldots, x_n]$.

In the following, R will be a Noetherian commutative ring and M will be a (left) R-module. We denote by L an additive monoid isomorphic to $\mathbb{Z}^k$, for some k. When R, or M, is graded by a monoid L, we denote by R_m, respectively M_m, the abelian group in R, respectively M, whose grading is $m \in L$. We denote by $(M_\bullet, \delta_\bullet)$ the chain complex whose modules are $\{M_i\}_{i\in\mathbb{Z}}$ and its maps are $\{\delta_i : M_i \to M_{i-1}\}_{i\in\mathbb{Z}}$, that is,

$$M_\bullet : \cdots \to M_i \xrightarrow{\delta_i} M_{i-1} \to \ldots$$

When the modules in $(M_\bullet, \delta_\bullet)$ are L-graded, we denote by $(M_i)_m$ the m-graded part of M_i.

For a natural number $N \in \mathbb{N}$, we use the abbreviation $[N] = \{1, \ldots, N\}$.

2.2 Ideals and quotient rings

The results in this section come mainly from [CLO15, Ch. 1], [Eis04, Ch. 0], [AM69, Ch. 1 & Ch. 2] and [Lan02, Ch. II, Ch. III & Ch. IV].

Let R be a commutative ring. For example, $R = \mathbb{K}[\boldsymbol{x}]$.

Definition 2.2.1. *An ideal $I \subset R$ is a set of elements in R such that*

- $\forall f \in I, \forall g \in R,\ f\,g \in I$.
- $\forall f_1, f_2 \in I,\ f_1 + f_2 \in I$.

Given elements $f_1, \ldots, f_k \in R$, we define their ideal to be

$$\langle f_1, \ldots, f_k \rangle := \left\{ \sum_{i=1}^{k} g_i f_i : g_1, \ldots, g_k \in R \right\}.$$

An ideal of a ring is analog to a subspace of a vector space over a field. Both are closed under addition and multiplication, but in the case of the subspaces we multiply by scalars, and in the case of ideals, we multiply by polynomials.

The Noetherian rings are commutative rings where every ideal is finitely generated.

Definition 2.2.2 (Noetherian ring)**.** *A ring is* Noetherian *if it satisfies the* ascending chain condition*. That is, for any infinite sequence of ideals $(I_i)_{i\in\mathbb{N}}$ such that*

$$I_0 \subseteq I_1 \subseteq I_2 \subseteq I_3 \subseteq \ldots,$$

there is a k such that the sequence "stabilizes", that is,

$$I_k = I_{k+1} = I_{k+2} = I_{k+3} = \ldots.$$

Proposition 2.2.3. *[Eis04, Ex 1.1] If R is a Noetherian ring, then every ideal I is finitely generated, that is there are $f_1, \ldots, f_k \in R$ such that $I = \langle f_1, \ldots, f_k \rangle$.*

An important class of Noetherian rings consists of the polynomial rings over a field.

Proposition 2.2.4 (Hilbert's Basis Theorem)**.** *[AM69, Thm. 7.5] If the ring R is Noetherian, then the polynomial ring $R[x_1, \ldots, x_n]$ is Noetherian. In particular, as every field is Noetherian, the polynomial ring $\mathbb{K}[\boldsymbol{x}]$ is a Noetherian ring, and so every ideal $I \subset \mathbb{K}[\boldsymbol{x}]$ is generated by a finite set of polynomials.*

The following proposition presents various operations between ideals.

Proposition 2.2.5. *[AM69, Chp. 1] Given two ideals $I, J \subset R$, we define the following related ideals;*

- *Addition: $I + J := \{f + g : f \in I, g \in J\}$.*
- *Product: $I\,J := \langle \{f\,g : f \in I, g \in J\} \rangle$ and, for $k > 1$, $J^k := J\,J^{k-1}$.*
- *Intersection: $I \cap J$.*

- *Radical of an ideal:* $\sqrt{I} := \{f \in R : \exists k \in \mathbb{N} \text{ s.t. } f^k \in I\}$.
- *Quotient (or Colon) ideal:* $I : J := \{f \in R : f \cdot J \subset I\}$.
- *Saturation:* $I : J^\infty := \{f \in R : \exists k \in \mathbb{N} \text{ s.t. } f \cdot J^k \subset I\}$.

Given an ideal I in R, we can define a new ring R/I, that we call quotient ring, which "erases" from R the elements in I.

Definition 2.2.6 (Quotient ring)**.** *Given an ideal I over R, we define an equivalence relation over R, $\sim_I$, where for each $f, g \in R$, $f \sim_I g$ if and only if $f - g \in I$. Also, for each $f \in R$, we define the coset $[f] := \{g \in R : f \sim_I g\}$. This equivalence relation is compatible with the addition and multiplication. That is, for every $f, g \in R$, $[f] + [g] = \{f' + g' : f' \in [f], g' \in [g]\} = [f + g]$ and $[f]\,[g] = \{f'\,g' : f' \in [f], g' \in [g]\} = [f\,g]$. Hence, we define the quotient ring R/I as the ring given by the sets of cosets of R together with the addition and multiplication.*

There is a natural way of thinking the ideals $J \subset R/I$. Consider the map $\phi : R \to R/I$ which maps each $f \in R$ to its coset $[f]$. This map is a surjective homomorphism.

Proposition 2.2.7. *[AM69, Prop. 1.1] There is a one-to-one order-preserving correspondence between the ideal $\bar{J}$ of R which contain I, and the ideal J of R/I, given by $\bar{J} = \phi^{-1}(J)$.*

The following "special" classes of ideals appear commonly throughout the thesis.

Definition 2.2.8 (Principal ideal)**.** *We say that an ideal I is principal if it is generated by one element, that is there is $f \in I$ such that $\langle f \rangle = I$.*

Definition 2.2.9 (Proper ideal)**.** *An ideal $I \subset R$ is proper if $I \neq R$.*

Definition 2.2.10 (Prime ideal)**.** *An ideal I is prime if and only if for all $f, g \in R$, $f\,g \in I$ implies $f \in I$ or $g \in I$.*

Definition 2.2.11 (Primary ideal)**.** *An ideal I is primary if and only if for all $f, g \in R$, $f\,g \in I$ implies $f \in I$ or there is $k \in \mathbb{N}$ such that $g^k \in I$.*

Definition 2.2.12 (Radical ideal)**.** *An ideal I is radical if and only if $I = \sqrt{I}$.*

These special kinds of ideals are related.

Proposition 2.2.13. *[CLO15, Lem. 4.8.2] Every prime ideal is radical and the radical of any primary ideal is prime.*

Proposition 2.2.14. *[CLO15, Prop. 4.3.16] Consider two ideals $I, J \subset R$. Then*

$$\sqrt{I} \cap \sqrt{J} = \sqrt{I \cap J}$$

Proposition 2.2.15. *[CLO15, Ex. 4.3.12] Consider two ideals $I, J \subset \mathbb{K}[\boldsymbol{x}]$ such that $I \subseteq \sqrt{J}$. Then, as they are finitely generated, there is $k \in \mathbb{N}$ such that $I^k \subseteq J$.*

Theorem 2.2.16 (Primary decomposition)**.** *[CLO15, Thm. 4.8.7] Consider an ideal $I \subset \mathbb{K}[\boldsymbol{x}]$. Then, there are primary ideals $Q_1, \ldots, Q_r \in R$ such that,*

$$I = Q_1 \cap \cdots \cap Q_r,$$

where $(\forall i)\, Q_i \not\supset \bigcap_{j \neq i} Q_j$ and $(\forall i \neq j)\, \sqrt{Q_i} \neq \sqrt{Q_j}$, for $i, j \in [r]$.

We call the set $(Q_1, \ldots, Q_r)$ a minimal primary decomposition *of I. The minimal primary decompositions of I are not unique, but they involve always the same number of primary ideals and the prime ideals $\sqrt{Q_1}, \ldots, \sqrt{Q_r}$, are the same. These prime ideals are called the* associated primes *of $\mathbb{K}[\boldsymbol{x}]/I$.*

Corollary 2.2.17. *[CLO15, Cor. 4.8.10] Let $I = \bigcap_{i=1}^{r} Q_i$ be a* minimal primary decomposition *of a proper radical ideal $I \subset \mathbb{K}[\boldsymbol{x}]$. Then, each Q_i is prime, and so the decomposition is unique.*

Given a ring, we can associate it a dimension.

Definition 2.2.18 (Krull dimension)**.** *[Har77, Page 6] In a commutative ring R, the* height *of a prime ideal $\mathfrak{p} \in R$ is the supremum of all integers n such that there exists a chain $\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_n = \mathfrak{p}$ of distinct prime ideals.*

The Krull dimension of R, $\dim_{Krull}(R)$, is the supremum of the heights of all prime ideals.

The Krull dimension carries a lot of information about the ideal. For example, in the zero-dimensional case we have the following proposition:

Proposition 2.2.19. *[AM69, Thm. 8.5 & Ex. 8.3] Consider an ideal $I \subset \mathbb{K}[\boldsymbol{x}]$ such that the Krull dimension of the quotient ring $\mathbb{K}[\boldsymbol{x}]/I$ is zero. Then, $\mathbb{K}[\boldsymbol{x}]/I$ is an Artinian ring (satisfies the descending chain condition for ideals) and a finite dimensional $\mathbb{K}$-vector space.*

We can relate the number of polynomials defining an ideal the Krull dimension of its associated quotient ring.

Definition 2.2.20 (Zero divisor)**.** *Consider a Noetherian ring R. We say that a non-zero $f \in R$ is a* zero divisor *if there is a non-zero $g \in R$ such that $g\, f = 0$. We say that a f is* regular *if it is not a zero divisor.*

Proposition 2.2.21 (Krull's principal ideal theorem)**.** *[AM69, Cor. 11.17] Consider a Noetherian ring R and let $f \in R$ such that f is neither a zero divisor nor a unit. Then, every minimal prime ideal $\mathfrak{p}$ containing f has height* 1.

Corollary 2.2.22. *Consider a proper principal ideal $I \subset R$, that is, an ideal $I = \langle f \rangle$ generated by $f \in R$, such that $\langle f \rangle \neq R$. Then,*

$$\dim_{Krull}(R) - 1 \leq \dim_{Krull}(R/I) \leq \dim_{Krull}(R).$$

Moreover, if f is a regular element in R, that is $(\forall g \in R \setminus \{0\})$ it holds $f\, g \neq 0$, then $\dim_{Krull}(R) - 1 = \dim_{Krull}(R/I)$.

Corollary 2.2.23. *Let $\langle f_1, \ldots, f_r \rangle$ be a proper ideal of $\mathbb{K}[\boldsymbol{x}]$. Then, $\dim_{Krull}(R/\langle f_1, \ldots, f_r \rangle) \geq n - r$.*

The converse of Cor. 2.2.22 does not necessarily hold. That is, for some rings R, there are zero divisors $f \in R$ such that $\langle f \rangle \subset R$ is a proper ideal and $\dim_{Krull}(R/\langle f \rangle) = \dim_{Krull}(R) - 1$.

Example 2.2.24. *Consider $R = \mathbb{K}[x,y]/\langle x^2, x\,y\rangle$. The Krull dimension of R is 1, $\dim_{Krull}(R) = 1$. The element y is a zero divisor, $x \notin R$ and $x\,y \in R$, however $dim(\mathbb{K}[x,y]/\langle x^2, x\,y, y\rangle) = 0$.*

Both, ideals and quotient rings, are special cases of R-modules. An R-module is a generalization of a $\mathbb{K}$-vector space over a ring.

Definition 2.2.25 (R-module)**.** *Let R be a ring with multiplicative identity 1_R. A (left) R-module M is an abelian group $(M,+)$ together with an operation $\cdot : R \times M \to M$ such that for every $x, y \in R$ and $f, g \in M$ it holds:*

- $x \cdot (f+g) = x \cdot f + x \cdot g$
- $(x+y) \cdot f = x \cdot f + y \cdot f$
- $(x\,y) \cdot f = x \cdot (y \cdot f)$
- $1_R \cdot f = f$

To simplify the notation, we write $x f$ instead of $x \cdot f$, for $x \in R$ and $f \in M$.

Definition 2.2.26 (Free module)**.** *A free R-module M is a module isomorphic to R^k, for a $k \in \mathbb{N}$. The rank of this free module M is k.*

A particular kind of R-modules are the R-algebras. An R-algebra is a ring A together with a ring homomorphism $\phi : R \to A$.

Definition 2.2.27 (R-Algebra)**.** *[AM69, Page 29] Let R and A be two commutative rings and consider a ring homomorphism $\phi : R \to A$. We define the multiplication of $r \in R$ and $f \in A$ in A as $\phi(r)\,f$. With this definition of multiplication, the R-algebra A has an R-module structure. In this case, we say that A is a commutative R-*algebra*.*

Example 2.2.28. *The polynomial ring $\mathbb{K}[\boldsymbol{x}]$ is a $\mathbb{K}$-algebra.*

Given a module M, we can construct algebras related to it which we can use to deduce some properties of M. To define these algebras, we need to introduce the tensor product of modules, which is related to the bilinear maps over these modules.

Definition 2.2.29 (Tensor product)**.** *Consider two R-modules M and N. The tensor product $M \otimes_R N$ is an R-module generated by $\{m \otimes n : m \in M, n \in N\}$ such that for all $m_1, m_2 \in M$, $n_1, n_2 \in N$ and $r \in R$, it holds*

- $(m_1 + m_2) \otimes_R n_1 = (m_1 \otimes_R n_1) + (m_2 \otimes_R n_1)$,
- $m_1 \otimes_R (n_1 + n_2) = (m_1 \otimes_R n_1) + (m_1 \otimes_R n_2)$, *and*
- $(r\,m_1 \otimes_R n_1) = (m_1 \otimes_R r\,n_1) = r\,(m_1 \otimes_R n_1)$.

When it is clear from the context, we skip the subindex from $\otimes_R$.

Proposition 2.2.30. *[AM69, Prop. 2.14] The tensor product is associative, meaning that for every three R-modules M_1, M_2, M_3, it holds $(M_1 \otimes M_2) \otimes M_3 \cong M_1 \otimes (M_2 \otimes M_3)$.*

Example 2.2.31. *The algebra $\mathbb{K}[\boldsymbol{x}]$ is isomorphic to $\mathbb{K}[x_1] \otimes_{\mathbb{K}} \mathbb{K}[x_2] \otimes_{\mathbb{K}} \cdots \otimes_{\mathbb{K}} \mathbb{K}[x_n]$.*

Remark 2.2.32. *Given a R-module M, $M \otimes_R R \cong M$.*

Definition 2.2.33 (Tensor, symmetric and exterior algebras)**.** *Given a R-module M, let $M^{\otimes 0} = R$ and, for $i > 0$, let $M^{\otimes i} = M \otimes_R \cdots \otimes_R M$ be i times the tensor product of M with itself. We define the tensor algebra $T(M)$ as*

$$T(M) := \bigoplus_{i \geq 0} M^{\otimes i} = (R) \oplus (M) \oplus (M \otimes M) \oplus (M \otimes M \otimes M) \oplus \dots$$

The product of $m_1 \otimes \cdots \otimes m_r \in M^{\otimes r}$ and $n_1 \otimes \cdots \otimes n_s \in M^{\otimes s}$ is $m_1 \otimes \cdots \otimes m_r \otimes n_1 \otimes \cdots \otimes n_s \in M^{\otimes r+s}$.

The symmetric algebra of M is the following quotient algebra:

$$S(M) := T(M)/\{n \otimes m - m \otimes n : m, n \in M\}.$$

We call symmetric product the product associated to this algebra and we represent it as $m * n$. This product is symmetric, that is $m * n = n * m$ in $S(M)$.

The exterior algebra of M is

$$\bigwedge M := T(M)/\{m \otimes m : m \in M\}.$$

We call exterior (or wedge) product the product associated to this algebra and we represent it as $m \wedge n$. This product is antisymmetric, that is $m \wedge n = -n \wedge m$ in $\bigwedge M$.

Example 2.2.34. *The polynomial ring $\mathbb{K}[x]$ is isomorphic to the symmetric algebra $S(\mathbb{K})$, where we consider $\mathbb{K}$ as an algebra over itself.*

Proposition 2.2.35. *[Eis04, Prop. A2.2] Let M and N be two R-modules. The direct sum commutes with the tensor product over symmetric and exterior algebras, that is $S(M \oplus N) = S(M) \otimes S(N)$ and $\bigwedge(M \oplus N) = \bigwedge(M) \otimes \bigwedge(N)$.*

Example 2.2.36. *In Ex. 2.2.31 we show that $\mathbb{K}[x_1, \dots, x_n] = \bigotimes_{k=1}^{q} \mathbb{K}[x_k]$, and in Ex. 2.2.34 we show that $\mathbb{K}[x_k] \cong S(\mathbb{K})$. We can combine both arguments to deduce that the polynomial ring $\mathbb{K}[x_1, \dots, x_n]$ is the symmetric algebra $S(\mathbb{K}^n)$, thus*

$$\mathbb{K}[x_1, \dots, x_n] = \bigotimes_{k=1}^{q} \mathbb{K}[x_k] \cong \bigotimes_{k=1}^{n} S(\mathbb{K}) = S(\mathbb{K} \oplus \cdots \oplus \mathbb{K}) = S(\mathbb{K}^n).$$

Definition 2.2.37 (Tensor product of linear maps)**.** *Let $F : M \to M'$ and $G : N \to N'$ be two linear maps. We define the tensor products of the linear maps as follows*

$$\begin{array}{rccl} F \otimes G : & M \otimes N & \to & M' \otimes N' \\ & (m \otimes n) & \mapsto & (F \otimes G)(m \otimes n) = F(m) \otimes G(n) \end{array}$$

Example 2.2.38. *Consider $F : \mathbb{K}[x] \to \mathbb{K}$ an evaluation morphism that specializes x to 1. For example, for $x^2 + 2 \in \mathbb{K}[x]$, $F(x^2 + 2) = 3$. Let $G : \mathbb{K}[y] \to \mathbb{K}[y]$ be a morphism that performs a linear change of coordinates, that is, it replaces y by $2\,y + 1$. For example, for $y^2 - 1 \in \mathbb{K}[y]$, $G(y^2 - 1) = 4\,y^2 + 4\,y$.*

We consider the map $F \otimes G : \mathbb{K}[x] \otimes \mathbb{K}[y] \to \mathbb{K} \otimes \mathbb{K}[y]$. Via the isomorphisms $\mathbb{K}[x] \otimes \mathbb{K}[y] \cong \mathbb{K}[x, y]$ and $\mathbb{K} \otimes \mathbb{K}[y] \cong \mathbb{K}[y]$, we can rewrite this map as $F \otimes G : \mathbb{K}[x, y] \to \mathbb{K}[y]$. For example, $F \otimes G$ maps $x^2\,y + 2\,x\,y^2 \in \mathbb{K}[x, y]$ to

$$(F \otimes G)(x^2\,y + 2\,x\,y^2) = (F \otimes G)(x^2\,y) + 2 \cdot (F \otimes G)(x\,y^2) =$$
$$F(x^2)\,G(y) + 2 \cdot F(x)\,G(y^2) = 1\,(2\,y + 1) + 2 \cdot 1\,(2\,y + 1)^2 = 8\,y^2 + 10\,y + 3.$$

2.3 Graded rings and modules

The results from this section come mainly from [CLO06, Ch. 6] and [Eis04, Ch. 1 & Ap. 2]

In some cases, we can write the rings or the modules as a direct sum of easier algebraic structures, for example $\mathbb{K}$-vector spaces. To index these structures, we use monoids, that is semigroups with identity. In this thesis, when we refer to a monoid related to a grading, we are considering a monoid isomorphic to the monoid given by $\mathbb{Z}^k$ and the operation of coordinate-wise addition, for some $k > 0$.

Definition 2.3.1 (Graded rings and modules)**.** *Let L be a monoid isomorphic to $\mathbb{Z}^k$, for $k \in \mathbb{N}$. A L-graded ring R is a ring that we can write as a direct sum of abelian groups (in our setting, $\mathbb{K}$-vector spaces) $R = \bigoplus_{m \in L} R_m$, such that for all $m_1, m_2 \in L$, $R_{m_1}\,R_{m_2} \subset R_{m_1+m_2}$. A L-graded R-module M is a module M that we can decompose as $\bigoplus_{m \in L} M_m$, where for every $m_1, m_2 \in L$, $R_{m_1} M_{m_2} \subset M_{m_1+m_2}$.*

Example 2.3.2 (Standard $\mathbb{Z}$ grading of $\mathbb{K}[\boldsymbol{x}]$)**.** *Consider the monoid $\mathbb{Z}$. Then, $\mathbb{K}[\boldsymbol{x}] = \bigoplus_{d \in \mathbb{Z}} \mathbb{K}[\boldsymbol{x}]_d$, where $\mathbb{K}[\boldsymbol{x}]_d$ is the $\mathbb{K}$-vector space given by the monomials in $\mathbb{K}[\boldsymbol{x}]$ of degree d, together with 0. As there are no monomials of negative degree, $\mathbb{K}[\boldsymbol{x}]_d = 0$, for $d < 0$.*

Example 2.3.3 (Grading by degree)**.** *Consider the monoid $\mathbb{Z}^2$. Then, $\mathbb{K}[x, y] = \bigoplus_{(d_x, d_y) \in \mathbb{Z}^2} \mathbb{K}[x]_{d_x} \otimes \mathbb{K}[y]_{d_y}$.*

Example 2.3.4 (Grading of $S(M)$, $T(M)$ and $\bigwedge M$)**.** *[Eis04, Cor A2.3] The tensor, symmetric and exterior algebras of M have a standard grading over $\mathbb{N}$ given by the direct sum of the abelian groups generated by the product of $i \in \mathbb{N}$ elements, that is*

$$T(M) \cong \bigoplus_{i \in \mathbb{N}} M^{\otimes i}, \quad S(M) \cong \bigoplus_{i \in \mathbb{N}} S_i(M), \text{ and } \bigwedge M \cong \bigoplus_{i \in \mathbb{N}} \overset{i}{\bigwedge} M,$$

where $S_0(M) = \overset{0}{\bigwedge} M = R$ and, for $i > 1$,

- *$S_i(M)$ is the abelian group generated by $\{a_{j_1} * \cdots * a_{j_i} : a_{j_1}, \ldots, a_{j_i} \in M\}$, the symmetric product of i elements, and*
- *$\overset{i}{\bigwedge} M$ is the abelian group generated by $\{a_{j_1} \wedge \cdots \wedge a_{j_i} : a_{j_1}, \ldots, a_{j_i} \in M\}$, the wedge product of i elements.*

In particular, if M is a free R-module of rank k, with basis $e_1, \dots, e_k$, then

- *The set $\{e_{j_1} \otimes \cdots \otimes e_{j_i} : j_1, \dots, j_i \in \{1, \dots, k\}\}$ is a basis of $M^{\otimes i}$, which is a free R-module of rank k^i.*
- *The set $\{\overbrace{e_1 * \cdots * e_1}^{j_1 \text{ times}} * \cdots * \overbrace{e_k * \cdots * e_k}^{j_k \text{ times}} : j_1 + \cdots + j_k = i\}$ is a basis of $S_i(M)$, which is a free R-module of rank $\binom{k+i-1}{k-1}$.*
- *The set $\{e_{j_1} \wedge \cdots \wedge e_{j_i} : 1 \leq j_1 < \cdots < j_i \leq k\}$ is a basis of $\bigwedge^i M$, which is a free R-module of rank $\binom{k}{i}$.*

Note that, for $i > k$, $\bigwedge^i M \cong 0$.

Definition 2.3.5 (Homogeneous ideal)**.** *Let L be a monoid such that R is L-graded. We say that an ideal $I \subset R$ is homogeneous if, when we consider it as R-module, it is L-graded.*

Proposition 2.3.6. *[CLO15, Thm. 8.3.2] Let L be a monoid such that R is L-graded. An element $f \in R_m$ is homogeneous if there is $m \in L$ such that $f \in R_m$. If R is also a Noetherian ring, an ideal I is homogeneous if and only if it is generated by homogeneous elements.*

Definition 2.3.7 (Grading for homogeneous ideals)**.** *A grading on R, with respect to a monoid L, induces a grading on its homogeneous ideals and on quotient rings of homogeneous ideals. Let $I \subset R$ be a homogeneous ideal, then $I = \bigoplus_{m \in L} I_m$ and $R/I = \bigoplus_{m \in L} R_m/I_m$, where $I_m := (I \cap R_m)$.*

Proposition 2.3.8. *Let $I \in R$ be a homogeneous ideal. Then, we can write every element $f \in R$ as a sum of homogeneous elements in the ideal, that is, we have $f = \sum_{m \in L} f^{(m)}$, where $f^{(m)} \in I_m$.*

As we will see later, when we introduce the twisted graded modules, the same monoid can be associated with different gradings of a module. In this thesis we will use three standard gradings for $\mathbb{K}[\boldsymbol{x}]$. In the first one, we consider the monoid $\mathbb{Z}$ and we graduate the ring with respect to the total degree of the monomials (see Ex. 2.3.2). In the second one, we consider the monoid $\mathbb{Z}^n$ and we graduate the ring with respect to the degree of each variable (see Ex. 2.3.3). The third one is a mix between the previous two and we will use it when we introduce multihomogeneous ideals in Sec. 2.10. In the following, if we do not make explicit the monoid of the grading, then we are considering the first grading related to the monoid $\mathbb{Z}$.

Example 2.3.9. *The ideal $I = \{x, x\,y\}$ is a homogeneous ideal with respect to the standard grading in $\mathbb{Z}$ and in $\mathbb{Z}^2$. For example, consider the polynomial $f = x^3\,y + x^4 \in I$. Then,*

- *As I is homogeneous with respect to the monoid $\mathbb{Z}$, $f \in I_4$.*
- *As I is homogeneous with respect to the monoid $\mathbb{Z}^2$, we can write f as a sum of two monomials which are $x^3\,y \in \mathbb{K}[\boldsymbol{x}]_{(3,1)}$ and $x^4 \in I_{(4,0)}$.*

The ideal in Ex. 2.3.9 is an example of a monomial ideal.

Definition 2.3.10 (Monomial ideal)**.** *A monomial ideal is a homogeneous ideal $I \subset \mathbb{K}[\boldsymbol{x}]$ with respect to the standard $\mathbb{Z}^n$ grading, that is, with respect to the degree of each variable as in Ex. 2.3.3. These ideals are always generated by monomials.*

The aforementioned gradings are special cases of the standard gradings for tensor products and direct sums.

Definition 2.3.11. *Let M and N be L-graded R-modules, for a monoid L. We define two gradings for $M \oplus N$ and $M \otimes N$ over L and over $L \times L$ as follows,*

	Over L	*Over $L \times L$*
$M \oplus N =$	$\displaystyle\bigoplus_{m \in L} M_m \oplus N_m$	$\displaystyle\bigoplus_{(m_1,m_2) \in L \times L} M_{m_1} \oplus N_{m_2}$
$M \otimes N =$	$\displaystyle\bigoplus_{m \in L} \left(\bigoplus_{\substack{m_1,m_2 \in L \\ m_1+m_2=m}} M_{m_1} \otimes N_{m_2} \right)$	$\displaystyle\bigoplus_{(m_1,m_2) \in L \times L} M_{m_1} \otimes N_{m_2}$

Definition 2.3.12 (Graded homomorphisms)**.** *Given two L-graded modules M, N and a homomorphism $\delta : M \mapsto N$, we say that δ is a graded homomorphism of degree $m \in L$ if, for every $n \in L$, then*

$$\delta(M_n) \subset N_{n+m}$$

Example 2.3.13. *Consider the endomorphism in $\mathbb{K}[\boldsymbol{x}]$ which multiplies a polynomial by x^d, that is, $\delta : \mathbb{K}[\boldsymbol{x}] \to \mathbb{K}[\boldsymbol{x}]$, such that $\delta(f) = x^d f$. This map is a graded homomorphism of degree d as, if the degree of f is D, the degree of $x^d f$ is $d + D$, that is $\delta(\mathbb{K}[\boldsymbol{x}]_D) \subset \mathbb{K}[\boldsymbol{x}]_{D+d}$.*

Given a graded homomorphism, we can modify the gradings of the domain and the codomain of the morphism to change its degree. To do so, we need to extend the monoid to a group. Let R be a graded ring with respect to a monoid L and let L' be the smallest abelian group that contains L. We can extended the grading of R to L' by setting $R_m = 0$ for every $m \in L' \setminus L$. In the following, we extend every grading in this way and, by abusing notation, we use the symbol L to denote the abelian group related to a grading.

Definition 2.3.14 (Twist)**.** *Let R be a graded ring with respect to an abelian group L. For any graded R-module M and $m \in L$, we define the twisted module $M(-m)$ as $\bigoplus_{m' \in L} M_{m'-m}$. Note that $M(-m)$ and M are isomorphic and their only difference is the grading which is "twisted" by $-m$.*

Example 2.3.15. *Consider the ideal $I \in \mathbb{K}[x, y]$ of Ex. 2.3.9, and the twisted ideal $I(-2)$. Then, $I(-2)_1 = I_{-1} = \{\}$, $I(-2)_3 = I_1 = \{x\}$, and $I(-2)_4 = I_2 = \{x^2, x\,y\}$.*

With the appropriate twisting, we can turn every graded homomorphism to a graded homomorphism of degree zero.

Example 2.3.16 (Continuation of Ex. 2.3.13)**.** *The map δ has degree 0 if we twist its domain by $-d$, that is we consider $\delta : \mathbb{K}[\boldsymbol{x}](-d) \to \mathbb{K}[\boldsymbol{x}]$. Then $\delta(\mathbb{K}[\boldsymbol{x}](-d)_D) \subset \mathbb{K}[\boldsymbol{x}](-d)_{D+d} = \mathbb{K}[\boldsymbol{x}]_D$.*

Definition 2.3.17 (Hilbert's function and series)**.** *Consider a graded ring R with respect to a monoid L and a graded R-module M. The Hilbert function, HF_M, is a function that maps $m \in L$ to the dimension of M_m as a $\mathbb{K}$-vector space, that is,*

$$\begin{aligned} HF_M : L &\to \mathbb{Z} \\ m &\mapsto HF_M(m) = \dim_{\mathbb{K}}(M_m). \end{aligned} \tag{2.1}$$

The Hilbert series is the formal series $HS_M(t) = \sum_{m \in L} HF_M(m)\, t^m$.

Proposition 2.3.18. *[CLO06, Prop. 6.4.7] Consider $R = \mathbb{K}[\boldsymbol{x}]$ and the standard $\mathbb{N}$ grading. For each finitely generated R-module M there is a polynomial, which we call the Hilbert polynomial HP_M, such that $HP_M(m) = HF_M(m)$ for big enough $m \in \mathbb{N}$.*

Example 2.3.19. *We consider $\mathbb{K}[\boldsymbol{x}]$ (as a $\mathbb{K}[\boldsymbol{x}]$-module) with respect to the $\mathbb{N}$-grading. Its Hilbert function corresponds to the number of monomials for each degree d; that is, $HF_{\mathbb{K}}[\boldsymbol{x}](d) = \binom{n-1+d}{n-1}$. In this case, the Hilbert function is a polynomial of degree $n-1$. So $HF_{\mathbb{K}[\boldsymbol{x}]}(d) = HP_{\mathbb{K}[\boldsymbol{x}]}(d) = \binom{n-1+d}{n-1}$, for every $d \in \mathbb{N}$. In particular, the Hilbert series is*

$$HS_{\mathbb{K}[\boldsymbol{x}]}(t) = \sum_{d \in \mathbb{N}} \binom{n-1+d}{n-1} t^d = \frac{1}{(1-t)^n}.$$

We can read many geometric and algebraic information of a module from its Hilbert series; for example the Krull dimension.

Proposition 2.3.20. *[BH98, Cor. 4.1.8] Consider a homogeneous ideal $I \subset \mathbb{K}[\boldsymbol{x}]$ and let d be the the Krull dimension of $\mathbb{K}[\boldsymbol{x}]/I$. Then, d is equal to the degree of $HP_{\mathbb{K}[\boldsymbol{x}]/I}(t)$ and there is a polynomial $H(t) \in \mathbb{Q}[t]$ such that $H(1) \neq 0$ and $HS_{\mathbb{K}[\boldsymbol{x}]/I}(t) = \frac{H(t)}{(1-t)^d}$.*

2.4 Affine varieties

The results of this section come mainly from [CLO15, Ch. 1, Ch. 4 & Ch. 5], [Har77, Sec. I.1], and [Per07, Ch. IV]

There is a strong relation between algebra and geometry. For example, when we consider univariate polynomials over $\mathbb{C}$, the fundamental theorem of algebra tell us that there is a one-to-one correspondence between a finite set of points in $\mathbb{C}$ and square-free monic polynomials in $\mathbb{C}[x]$. The objective of this section is to formalize this relation and to introduce a multivariate generalization, that is *Hilbert's Nullstellensatz*.

Definition 2.4.1 (Affine variety)**.** *Given an ideal $I \in \mathbb{K}[\boldsymbol{x}]$, the affine variety $\mathbb{V}_{\mathbb{K}^n}(I)$ is the set of all the points over $\mathbb{K}^n$ such that every polynomial in I vanishes at these points, that is,*

$$\{\boldsymbol{p} \in \mathbb{K}^n : (\forall f \in I)\ f(\boldsymbol{p}) = 0\}.$$

We take the previous definition from [CLO15] and [AM69]. Some other authors, as [Eis04] and [Har77], call these objects algebraic sets and reserve the word affine variety for the irreducible varieties.

In what follows, we write $\mathbb{V}_{\mathbb{K}^n}(f_1, \ldots, f_k)$ to refer to the variety $\mathbb{V}_{\mathbb{K}^n}(\langle f_1, \ldots, f_k\rangle)$.

Definition 2.4.2 (Irreducible variety)**.** *We say that an affine variety is irreducible if and only if we cannot write it as a union of two (strict) subvarieties. That is, V is irreducible if for every two affine varieties $V_1, V_2 \subset V$ such that $V = V_1 \cup V_2$, it holds $V = V_1$ or $V = V_2$.*

Example 2.4.3 (Reducible variety)**.** *The variety $\mathbb{V}_{\mathbb{K}^n}(x\,y) \subset \mathbb{K}^2$ is not irreducible. This variety equals to the points that have at least one coordinate equal to 0. So we can write the variety as,*

$$\mathbb{V}_{\mathbb{K}^n}(x\,y) = \{(a,b) \in \mathbb{K}^2 : a = 0 \text{ or } b = 0\} =$$
$$\{(a,b) \in \mathbb{K}^2 : a = 0\} \cup \{(a,b) \in \mathbb{K}^2 : y = 0\} = \mathbb{V}_{\mathbb{K}^n}(x) \cup \mathbb{V}_{\mathbb{K}^n}(y).$$

Notice that $\mathbb{V}_{\mathbb{K}^n}(x) \neq \mathbb{V}_{\mathbb{K}^n}(x\,y)$ and $\mathbb{V}_{\mathbb{K}^n}(y) \neq \mathbb{V}_{\mathbb{K}^n}(x\,y)$ because $\{(1,0),(0,1)\} \subset \mathbb{V}_{\mathbb{K}^n}(x\,y)$ but $(1,0) \notin \mathbb{V}_{\mathbb{K}^n}(x)$, and $(0,1) \notin \mathbb{V}_{\mathbb{K}^n}(y)$.

Given a set of points W in $\mathbb{K}^n$, we can consider the set of all the polynomials that vanish at W.

Proposition 2.4.4 (Ideal of a set of points)**.** *[CLO15, Thm. 4.2.7] Consider a set $W \subset \mathbb{K}^n$. The set of polynomials that vanish at W define a radical ideal $\mathbb{I}(W)$, where*

$$\mathbb{I}(W) := \{f \in \mathbb{K}[\boldsymbol{x}] : (\forall \boldsymbol{p} \in W)\ f(\boldsymbol{p}) = 0\}.$$

Example 2.4.5. *Given a point $(a,b) \in \mathbb{K}^2$, its defining ideal is*

$$\mathbb{I}(\{(a,b)\}) = \langle x - a, y - b\rangle \subset \mathbb{K}[x,y].$$

Proposition 2.4.6. *[CLO15, Prop. 4.5.3] An affine variety V is irreducible if and only if $\mathbb{I}(V)$ is prime.*

As every ideal is finitely generated (Prop. 2.2.3) we can think the varieties as the solution set of a system of polynomial equations.

Definition 2.4.7 (Polynomial system)**.** *A polynomial system, or a system of polynomial equations, is a set of equations defined by polynomials $f_1, \ldots, f_r \in \mathbb{K}[\boldsymbol{x}]$ as*

$$\begin{cases} f_1(x_1, \ldots, x_n) & = 0 \\ \quad\vdots & \\ f_r(x_1, \ldots, x_n) & = 0 \end{cases}. \tag{2.2}$$

The solutions of the system form an affine variety over $\mathbb{K}^n$ given by $\mathbb{V}_{\mathbb{K}^n}(f_1, \ldots, f_r)$. In the following, we write $\{f_1, \ldots, f_r\} \in \mathbb{K}[\boldsymbol{x}]$ to refer to this polynomial system.

Despite of the classical topology that $\mathbb{K}^n$ might have, we can define another topology for $\mathbb{K}^n$ related to algebraic varieties.

Definition 2.4.8 (Zariski topology)**.** *The intersection and finite union of affine varieties is again a variety. Moreover, the empty set and $\mathbb{K}^n$ are affine varieties given by $\mathbb{V}_{\mathbb{K}^n}(\langle 1\rangle)$ and $\mathbb{V}_{\mathbb{K}^n}(\langle 0\rangle)$, respectively. The Zariski topology over $\mathbb{K}^n$ is the topology where the closed sets are the affine varieties and the open sets are their complements.*

In the Zariski topology some sets over $\mathbb{K}^n$ are neither open nor close. For this reason, we define the Zariski closure of a set, which is the smallest closed set in this topology that contains it.

Definition 2.4.9 (Zariski closure)**.** *Given a set $W \subset \mathbb{K}^n$, we denote by $\overline{W}^Z$ the smallest affine variety that contains W, that is*

$$\overline{W}^Z := \mathbb{V}_{\mathbb{K}^n}(\mathbb{I}(W)).$$

There is a one-to-one correspondence between radical ideals and affine varieties.

Proposition 2.4.10 (Strong Hilbert's Nullstellensatz)**.** *[CLO15, Thm. 4.2.6] Given an ideal $I \subset \mathbb{K}[\boldsymbol{x}]$, then*

$$\mathbb{I}(\mathbb{V}(I)) = \sqrt{I}.$$

In other words, for every I, if $f \in \mathbb{K}[\boldsymbol{x}]$ vanishes over the variety $\mathbb{V}_{\mathbb{K}^n}(I)$, there is a $k \in \mathbb{N}$, such that $f^k \in I$.

Proposition 2.4.11 (Ideal-Variety correspondence)**.** *[CLO15, Thm. 4.2.7] For any two ideals $I, J \subset \mathbb{K}[\boldsymbol{x}]$ such that $I \subset J$, it holds $\mathbb{V}_{\mathbb{K}^n}(I) \supset \mathbb{V}_{\mathbb{K}^n}(J)$. For any two affine varieties $V, W \subset \mathbb{K}^n$ such that $V \subset W$, it holds $\mathbb{I}(V) \supset \mathbb{I}(W)$.*

Corollary 2.4.12. *For any two ideals $I, J \subset \mathbb{K}[\boldsymbol{x}]$ such that $\sqrt{I} = \sqrt{J}$, it holds $\mathbb{V}_{\mathbb{K}^n}(I) = \mathbb{V}_{\mathbb{K}^n}(J)$.*

Corollary 2.4.13 (Weak Hilbert's Nullstellensatz)**.** *[CLO15, Thm. 4.1.1] Let I be an ideal such that $\mathbb{V}_{\mathbb{K}^n}(I) = \emptyset$, then $1 \in I$.*

The next table, extracted from [CLO15, Ch. 4.9], summarizes the relations between affine varieties and ideals.

Algebra		Geometry
Radical ideals		Varieties
I	$\rightarrow$	$\mathbb{V}_{\mathbb{K}^n}(I)$
$\mathbb{I}(V)$	$\leftarrow$	V
Inclusion of ideals		Inclusion of Varieties
$I \subset J$	$\rightarrow$	$\mathbb{V}_{\mathbb{K}^n}(I) \supset \mathbb{V}_{\mathbb{K}^n}(J)$
$\mathbb{I}(V) \subset \mathbb{I}(W)$	$\leftarrow$	$V \supset W$
Addition of ideals		Intersection of varieties
$I + J$	$\rightarrow$	$\mathbb{V}_{\mathbb{K}^n}(I) \cap \mathbb{V}_{\mathbb{K}^n}(J)$
$\sqrt{\mathbb{I}(V) + \mathbb{I}(W)}$	$\leftarrow$	$V \cap W$
Product of ideals		Union of varieties
$I\,J$	$\rightarrow$	$\mathbb{V}_{\mathbb{K}^n}(I) \cup \mathbb{V}_{\mathbb{K}^n}(J)$
$\mathbb{I}(V) \cap \mathbb{I}(W)$	$\leftarrow$	$V \cup W$
Ideal quotient		Difference of varieties
$I : J$	$\rightarrow$	$\overline{\mathbb{V}_{\mathbb{K}^n}(I) \setminus \mathbb{V}_{\mathbb{K}^n}(J)}^Z$
$\mathbb{I}(V) : \mathbb{I}(W)$	$\leftarrow$	$\overline{V \setminus W}^Z$
Prime ideal	$\leftrightarrow$	Irreducible varieties

As varieties correspond to radical ideals and radical ideals can be written uniquely as the intersection of prime ideals (Cor. 2.2.17), then every variety can be written uniquely as the intersection of irreducible varieties.

Proposition 2.4.14 (Minimal irreducible decomposition)**.** *[CLO15, Thm. 4.6.4] Given an affine variety $V \subset \mathbb{K}^n$, there are unique irreducible varieties $V_1, \ldots, V_r$ such that $(\forall i \neq j)\, V_i \not\subset V_j$ and*

$$V = V_1 \cup \cdots \cup V_r.$$

Varieties are geometric objects over the topological space $\mathbb{K}^n$. Hence, we can associate to them a dimension, see [Har77, Page 5].

Definition 2.4.15 (Dimension)**.** *The dimension of an affine variety $V \subseteq \mathbb{K}^n$, $\dim(V)$, is the supremum of all the integers $d \geq 0$ such that there is a chain $W_0 \subset W_1 \subset \cdots \subset W_d \subseteq V$ of different irreducible affine varieties contained in V.*

It follows directly from the definition that we can deduce the dimension of a variety from the dimension of its affine irreducible pieces.

Proposition 2.4.16. *[Per07, Prop. IV.1.4] Let $V \subset \mathbb{K}^n$ be an affine variety and consider its minimal irreducible decomposition $V = \bigcup_{i \leq r} V_i$, see Prop. 2.4.14. The dimension of V equals the maximal of the dimensions of the irreducible components, that is,*

$$\dim(V) = \max_i(\dim(V_i)).$$

Up to now, given an ideal $I \subset \mathbb{K}[\boldsymbol{x}]$, we associated to it a quotient ring $\mathbb{K}[\boldsymbol{x}]/I$ (Def. 2.2.6) and an affine variety $\mathbb{V}_{\mathbb{K}^n}$. For both objects, we defined a notion of dimension (Def. 2.2.18). In what follows, we show that, for affine varieties, these concepts are the same.

By Hilbert's Nullstellensatz, different ideals correspond to the same affine variety. To make this correspondence one-to-one, to each variety we will associate a radical ideal (Prop. 2.4.4) and consider its quotient ring. This ring, that we call *coordinate ring*, is isomorphic to the ring of polynomial mappings (regular functions) from the variety to $\mathbb{K}$ [Har77, Thm. I.3.2]. For more information about this relation see [CLO15, Ch. 5.1].

Definition 2.4.17 (Coordinate Ring)**.** *Given a variety $V \in \mathbb{K}^n$, its coordinate ring is*

$$\mathbb{K}[V] = \mathbb{K}[\boldsymbol{x}]/\mathbb{I}(V).$$

Proposition 2.4.18 (Dimension of a variety and a coordinate ring)**.** *[Har77, Prop. I.1.7] The dimension of an affine variety V equals the Krull dimension of its coordinate ring, that is,*

$$\dim(V) = \dim_{Krull}(\mathbb{K}[V]).$$

As a consequence of this relation, using Cor. 2.2.23, we can bound the dimension of a variety with respect to the number of generators of an ideal.

Proposition 2.4.19. *[Har77, Prop. I.1.13] We say that an affine variety $V \subset \mathbb{K}^n$ is a* hypersurface*, if there is a polynomial $f \in \mathbb{K}[\boldsymbol{x}]$ such that $V = \mathbb{V}_{\mathbb{K}^n}(\langle f \rangle)$, that is, the variety is defined by a principal ideal. The dimension of an hypersurface is $n-1$. Moreover, if the dimension of $V \subset \mathbb{K}^n$ is $n-1$, then V is an hypersurface.*

To generalize the previous proposition we need to introduce the concept of equidimensional variety.

Definition 2.4.20 (Equidimensional variety)**.** *Let $V \subset \mathbb{K}^n$ be an affine variety and consider $V = \bigcup_{i \leq r} V_i$ its minimal irreducible decomposition. We say that V is* equidimensional *if the dimension of each irreducible component is the same, that is,*

$$(\forall i)\ \dim(V_i) = \dim(V).$$

Example 2.4.21. *The variety $\mathbb{K}^n$ is equidimensional.*

Proposition 2.4.22. *[Har77, Prop. I.7.1] Consider two equidimensional varieties $V, W \subset \mathbb{K}^n$ such that $V \cap W \neq 0$. Consider the irreducible decomposition of $V \cap W = U_1 \cup \cdots \cup U_r$ (Prop. 2.4.14). Then, the dimension of each irreducible component U_i is lower bounded by $\dim(V) + \dim(W) - n$, that is,*

$$(\forall i)\ \dim(U_i) \geq \dim(V) + \dim(W) - n.$$

Using this proposition, we can bound the dimension of a variety with respect to the number of generators.

Corollary 2.4.23 (Bounds for the dimension)**.** *[Har77, Ex. I.1.9] Consider $I = \langle f_1, \ldots, f_r \rangle \subset \mathbb{K}[\boldsymbol{x}]$ such that $r \leq n$ and I is a proper ideal. Then, the dimension of $\mathbb{V}_{\mathbb{K}^n}(I)$ is at least $n - r$.*

From Prop. 2.2.19, if the variety is zero-dimensional, its coordinate ring is a finite dimensional vector space.

Proposition 2.4.24. *[Cox05, Thm 2.1.2] The Krull dimension of $\mathbb{K}[\boldsymbol{x}]/I$ is zero if and only if $\mathbb{V}_{\mathbb{K}^n}(I)$ is finite. The dimension of $\mathbb{K}[\boldsymbol{x}]/I$ as $\mathbb{K}$-vector space is bigger or equal to the number of points in $\mathbb{V}_{\mathbb{K}^n}(I)$. In particular, these two numbers are the same if and only if I is radical.*

Example 2.4.25. *Consider the ideals $I_{(1)} = \langle x, y \rangle$, $I_{(2)} = \langle x, y^2 \rangle$ and $I_{(3)} = \langle x^2, x\,y, y^2 \rangle$ in $\mathbb{K}[x, y]$. These three ideals describe the variety $\{(0,0)\} \in \mathbb{K}^2$ and so their Krull dimension is zero. However, their quotient rings are not the same, as $\dim_{\mathbb{K}}(\mathbb{K}[x,y]/I_{(1)}) = 1, \dim_{\mathbb{K}}(\mathbb{K}[x,y]/I_{(2)}) = 2$ and $\dim_{\mathbb{K}}(\mathbb{K}[x,y]/I_{(3)}) = 3$. The only radical ideal is $I_{(1)}$, and so its number of solutions equals its dimension as a $\mathbb{K}$-vector space.*

As we will see in Sec. 5.2, one approach to solve zero-dimensional polynomial systems is to consider the ideal I that they define and its quotient ring $\mathbb{K}[\boldsymbol{x}]/I$. Proposition 2.4.24 tell us that, in general, $\mathbb{K}[\boldsymbol{x}]/I$ has more information than the solutions of the system, that is, the variety $\mathbb{V}_{\mathbb{K}^n}(I)$. Only when I is radical, and so $\mathbb{K}[\boldsymbol{x}]/I$ is a coordinate ring, we can expect $\mathbb{K}[\boldsymbol{x}]/I$ to contain no extra information. When the ideal I is not radical, we can still associate to $\mathbb{K}[\boldsymbol{x}]/I$ a geometric object, but this is not a variety anymore, it is an affine scheme. Roughly speaking, a scheme is a variety on which we can associate to each point a multiplicity. As a variety $\mathbb{V}_{\mathbb{K}^n}(I)$ is isomorphic to the set of maximal ideals of $\mathbb{K}[\boldsymbol{x}]/I$, a scheme takes

into account every prime ideal. Schemes are central objects in Algebraic Geometry. However, we will not elaborate on them and we refer the reader to [Har77] for a classical introduction to the subject.

The degree of a univariate polynomial give us its number of roots, counted multiplicities, over the algebraic closure of its coefficients' field. Unfortunately, we cannot do a similar thing when we work with affine varieties given by non-homogeneous ideals.

Example 2.4.26. *Let $a \in \mathbb{Z}$ and consider the ideal $I_{(a)} = \langle x^2 - y^2 + a\,x\,y, x - y - 1\rangle$. According to the value of a the number of elements in $\mathbb{V}_{\mathbb{K}^n}(I_{(a)})$ changes:*

- *For $a = 0$, $\mathbb{V}_{\mathbb{K}^n}(I_{(0)}) = \{(\frac{1}{2}, -\frac{1}{2})\}$ and, as $I_{(0)}$ is radical, $\dim_{\mathbb{K}}(\mathbb{K}[x, y]/I_{(0)}) = 1$*
- *For $a = 2$, $\mathbb{V}_{\mathbb{K}^n}(I_{(2)}) = \{(\frac{2}{\sqrt{2}}, -1+\frac{2}{\sqrt{2}}), (\frac{-2}{\sqrt{2}}, -1-\frac{2}{\sqrt{2}})\}$ and, as $I_{(2)}$ is radical, $\dim_{\mathbb{K}}(\mathbb{K}[x, y]/I_{(2)}) = 2$.*

2.5 Projective varieties

The results of this sections come mainly from [CLO15, Ch. 8], [Eis04, Ch. 1] and [Har77, Ch. 1].

In the following example we explain why in Ex. 2.4.26 we obtained different number of solutions.

Example 2.5.1 (Continuation of Ex. 2.4.26)**.** *We want to study the solutions of the polynomial system $\{x^2 - y^2 + a\,x\,y = 0, x - y - 1 = 0\}$, with respect to the parameter a. We use the second equation to eliminate x from the first equation, that is, we use the equality $x = y + 1$, to write $x^2 - y^2 + a\,x\,y = 0$ as $a\,y^2 + (a + 2)\,y + 1 = 0$. We use the quadratic formula to solve this polynomial and deduce that the solutions of this polynomial are $\beta_1 = \frac{a+2-\sqrt{a^2+4}}{-2\,a}$, and $\beta_2 = \frac{a+2+\sqrt{a^2+4}}{-2\,a}$. Hence, we deduce that the solutions of the original system are $(\alpha_1, \beta_1) = (\frac{a+2-\sqrt{a^2+4}}{-2\,a} + 1, \frac{a+2-\sqrt{a^2+4}}{-2\,a})$, and $(\alpha_2, \beta_2) = (\frac{a+2+\sqrt{a^2+4}}{-2\,a} + 1, \frac{a+2+\sqrt{a^2+4}}{-2\,a})$, We study the limit of the solutions when a goes to zero. The limit of (α_1, β_1) is $(\frac{1}{2}, -\frac{1}{2})$, meanwhile the limit of (α_2, β_2) does not exist. Somehow, (α_2, β_2) is a solution of the system $\{x^2 - y^2 = 0, x - y - 1 = 0\}$ at infinity.*

In this section we introduce projective varieties. These varieties allow us to formalize the notion of "solutions at infinity" that we observed in Ex. 2.5.1. Projective varieties are related to homogeneous ideals. In the following, we will consider the polynomial ring $\mathbb{K}[\boldsymbol{x}][x_0]$ together with its standard gradings over $\mathbb{N}$ (Ex. 2.3.2).

Definition 2.5.2 (Projective space)**.** *[CLO15, Def. 8.2.1] The projective space $\mathbb{P}^n$ represents the space of lines in $\mathbb{K}^{n+1}$ through the origin. Let $\sim$ be an equivalence relation over $\mathbb{K}^{n+1} \setminus \{0\}$ such that $a \sim b$ if there exists $\lambda \in \mathbb{K}$ such that $a = \lambda\, b$. We define the projective n-space $\mathbb{P}^n$ as the quotient of $\mathbb{K}^{n+1}$ (excluding zero) by the equivalence relation $\sim$, that is*

$$\mathbb{P}^n := (\mathbb{K}^{n+1} \setminus \{0\})/\sim .$$

Each $(n + 1)$-tuple $(a_0, \ldots, a_n) \in \mathbb{K}^{n+1}$ defines a point $\boldsymbol{a} \in \mathbb{P}^n$ and we say that $(a_0 : \cdots : a_n)$ are the homogeneous coordinates *of $\boldsymbol{a}$.*

Remark 2.5.3. *Each $(n+1)$-tuples defines homogeneous coordinates of a point in $\mathbb{P}^n$, but many tuples define coordinates of the same point. More precisely, two tuples define the same point if the tuples are equivalent with respect to $\sim$.*

Definition 2.5.4 (Projective variety)**.** *Given a homogeneous ideal $I \subset \mathbb{K}[\boldsymbol{x}][x_0]$, we define the projective variety $\mathbb{V}_{\mathbb{P}^n}(I)$ as the zero set of the homogeneous polynomials in I over $\mathbb{P}^n$, that is*

$$\mathbb{V}_{\mathbb{P}^n}(I) := \{\boldsymbol{p} \in \mathbb{P}^n : (\forall \text{ homogeneous } f \in I)\ f(\boldsymbol{p}) = 0\}.$$

We can define the Zariski topology for projective varieties similarly to the way that we did it for the affine varieties (see Def. 2.4.8).

Definition 2.5.5 (Zariski topology)**.** *[Har77, Page 10] We define the Zariski topology on $\mathbb{P}^n$ by considering the open sets to be the complement of the projective varieties.*

When we work with a projective variety, we might want to study the affine variety related to the homogeneous ideal that defines it. We call this affine variety the affine cone.

Definition 2.5.6 (Affine cone)**.** *Let $V \subset \mathbb{P}^n$ be a projective variety such that $\mathbb{V}_{\mathbb{P}^n}(I) = V$. We define the affine cone of V, C_V, as the affine variety in $\mathbb{K}^{n+1}$ that is the zero set of I. Equivalently, $C_V := \mathbb{V}_{\mathbb{K}^{n+1}}(I)$.*

The empty set $\emptyset \subset \mathbb{P}^n$ is a projective variety as $\emptyset = \mathbb{V}_{\mathbb{P}^n}(\langle 1\rangle)$ and $\emptyset = \mathbb{V}_{\mathbb{P}^n}(\langle x_0, x_1, \ldots, x_n\rangle)$. If we apply Hilbert's Nullstellensatz in the same way as we do in the affine setting, as $\langle 1\rangle$ and $\langle x_0, x_1, \ldots, x_n\rangle$ describe the same variety, we should conclude that the corresponding radical ideals are the same. But this is not the case, as $\sqrt{\langle x_0, x_1, \ldots, x_n\rangle} \neq \sqrt{\langle 1\rangle}$. Hence, we cannot extend directly Hilbert's Nullstellensatz to the projective setting.

Definition 2.5.7 (Irrelevant ideal)**.** *The irrelevant ideal is the biggest ideal that defines an empty variety. For projective varieties over $\mathbb{P}^n$, the irrelevant ideal is $\langle x_0, x_1, \ldots, x_n\rangle = \bigoplus_{i\geq 1} \mathbb{K}[\boldsymbol{x}][x_0]_i$.*

Definition 2.5.8 (Ideal of a projective variety)**.** *Let $V \subset \mathbb{P}^n$ be a projective variety. If $V \neq \emptyset$, then the ideal associated to V is the ideal generated by all the homogeneous polynomials that vanish at every point of V. If $V = \emptyset$, then its associated ideal is the irrelevant ideal $\langle x_0, \ldots, x_n\rangle$.*

When the projective variety is not empty, the associated ideal equals the associated ideal of the affine cone. Hence, by abusing notation, we denote by $\mathbb{I}(V)$ the associated ideal of a projective variety V.

Now we can state the, so-called, projective Hilbert's Nullstellensatz. It provides an one-to-one correspondence between projective varieties and radical ideals contained in the irrelevant ideal.

Proposition 2.5.9 (Projective Hilbert's Nullstellensatz)**.** *[CLO15, Thm. 8.3.9] Let $I \subset \langle x_0, \ldots, x_n\rangle \subset \mathbb{K}[\boldsymbol{x}][x_0]$ be an homogeneous ideal. Then,*

$$\mathbb{I}(\mathbb{V}_{\mathbb{P}^n}(I)) = \sqrt{I}.$$

If $\mathbb{V}_{\mathbb{P}^n}(I) = \emptyset$, then there is a $k \in \mathbb{N}$ such that, for every $0 \leq i \leq n$, $x_i^k \in I$.

One might ask what happened with the $1 \in \mathbb{K}[\boldsymbol{x}][x_0]$ that appears in any ideal defining an empty affine variety. To answer this question we have to recall the concept of saturation of an ideal, see prop. 2.2.5. Let $V \subset \mathbb{K}^n$ be a variety and $I \subset \mathbb{K}[\boldsymbol{x}]$ an ideal. The ideal I has no solutions outside V, that is $\mathbb{V}(I) \cap (\mathbb{K}^n \setminus V) = \emptyset$, if and only if $\mathbb{V}(I) \subset V$. By Prop. 2.4.11, this is equivalent to the condition $\mathbb{I}(V) \subset \sqrt{I}$. By Prop. 2.2.15, there is a $k \in \mathbb{N}$ such that $\mathbb{I}(V)^k \subset (\sqrt{I})^k \subset I$, and so $(I : \mathbb{I}(V)^k) = \langle 1 \rangle$. Therefore, the lack of solutions outside V is equivalent to $(I : \mathbb{I}(V)^\infty) \supseteq (I : \mathbb{I}(V)^k) = \langle 1 \rangle$.

Proposition 2.5.10. *An ideal $I \in \mathbb{K}[\boldsymbol{x}]$ has no solutions outside $V \subset \mathbb{K}^n$, that is $\mathbb{V}_{\mathbb{K}^n}(I) \cap (\mathbb{K}^n \setminus V) = \emptyset$, if and only if $(I : \mathbb{I}(V)^\infty) = \langle 1 \rangle$.*

Hence, a proper homogeneous ideal over $\mathbb{K}[\boldsymbol{x}][x_0]$ defines an empty projective variety if its affine cone contains only the point $\mathbf{0} \in \mathbb{V}_{\mathbb{K}^{n+1}}(\langle x_0, \ldots, x_n \rangle)$.

Projective varieties are obtained from "gluing" together affine varieties. These affine varieties are open subsets of the projective varieties.

Proposition 2.5.11 ([CLO15, Ex. 8.2.9])**.** *For each $0 \leq i \leq n$, we define $U_i \subset \mathbb{P}^n$ as the open set of $\mathbb{P}^n$ formed by the points whose i-th coordinate is different to zero. Equivalently,*

$$U_i = \mathbb{P}^n \setminus \mathbb{V}_{\mathbb{P}^n}(\langle x_i \rangle).$$

For each projective variety $V \subseteq \mathbb{P}^n$, $V \cap U_i$ is an affine variety. We say that the affine varieties $\{V \cap U_i\}_i$ form an affine open cover *of V.*

We can think of this affine open covering as follows: Given a homogeneous ideal $I \in \mathbb{K}[\boldsymbol{x}][x_0]$, consider the affine variety $\mathbb{V}_{\mathbb{P}^n}(I) \cap U_0$ given by the intersection of the affine cone $\mathbb{V}_{\mathbb{K}^{n+1}}(I)$ with the hyperplane $x_0 = 1$. Hence, we have correspondence

$$(a_0 : \cdots : a_n) \in \mathbb{V}_{\mathbb{P}^n}(I) \cap U_0 \leftrightarrow \left(1, \frac{a_1}{a_0}, \ldots, \frac{a_n}{a_0}\right) \in \mathbb{V}_{\mathbb{K}^{n+1}}(I + \langle x_0 - 1 \rangle).$$

There is a similar correspondence for each variable x_i.

We can also use the affine open coverings to define varieties. The way of doing so corresponds to "gluing" the affine varieties, that is, identify the affine pieces $(\mathbb{V}_{\mathbb{P}^n}(I) \cap U_i) \cap U_j$ and $(\mathbb{V}_{\mathbb{P}^n}(I) \cap U_j) \cap U_i$. For more information see [CLS11, Sec. 3.0].

Example 2.5.12 (Continuation of Ex. 2.5.1)**.** *We consider the homogeneous ideal $I := \langle x^2 - y^2, x - y - z \rangle \subset \mathbb{K}[x, y, z]$. This ideal defines a projective variety $\mathbb{V}_{\mathbb{P}^2}(I)$ which consists of two points. Let U_x, U_y, U_z be the open sets from Prop. 2.5.11.*

The affine variety defined by $\mathbb{V}_{\mathbb{P}^2}(I) \cap U_z$ is isomorphic to the affine variety associated to the ideal $I_z := \{x^2 - y^2, x - y - 1\} \subset \mathbb{K}[x, y]$ by setting $z = 1$. Note that this ideal agrees with the one obtained in Ex. 2.5.1 by setting $a = 0$. Hence, $\mathbb{V}_{\mathbb{K}^2}(I_z) = \{(\frac{1}{2}, -\frac{1}{2})\}$. Over the projective space, the homogeneous coordinates of this point are $\{(\frac{1}{2} : -\frac{1}{2} : 1)\} \subset \mathbb{V}_{\mathbb{P}^2}(I)$.

The affine variety defined by $\mathbb{V}_{\mathbb{P}^2}(I) \cap U_y$ is isomorphic $I_y := \{x^2 - 1, x - 1 - z\}$ by setting $y = 1$. In this case $\mathbb{V}_{\mathbb{K}^2}(I_y) = \{(1, 0), (-1, -2)\}$. Over the projective spaces, we write these points in homogeneous coordinates as $\{(1 : 1 : 0), (-1 : 1 : -2)\} \subset \mathbb{V}_{\mathbb{P}^2}(I)$. Note that in this case, we obtained two points as we had expected in Ex. 2.5.1.

The affine variety defined by $\mathbb{V}_{\mathbb{P}^2}(I) \cap U_x$ is isomorphic $I_x := \{1-y^2, 1-y-z\}$ by setting $x = 1$. In this case $\mathbb{V}_{\mathbb{K}^2}(I_x) = \{(1,0),(-1,2)\}$. Over the projective spaces, we write these points in homogeneous coordinates as $\{(1:1:0),(1:-1:2)\} \subset \mathbb{V}_{\mathbb{P}^2}(I)$.

Hence, we conclude that $\{(1:1:0),(\frac{1}{2}:-\frac{1}{2}:1)\} = \mathbb{V}_{\mathbb{P}^2}(I)$ as $\mathbb{V}_{\mathbb{P}^2}(I) \cap U_x \subset \mathbb{V}_{\mathbb{P}^2}(I) \cap U_y = \mathbb{V}_{\mathbb{P}^2}(I) \cap U_z$. The point $(1:1:0)$ is the solution at infinity that we mention in Ex. 2.5.1.

Many definitions and properties from affine varieties extend naturally to the projective case.

Proposition 2.5.13 (Irreducible variety)**.** *[CLO15, Ex. 8.3.11] We say that a projective variety is irreducible if and only if we cannot write it as a union of two projective (strictly) subvarieties.*

The projective variety V is irreducible if and only if the homogeneous ideal $\mathbb{I}(V)$ is prime.

Proposition 2.5.14 (Minimal irreducible decomposition)**.** *Given a projective variety $V \subset \mathbb{K}^n$, there are unique irreducible projective varieties $V_1, \ldots, V_r$ such that $(\forall i \neq j)\, V_i \not\subset V_j$ and*

$$V = V_1 \cup \cdots \cup V_r.$$

Definition 2.5.15 (Dimension)**.** *The dimension of an projective variety $V \subseteq \mathbb{K}^n$, $\dim(V)$, is the supremum of all the integers $n \geq 0$ such that there is a chain $W_0 \subset W_1 \subset \cdots \subset W_n \subseteq V$ of different irreducible projective varieties contained in V.*

As we did in the affine case, we can relate the geometric dimension of a projective variety with the Krull dimension of its *homogeneous coordinate ring*.

Definition 2.5.16 (Homogeneous coordinate ring)**.** *The homogeneous coordinate ring of a projective variety $V \subset \mathbb{P}^n$ is the quotient ring $\mathbb{K}[\boldsymbol{x}][x_0]/\mathbb{I}(V)$. We denote this quotient ring as $\mathbb{K}[V]$.*

Note that the homogeneous coordinate ring of V is isomorphic to the coordinate ring of the affine cone of V, $\mathbb{K}[C_V]$, see [Har77, Ex. I.2.10]

Proposition 2.5.17. *[Har77, Ex. I.2.6,Ex. I.2.10] The dimension of a projective variety V equals the Krull dimension of its homogeneous coordinate ring minus 1, that is,*

$$\dim(V) = \dim_{Krull}(\mathbb{K}[V]) - 1.$$

As in the affine case, see Def. 2.4.20, we say that a projective variety is *equidimensional* when all its irreducible components have the same dimension. We can extend Prop. 2.4.22 to the projective case. In this case, the theorem is stronger because we can tell when the intersection of two projective varieties is non-empty.

Proposition 2.5.18. *[Har77, Prop. I.7.2] Consider two equidimensional projective varieties $V, W \subset \mathbb{P}^n$. Consider the irreducible decomposition of $V \cap W = U_1 \cup \cdots \cup U_r$, see Prop. 2.5.14. Then, the dimension of each irreducible projective component U_i is lower bounded by $\dim(V) + \dim(W) - n$, that is,*

$$(\forall i)\ \dim(U_i) \geq \dim(V) + \dim(W) - n.$$

Moreover, if $\dim(V) + \dim(W) - n \geq 0$, the intersection $V \cap W$ is not empty.

For projective varieties, we can have a stronger form of Cor. 2.4.23.

Corollary 2.5.19. *Consider a homogeneous ideal $I = \langle f_1, \ldots, f_r \rangle \subset \mathbb{K}[\boldsymbol{x}]$ such that $r \leq n$ and each f_i is homogeneous. Then, $\mathbb{V}_{\mathbb{P}^n}(I)$ is not empty and its dimension is at least $n - r$.*

As we discussed at the end of Sec. 2.4, similarly to the univariate case, we expect to have a relation between the degrees of the polynomials defining a zero-dimensional variety and the number of points in this variety. We already saw that, when we work with affine varieties, such a relation does not exist, see Ex. 2.4.26. In what follows, we introduce the concept of *degree* of a projective variety which formalizes this expected relation.

Let $V \subset \mathbb{P}^n$ be a projective variety. If V is zero-dimensional, we define its *degree* as the number of points that V contains. When V is not zero-dimensional, its degree corresponds to the expected number of points, counted with multiplicity, when we intersect V with $\dim(V)$ generic hyperplanes, see Sec. 2.13. We present an equivalent definition of *degree of a projective variety*. Our definition relies on the Hilbert series of its homogeneous coordinate ring, see Def. 2.3.17.

Definition 2.5.20 (Degree of a projective variety)**.** *Let $V \subset \mathbb{P}^n$ be a projective variety. Consider the Hilbert series of its homogeneous coordinate ring, $HS_{\mathbb{K}[V]} = \frac{H(t)}{(1-t)^d}$, where d is equals the Krull dimension of $\mathbb{K}[V]$ and $H(t) \in \mathbb{Q}[t]$ is a polynomial such that $H(1) \neq 0$, see Prop. 2.3.20. We define the* degree *of V as $\deg(V) = H(1)$.*

The degree of a zero-dimensional projective variety is an upper bound on the number points that it contains. This bound is tight if and only if the variety is irreducible, see Prop. 2.4.24.

2.6 Homological algebra

The results of this section come mainly from [CLO06, Ch. 6], [Eis04, Ap. 3] and [Lan02, Ch. XX & Ch. XXI].

First, we introduce chain complexes and resolutions. Roughly speaking, a chain complex is a sequence of modules together with maps between them, such that each map generates part of the kernel of the next map. A *resolution* of a module M is a chain complex where each map generates exactly the kernel of the next map and the image of the last map is M. In this way, we can study M by "approximating" it by a sequences of modules. From a *minimal* resolution of a graded module M, we can deduce its Hilbert series, its Betti numbers, and its Castelnuovo-Mumford regularity.

Definition 2.6.1 (Chain complex)**.** *A chain complex $(M_\bullet, \delta_\bullet)$ is a sequence of R-modules $\{M_i\}_{i\in\mathbb{Z}}$ together with a sequence of homomorphisms between them $\delta_\bullet = \{\delta_i : M_i \to M_{i-1}\}_{i\in\mathbb{Z}}$, such that $\forall i \in \mathbb{Z}$, $\delta_{i-1} \circ \delta_i = 0$, that is, $\mathtt{Im}(\delta_i) \subset \mathtt{Ker}(\delta_{i-1})$. We usually write $M_\bullet$ as*

$$M_\bullet : \ldots \xrightarrow{\delta_{i+1}} M_i \xrightarrow{\delta_i} M_{i-1} \xrightarrow{\delta_{i-1}} \ldots$$

We say that a chain complex is bounded when there are two constants $a, b \in \mathbb{Z}$, such that, if $i < a$ or $i > b$, then $M_i = 0$. In this case, we write $M_\bullet$ as

$$M_\bullet : 0 \to M_b \xrightarrow{\delta_b} \ldots \xrightarrow{\delta_{i+1}} M_i \xrightarrow{\delta_i} M_{i-1} \xrightarrow{\delta_{i-1}} \ldots \xrightarrow{\delta_{a-1}} M_a \to 0$$

Definition 2.6.2 (Exact chain complex)**.** *We say that a complex is exact when,*

$$\forall i \in \mathbb{Z},\ \texttt{Im}(\delta_i) = \texttt{Ker}(\delta_{i-1}).$$

We can characterize how far a chain complex is from being exact by studying its homologies.

Definition 2.6.3 (Homology)**.** *The* i-th homology *of a chain complex* $(M_\bullet, \delta_\bullet)$ *is the quotient between the kernel and the image of consecutive maps, that is*

$$H_i := \texttt{Ker}(\delta_i)/\texttt{Im}(\delta_{i+1}).$$

A chain complex is an exact sequence *if and only if every homology vanishes.*

Definition 2.6.4 (Free Resolution)**.** *Given a* R*-module* M *a free resolution for* M *is a chain complex* $(M_\bullet, \delta_\bullet)$ *of the form*

$$M_\bullet : \cdots \to M_2 \to M_1 \to M_0 \to 0,$$

where,

- *the* 0*-homology equals* M*, that is* $H_0 = M$*,*
- *for* $i > 0$*, the* i*-homology vanishes, that is,* $H_i = 0$*, and*
- *every module* M_i *is free, that is, each module* M_i *is isomorphic to* R^{k_i}*, where* k_i *is the rank of* M_i*.*

The augmented free resolution *of* M*, is the exact sequence*

$$M_\bullet : \cdots \to M_2 \to M_1 \to M_0 \to M \to 0$$

Proposition 2.6.5. *[Eis04, Cor. 19.6] Every finitely generated module over* $\mathbb{K}[\boldsymbol{x}]$ *has a finite free resolution.*

The resolution of an ideal $I \subset \mathbb{K}[\boldsymbol{x}]$ is closely related to the resolution of $\mathbb{K}[\boldsymbol{x}]/I$.

Proposition 2.6.6. *[CLO06, Ex. 6.1.8] Consider a free module* $M \subset R^m$ *and let* $(M_\bullet, \delta_\bullet)$ *be the free resolution of* M*. Then, the following chain complex is an augmented free resolution of* R^m/M*.*

$$\cdots \to M_2 \to M_1 \to M_0 \to R^m \to R^m/M \to 0.$$

When we consider the free resolution of a graded module the resolution can inherit the grading. For this, we have to twist the modules so that the maps in the resolution are graded and have degree zero. We illustrate this in Ex. 2.7.6 by considering the Koszul complex of a sequence of three homogeneous polynomials in $\mathbb{K}[\boldsymbol{x}]$. Among the resolutions of a module, we are interested in the *minimal* ones.

Definition 2.6.7 (Minimal free resolution)**.** *We say that a graded free resolution for a* $\mathbb{Z}$*-graded* $\mathbb{K}[\boldsymbol{x}][x_0]$*-module* M*,* $(M_\bullet, \delta_\bullet)$*, is minimal if*

$$(\forall i)\ \texttt{Im}(\delta_i) \subset \langle x_0, \ldots, x_n \rangle M_{i-1},$$

where $\langle x_0, \ldots, x_n \rangle M_{i-1}$ *is the module* $x_0 M_{i-1} + \cdots + x_n M_{i-1}$*.*

Minimal free resolutions are unique up to isomorphism, see [Eis05, Thm. 1.1.6].

Definition 2.6.8 (Strand of a complex)**.** *Consider a chain complex $(M_\bullet, \delta_\bullet)$ such that all the modules M_i are L-graded and all the maps δ_i are L-graded of degree $0 \in L$, see Ex. 2.3.13. Then, for each $m \in L$, the* strand *of $(M_\bullet, \delta_\bullet)$ is the chain complex $(M_\bullet, \delta_\bullet)_m$, which correspond to the restriction of $(M_\bullet, \delta_\bullet)$ to the m-graded pieces of the modules, that is,*

$$(M_\bullet)_m : \ldots \xrightarrow{\delta_{i+1}} (M_i)_m \xrightarrow{\delta_i} (M_{i-1})_m \xrightarrow{\delta_{i-1}} \ldots$$

We warn the reader that in the bibliography, for example in [Eis04, Eis05], the word strand is used in a different, but closely related, sense.

Let us also briefly mention the *determinant of a complex*. For a particular class of bounded complexes, called *generically exact* [GKZ08, App. A], we can extend the definition of the determinant of matrices to complexes. The non-vanishing of this determinant is related to the exactness of the complex. When there are only two non-zero free modules in the complex, that is, all the other modules are zero, we can define the determinant of the complex if and only if both the non-zero free modules have the same rank. In this case, the determinant of the complex reduces to the determinant of the (matrix of the) map between the two non-zero free modules, [GKZ08, Prop. A.8]. We refer the reader to [AMS09] for an accessible introduction to the determinant of a complex and to [GKZ08, App. A] for a complete formalization.

2.7 Regular sequences and Koszul complex

The results of this section come mainly from [Eis04, Ch. 17], [Lan02, Ch. XXI] and [BH98, Ch. 1].

A regular sequence is a sequence of polynomials that shares many algebraic properties with a sequence of unknowns. The Koszul complex is a chain complex that it is also a resolution for regular sequences.

Definition 2.7.1 (Regular sequence)**.** *An element $f \in \mathbb{R}$ is* regular *in a ring R if it is not a zero divisor, that is $(\forall r \in R \setminus \{0\})\ r\, f \neq 0$. Given elements $f_1, \ldots, f_r \in R$, we say that $(f_1, \ldots, f_r)$ is a regular sequence if*

- *$R/\langle f_1, \ldots, f_r\rangle \neq 0$, and*
- *for each $1 \leq i \leq r$, f_i is a regular element in the quotient ring $R/\langle f_1, \ldots, f_{i-1}\rangle$.*

Note that for regular sequences the order of the polynomials in the sequence matters. In Cor. 2.7.10, we will discuss for which regular sequences, when we permute their polynomials, the resulting sequences are also regular.

Remark 2.7.2. *If $(f_1, \ldots, f_r)$ is a regular sequence over R, then for $i \leq r$, $(f_1, \ldots, f_i)$ is a regular sequence too.*

If $(f_1, \ldots, f_r) \subset \mathbb{K}[\boldsymbol{x}]$ is a regular sequence, then the Koszul complex provides a resolution for $\mathbb{K}[\boldsymbol{x}]/\langle f_1, \ldots, f_r\rangle$.

Definition 2.7.3. *Let E be a $\mathbb{K}$-vector space generated by $\{e_1, \dots, e_n\}$. We define the map Φ_s as the (-1)-graded map such that, for each i and $1 \leq j_1 < \cdots < j_i \leq n$,*

$$\begin{array}{rccl} \Phi_s : & \bigwedge^i E & \to \bigwedge^{i-1} E & \\ & e_{j_1} \wedge \cdots \wedge e_{j_i} & \mapsto \Phi_s(e_{j_1} \wedge \cdots \wedge e_{j_i}) = \begin{cases} (-1)^{k+1} e_{j_1} \wedge \cdots \wedge e_{j_{k-1}} \wedge e_{j_{k+1}} \wedge \cdots \wedge e_{j_i} & \text{if } j_k = s \\ 0 & \text{otherwise} \end{cases} \end{array}$$

Definition 2.7.4 (Multiplication operator)**.** *For each $f \in \mathbb{K}[\boldsymbol{x}]_d$, we define the multiplication operator μ_f, which is a graded homomorphism of degree d, that corresponds to the multiplication by f.*

$$\begin{array}{rcl} \mu_f : \mathbb{K}[\boldsymbol{x}]_D & \to \mathbb{K}[\boldsymbol{x}]_{D+d} \\ g & \mapsto f\, g \end{array}$$

Definition 2.7.5 (Koszul complex)**.** *Let E be a $\mathbb{K}$-vector space of dimension r, with basis $e_1, \dots, e_r$. Given a set of polynomials $f_1, \dots, f_r \in R$, we define the* Koszul complex $\mathcal{K}_\bullet(f_1, \dots, f_r)$ *as the chain complex where $\mathcal{K}_0 = R$, $\mathcal{K}_i = R \otimes \bigwedge^i E \cong R^{\binom{r}{i}}$, when $r \geq i \geq 1$, or 0 otherwise.*

The differential $\delta_i : \mathcal{K}_i \to \mathcal{K}_{i-1}$ are homomorphisms such that

$$\delta_i = \sum_{s=1}^{r} \mu_{f_s} \Phi_s.$$

Equivalently, for each $1 \leq j_1 < \cdots < j_i \leq r$ and $g \in \mathbb{K}[\boldsymbol{x}]$,

$$\begin{aligned} \delta_i(g \otimes e_{j_1} \wedge \cdots \wedge e_{j_i}) &= \sum_{s=1}^{r} \mu_{f_s}(g) \otimes \Delta_s(e_{j_1} \wedge \cdots \wedge e_{j_i}) \\ &= \sum_{k=1}^{i} (-1)^{k+1} f_{j_k}\, g \otimes e_{j_1} \wedge \cdots \wedge e_{j_{k-1}} \wedge e_{j_{k+1}} \wedge \cdots \wedge e_{j_i}. \end{aligned}$$

When the polynomials $f_1, \dots, f_r$ are homogeneous, then the modules of the Koszul complex are homogeneous too, $\mathcal{K}_i \cong \bigoplus_{1 \leq j_1 < \cdots < j_i \leq r} R(-j_1 - \cdots - j_i) \otimes e_{j_1} \wedge \cdots \wedge e_{j_i}$.

We denote by $\mathcal{H}_i(f_1, \dots, f_r)$ the i-th homology of $\mathcal{K}(f_1, \dots, f_r)$, that is,

$$\mathcal{H}_i(f_1, \dots, f_r) = \mathtt{Ker}(\delta_i)/\mathtt{Im}(\delta_{i+1}).$$

Example 2.7.6. *Let $f_1, f_2, f_3 \in \mathbb{K}[\boldsymbol{x}]$ be a sequence of polynomials. The Koszul complex of f_1, f_2, f_3 is as follows:*

$$0 \to R \otimes e_1 \wedge e_2 \wedge e_3 \xrightarrow[\delta_3]{\begin{bmatrix} f_3 & -f_2 & f_1 \end{bmatrix}} \begin{array}{c} R \otimes e_1 \wedge e_2 \\ \oplus \\ R \otimes e_1 \wedge e_3 \\ \oplus \\ R \otimes e_2 \wedge e_3 \end{array} \xrightarrow[\delta_2]{\begin{bmatrix} -f_2 & f_1 & 0 \\ -f_3 & 0 & f_1 \\ 0 & -f_3 & f_2 \end{bmatrix}} \begin{array}{c} R \otimes e_1 \\ \oplus \\ R \otimes e_2 \\ \oplus \\ R \otimes e_3 \end{array} \xrightarrow[\delta_1]{\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}} R \to 0$$

If we assume that the polynomials $f_1, f_2, f_3 \in \mathbb{K}[\boldsymbol{x}]$ are homogeneous of degree d_1, d_2, d_3, then we can make the Koszul complex 0-graded.

To do so, we start analyzing the last map of the Koszul complex, δ_1. This map is equivalent to the map $(g_1, g_2, g_3) \mapsto \sum_{k=1}^{3} f_k g_k$. In order to make δ_1 graded, the degrees of $f_1 g_1, f_2 g_2, f_3 g_3$ must be the same. Hence, if the degree of $\sum_{k=1}^{3} f_k g_k$ is d, the degree of g_k must be $d - d_k + \deg(\delta_1)$. If we want the degree of δ_1 to be zero, then the first map of the Koszul complex must be

$$R(-d_1) \oplus R(-d_2) \oplus R(-d_3) \xrightarrow[\delta_1]{\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}} R.$$

Following the same reasoning, the second map of the Koszul complex is the map

$$(h_{1,2}, h_{1,3}, h_{2,3}) \mapsto (-h_{1,2}\, f_2 - h_{1,3}\, f_3,\ h_{1,2}\, f_1 - h_{2,3}\, f_3,\ h_{1,3}\, f_1 + h_{2,3}\, f_2).$$

For this map to be 0-graded, we should find constants $D_{1,2}, D_{1,3}, D_{2,3}$ such that, for $d \in \mathbb{Z}$, it holds

$$\delta_2 \left((R(-D_{1,2}) \oplus R(-D_{1,3}) \oplus R(-D_{2,3}))_d \right) \subset (R(-d_1) \oplus R(-d_2) \oplus R(-d_3))_d$$

So, for $h_{1,2} \in R(-D_{1,2})_d$ and $h_{1,3} \in R(-D_{1,3})_d$, we want $h_{1,2}\, f_2 + h_{1,3}\, f_3 \in R(-d_1)_d$. Hence, the degree of $h_{1,2}\, f_2$ should be equal to the degree of $h_{1,3}\, f_3$, that is $-d_1 + d$. Thus,

$$d - D_{1,2} + d_2 = d - D_{1,3} + d_3 = d - d_1.$$

Using the same arguments for $h_{1,3}$ and $h_{2,3}$, we conclude that for each d, we want $D_{1,2}, D_{1,3}, D_{2,3}$ such that,

$$\begin{cases} d - D_{1,2} + d_2 = d - D_{1,3} + d_3 = d - d_1 \\ d - D_{1,2} + d_1 = d - D_{2,3} + d_2 = d - d_2 \\ d - D_{1,3} + d_1 = d - D_{2,3} + d_3 = d - d_3 \end{cases},$$

which leads to

$$\begin{cases} D_{1,2} = d_1 + d_2 \\ D_{1,3} = d_1 + d_3 \\ D_{2,3} = d_2 + d_3 \end{cases}.$$

Hence, the map δ_2 is a graded map if we consider the following graded modules

$$\begin{matrix} R(-d_1 - d_2) \\ \oplus \\ R(-d_1 - d_3) \\ \oplus \\ R(-d_2 - d_3) \end{matrix} \xrightarrow[\delta_2]{\begin{bmatrix} -f_2 & f_1 & 0 \\ -f_3 & 0 & f_1 \\ 0 & -f_3 & f_2 \end{bmatrix}} \begin{matrix} R(-d_1) \\ \oplus \\ R(-d_2) \\ \oplus \\ R(-d_3). \end{matrix}$$

Finally, we do the same analysis for δ_3, which represents the map $g \mapsto (g\, f_3, -g\, f_2, g\, f_1)$. We want to find a D such that, for every d, it holds

$$\delta_3(R(-D)_d) \subset R(-d_1 - d_2) \oplus R(-d_1 - d_3) \oplus R(-d_2 - d_3).$$

Equivalently, we are looking for a D such that, for any d and $g \in R(-D)_d$, it holds $(g\, f_3, -g\, f_2, g\, f_1) \in (R(-d_1-d_2) \oplus R(-d_1-d_3) \oplus R(-d_2-d_3))_d$. Hence, the degrees must satisfy the equations,

$$\begin{cases} d - D = d - d_2 - d_3 \\ d - D = d - d_1 - d_3 \\ d - D = d - d_1 - d_2 \end{cases}$$

Therefore, $D = d_1 + d_2 + d_3$, which results to

$$R(-d_1-d_2-d_3) \xrightarrow[\delta_3]{\begin{bmatrix} f_3 & -f_2 & f_1 \end{bmatrix}} \begin{matrix} R(-d_1-d_2) \\ \oplus \\ R(-d_1-d_3) \\ \oplus \\ R(-d_2-d_3) \end{matrix}$$

Hence, our graded Koszul complex is isomorphic to,

$$0 \to R(-d_1-d_2-d_3) \xrightarrow[\delta_3]{\begin{bmatrix} f_3 & -f_2 & f_1 \end{bmatrix}} \begin{matrix} R(-d_1-d_2) \\ \oplus \\ R(-d_1-d_3) \\ \oplus \\ R(-d_2-d_3) \end{matrix} \xrightarrow[\delta_2]{\begin{bmatrix} -f_2 & f_1 & 0 \\ -f_3 & 0 & f_1 \\ 0 & -f_3 & f_2 \end{bmatrix}} \begin{matrix} R(-d_1) \\ \oplus \\ R(-d_2) \\ \oplus \\ R(-d_3) \end{matrix} \xrightarrow[\delta_1]{\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}} R \to 0$$

Remark 2.7.7. *The 0-th homology of the Koszul complex $\mathcal{K}(f_1,\dots,f_r)$ is $\mathcal{H}_0(f_1,\dots,f_r) = R/\langle f_1,\dots,f_r\rangle$.*

Regular sequences are closely related to the Koszul complex.

Proposition 2.7.8 (Koszul complex of a regular sequences)**.** *[Lan02, Thm. XXI.4.6] Consider $f_1,\dots,f_r \in \mathbb{K}[\boldsymbol{x}]$. If $(f_1,\dots,f_r)$ is a regular sequence over $\mathbb{K}[\boldsymbol{x}]$, then $\mathcal{H}_i(f_1,\dots,f_r) = 0$, for $i > 0$.*

If $f_1,\dots,f_r$ are homogeneous polynomials, then converse statement also holds, that is, if $\mathcal{H}_i(f_1,\dots,f_r) = 0$, for $i > 0$, then $(f_1,\dots,f_r)$ is a regular sequence over $\mathbb{K}[\boldsymbol{x}]$. Moreover, when the polynomials are homogeneous, if $\mathcal{H}_1(f_1,\dots,f_r) = 0$ then $\mathcal{H}_j(f_1\dots,f_r) = 0$, for $j > 1$.

In particular, if I is generated by a regular sequence, then the Koszul complex provide us with an augmented free resolution for $\mathbb{K}[\boldsymbol{x}]/I$.

Corollary 2.7.9. *If $(f_1,\dots,f_r)$ is a regular sequence over $\mathbb{K}[\boldsymbol{x}]$, then the Koszul complex is an augmented free resolution of the quotient ring $\mathbb{K}[\boldsymbol{x}]/\langle f_1,\dots,f_r\rangle$. Moreover, the following chain complex is an augmented free resolution of I,*

$$0 \to \cdots \to \mathcal{K}_2 \to \mathcal{K}_1 \to I \to 0.$$

If $(f_1,\dots,f_r)$ are homogeneous, then the free resolutions are minimal*, see Def. 2.6.7.*

The Koszul complex does not take into account the order of the polynomials $f_1,\dots,f_r$ in the sequence. Hence, when we have a homogeneous regular sequence, any permutation of its polynomials result a regular sequence.

Corollary 2.7.10 (Permutation of homogeneous regular sequences)**.** *Let $\sigma \in S_r$, where S_r is the symmetric group of $\{1,\dots,r\}$, that is, σ is a permutation of the set $\{1,\dots,r\}$. Hence, if $f_1,\dots,f_r$ are homogeneous polynomials and $(f_1,\dots,f_r)$ is a regular sequence, then $(f_{\sigma(1)},\dots,f_{\sigma(r)})$ is a regular sequence, too.*

When a regular sequence does not consist of homogeneous polynomials, then a permutation of its elements might not result in a regular sequence.

Example 2.7.11 (Permutation of affine regular sequences)**.** *Consider the polynomial sequence $(z, x\,(z+1), y\,(z+1))$ in $\mathbb{K}[x,y,z]$. This sequence is regular because $\mathbb{K}[x,y,z]/\langle z\rangle \cong \mathbb{K}[x,y]$ and $x\,(z+1) = x$ and $y\,(z+1) = y$ in $\mathbb{K}[x,y,z]/\langle z\rangle$. But $(x\,(z+1), y\,(z+1), z)$ is not a regular sequence as $y\,(z+1)$ is a zero divisor in $\mathbb{K}[x,y,z]/\langle x\,(z+1)\rangle$, $x\,(y\,(z+1)) = y\,(x\,(z+1)) = 0$ in $\mathbb{K}[x,y,z]/\langle x\,(z+1)\rangle$. Hence, Cor. 2.7.10 might not hold when the polynomials are not homogeneous.*

An ideal of $\mathbb{K}[\boldsymbol{x}]$ generated by a regular sequence of r elements describe an equidimensional variety of dimension $n-r$.

Proposition 2.7.12. *[Per07, Thm. IV.2.3] Consider polynomials $f_1,\dots,f_r \in \mathbb{K}[\boldsymbol{x}]$ such that $(f_1,\dots,f_r)$ is a regular sequence. Then, if $\mathbb{V}_{\mathbb{K}^n}(\langle f_1,\dots,f_r\rangle)$ is not empty, then it is equidimensional and has dimension $n-r$.*

Moreover, if the ideal is homogeneous, then the previous condition is an "if a only if".

Proposition 2.7.13. *[Hoc16, Page 15] Consider homogeneous polynomials $f_1,\dots,f_r \in \mathbb{K}[\boldsymbol{x}][x_0]$. Then, the following three conditions are equivalent:*

- *$(f_1,\dots,f_r)$ is a regular sequence.*
- *$\dim(\mathbb{V}_{\mathbb{P}^n}(\langle f_1,\dots,f_r\rangle)) = n-r$.*
- *$\dim(\mathbb{V}_{\mathbb{P}^n}(\langle f_1,\dots,f_r\rangle)) = n-r$ and $\mathbb{V}_{\mathbb{P}^n}(\langle f_1,\dots,f_r\rangle)$ is equidimensional.*

For regular sequences, because we know their minimal resolution, we can also predict their Hilbert series and their degree [CLO06, Thm 6.4.4].

Proposition 2.7.14 (Hilbert series of a regular sequence)**.** *[CLO15, Lem. 10.2.5] Let $(f_1,\dots,f_r)$ be a regular sequence over $\mathbb{K}[\boldsymbol{x}]$ given by homogeneous polynomials of degrees $d_1,\dots,d_r$, respectively, then the Hilbert series of the corresponding quotient ring is*

$$HS_{\mathbb{K}[\boldsymbol{x}]/\langle f_1,\dots,f_r\rangle} = \frac{\prod_{i=1}^{r}(1-t^{d_i})}{(1-t)^n} = \frac{\prod_{i=1}^{r}\sum_{j=0}^{d_i-1} t^j}{(1-t)^{n-r}}$$

Corollary 2.7.15 (Bézout bound)**.** *Let $(f_1,\dots,f_r)$ be a regular sequence over $\mathbb{K}[\boldsymbol{x}]$ given by homogeneous polynomials of degrees $d_1,\dots,d_r$, respectively, then the degree of the variety $\mathbb{V}_{\mathbb{P}^n}(\langle f_1,\dots,f_r\rangle)$, see Def. 2.5.20, is the product of the degrees of the polynomials, that is,*

$$\texttt{Degree}(\mathbb{V}_{\mathbb{P}^n}(\langle f_1,\dots,f_r\rangle)) = \left((1-t)^{n-r}\frac{\prod_{i=1}^{r}\sum_{j=0}^{d_i-1} t^j}{(1-t)^{n-r}}\right)\Bigg|_{t=1} = \prod_{i=1}^{r} d_i.$$

In particular, if $r=n$, then the system $\{f_1,\dots,f_n\}$ has at most $\prod_{i=1}^{n} d_i$ different solutions.

The Bézout bound is tight when the ideal $\langle f_1,\dots,f_n\rangle$ is radical or if we count the number of solutions taking into account their *multiplicities*, see [CLO06, Sec. 4.2].

2.8 Betti numbers and Castelnuovo-Mumford regularity

The results of this section come mainly from [Eis04, Ch. 20], [CLO06, Ch. 6] and [BH98, Ch. 1 & Ch. 4].

Among the free resolutions of graded modules we can always find a minimal free resolution, which is unique up to isomorphism, see Def. 2.6.7. From the minimal resolution we can read the Betti numbers, which are important invariants of a module. Using the Betti numbers, we can recover the Hilbert series and define the Castelnuovo-Mumford regularity.

Proposition 2.8.1 (Betti numbers)**.** *[Eis05, Thm. 1.1.6] If there is a finite minimal free resolution for a graded R-module M, then the minimal free resolution is unique up to isomorphisms. In this case, we can always write it as*

$$0 \to \bigoplus_{j\in\mathbb{Z}} R(-j)^{\beta_{pd(M),j}} \to \cdots \to \bigoplus_{j\in\mathbb{Z}} R(-j)^{\beta_{0,j}} \to M \to 0,$$

where only a finite number of $\beta_{i,j}$ are not zero.

The $\{\beta_{i,j}\}_{i,j}$ are the graded Betti numbers.

The graded Betti numbers define the Hilbert Series.

Proposition 2.8.2. *[BH98, Lem. 4.1.13] Let $\{b_{i,j}\}$ be the graded Betti numbers of a graded R-module M, then*

$$HS_M(t) = HS_R(t) \sum_{i=0}^{n} \sum_{j\in\mathbb{Z}} (-1)^i \beta_{i,j}\, t^j.$$

When $R = \mathbb{K}[\boldsymbol{x}]$, then $HS_R(t) = \frac{1}{(1-t)^n}$, and so

$$HS_M(t) = \frac{1}{(1-t)^n} \sum_{i=0}^{n} \sum_{j\in\mathbb{Z}} (-1)^i \beta_{i,j} t^j.$$

Also, the graded Betti numbers allow us to define the Castelnuovo-Mumford regularity. As we will see in Sec. 4.5, this regularity is related to the complexity of solving polynomial systems.

Definition 2.8.3 (Castelnuovo-Mumford regularity)**.** *Consider a graded R-module M and let $\{\beta_{i,j}\}$ be its set of graded Betti numbers. Then, the Castelnuovo-Mumford regularity of M is,*

$$reg_{\mathfrak{CM}}(M) = \max_{i,j}\{j - i : \beta_{i,j} \neq 0\}.$$

The Castelnuovo-Mumford regularity tell us when the Hilbert polynomial and the Hilbert function agree.

Corollary 2.8.4. *[Eis05, Thm. 4.2] [BS13, Ex. 17.1.10] If $d > reg_{\mathfrak{CM}}(M)$, then the Hilbert polynomial and Hilbert function agree, that is, $HP_M(d) = HF_M(d)$.*

For homogeneous regular sequences $(f_1, \ldots, f_r)$, the Koszul complex give us minimal free resolutions for $\langle f_1, \ldots, f_r\rangle$ and $\mathbb{K}[\boldsymbol{x}]/\langle f_1, \ldots, f_r\rangle$, see Cor. 2.7.9. Hence, we can deduce the Castelnuovo-Mumford regularity.

Proposition 2.8.5 (Macaulay bound)**.** *Let $(f_1, \ldots, f_r)$ be a homogeneous regular sequence in $\mathbb{K}[\boldsymbol{x}]$ where the polynomials have degrees $d_1, \ldots, d_r$, respectively. Then, the Castelnuovo-Mumford regularity of the ideal $\langle f_1, \ldots, f_r\rangle$ and the corresponding quotient ring is*

$$\begin{cases} reg_{\mathfrak{CM}}(\langle f_1, \ldots, f_r\rangle) = \sum_{i=1}^{r} d_i - r + 1. & (\textit{Macaulay bound, } \texttt{MB}) \\ reg_{\mathfrak{CM}}(\mathbb{K}[\boldsymbol{x}]/\langle f_1, \ldots, f_r\rangle) = \sum_{i=1}^{r} d_i - r. \end{cases}$$

2.9 Local cohomology and Castelnuovo-Mumford regularity

The results of this section come mainly from [BS13, Sec. 5.1], [Bus06, Ch .3], and [BH98, Sec. 3.5]

In this section we will talk about local cohomology and its relation with the Castelnuovo-Mumford regularity. For this, we need to introduce *localization*. We say that a set $U \subset R$ is *multiplicatively closed* if $1 \in U$ and $(\forall a, b \in U)\ a\,b \in U$. Roughly speaking, to localize a module M at a multiplicative set U is like inverting the elements of U in M.

Definition 2.9.1 (Localization)**.** *Given a R-module M and a multiplicatively closed set $U \subset M$, we define the localization of M at U, $M[U^{-1}]$, as the set of equivalence classes of pairs (f, u), with $f \in M$ and $u \in U$. The equivalence relation $\sim$ is such that $(f, u) \sim (f', u')$ if and only if there is an element $v \in U$ such that $v\,(u'\,f - u\,f') = 0$ in M.*

Proposition 2.9.2. *Consider $x \in U \subset R$ and $f \in M$ such that $x\,f = 0$ in M. Then $f = 0$ in $M[U]$.*

If the ring R is $\mathbb{Z}^k$-graded and the elements of the multiplicatively closed set U are homogeneous with respect to the same $\mathbb{Z}^k$-grading, then the localization $R[U^{-1}]$ is $\mathbb{Z}^k$-graded and the degree of $(f, u) \in R[U^{-1}]$ is $\deg(f) - \deg(u)$.

In most of the cases, we localize with respect to an element or the complement of a prime ideal.

Definition 2.9.3 (Localization at an element)**.** *Given an element $x \in R$, we define $M[x^{-1}]$ as the localization of M at the multiplicative closed set $U = \{x^i : i \in \mathbb{N}\}$ (we consider $0 \in \mathbb{N}$).*

Given a prime ideal $\mathfrak{p} \subset R$, the complement (set theoretically) of the ideal, $R \setminus \mathfrak{p}$, is a multiplicative closed set.

Definition 2.9.4 (Localization at a prime ideal)**.** *Given a prime ideal $\mathfrak{p} \subset R$, we define $M[\mathfrak{p}]$ as the localization of M at the multiplicative closed set $U = R \setminus \mathfrak{p}$.*

Similarly to the Koszul complex, the *Čech complex* is another central chain complex in algebraic geometry.

Definition 2.9.5 (Čech complex)**.** *[Bus06, Def. 3.1.3] Let R be a ring, $\boldsymbol{f} := (f_1, \ldots, f_r)$ be a sequence of elements in R and M a R-module. The Čech complex of $\boldsymbol{f}$ over M is the complex $\mathcal{C}^\bullet(\boldsymbol{f}; M)$ whose modules are*

$$\mathcal{C}^0(\boldsymbol{f}, M) := M \qquad \textit{and} \qquad \mathcal{C}^i(\boldsymbol{f}, M) := \bigoplus_{1 \le j_1 < \cdots < j_i \le r} M\left[\left(\prod_{s=1}^{i} f_{j_s}\right)^{-1}\right]\ (1 \le i \le r)$$

and the maps $d^i : \mathcal{C}^i(\boldsymbol{f}; M) \to \mathcal{C}^{(i+1)}(\boldsymbol{f}; M)$ *are*

$$d^0(m) = \sum_{i=1}^{r} \frac{m}{1} \text{ and } d^i(m_{j_1,\dots,j_i}) = \sum_{k \notin \{j_1,\dots,j_i\}} (-1)^{\tau(k)} \phi_k(m_{j_1,\dots,j_i})$$

where $m_{j_1,\dots,j_i} \in M\left[\left(\prod_{s=1}^{i} f_{j_s}\right)^{-1}\right]$, $\phi_k(m_{j_1,\dots,j_i}) \in M\left[\left(f_k \prod_{s=1}^{i} f_{j_s}\right)^{-1}\right]$, *and* $\tau(k)$ *is such that* $j_{\tau(k)} < k < j_{\tau(k)+1}$.

Definition 2.9.6 (Local cohomology)**.** *The* i*-th local cohomology* R*-module with support in* $\langle \boldsymbol{f} \rangle$ *is the* i*-th cohomology* R*-module* $H^i_{\langle \boldsymbol{f} \rangle}(M) = \mathtt{Ker}(d^i)/\mathtt{Im}(d^{(i-1)})$ *(*$H^0_{\langle \boldsymbol{f} \rangle}(M) = \mathtt{Ker}(d^0)$*) of the Čech complex* $\mathcal{C}^\bullet(\boldsymbol{f}; M)$.

The local cohomology encodes geometric information of the variety defined by $\langle \boldsymbol{f} \rangle$. Hence, $H^i_{\langle \boldsymbol{f} \rangle}(M)$ is independent of the generators of $\langle \boldsymbol{f} \rangle$ that we choose to construct the Čech complex and only depends on the radical of the ideal $\langle \boldsymbol{f} \rangle$. Hence, we will write the local cohomologies $H^i_I(M)$ without specifying a particular set of generators of I. In this thesis we only use some basic properties of the local cohomology; we refer the reader to [BS13], [BH98, Chp. 3.5], and [Har77, Sec. III.4] for a detailed representation.

Proposition 2.9.7. *[Bus06, Prop. 3.5] Consider two ideals over* R, $I = \langle f_1, \dots, f_r \rangle$ *and* $J = \langle g_1, \dots, g_s \rangle$, *such that* $\sqrt{I} = \sqrt{J}$. *Then* $H^i_I(M) \cong H^i_J(M)$.

The 0-th local cohomology of I over R/J is isomorphic the saturation of J with respect to I.

Proposition 2.9.8. *[Bus06, Eq. (1.1)] The* 0*-th local cohomology module is the set of elements in* M *annihilated by a power of the ideal* $I = \{f_1, \dots, f_r\}$, *that is*

$$H^0_I(M) = \{m \in M \text{ s.t. } \exists k \in \mathbb{N} : f_i^k\, m = 0 \text{ in } M \text{ for all } i = 1 \dots r\}$$

When $M = R/J$, *then the* 0*-th local cohomology is the saturation of the ideal* J *with respect to* I, *that is* $H^0_I(R/J) = (J : I^\infty)$.

Given a $\mathbb{Z}$-graded finitely generated $\mathbb{K}[\boldsymbol{x}]$-modules M, the local cohomologies of $\langle x_1, \dots, x_n \rangle$ over M vanish at big degrees.

Proposition 2.9.9. *[BS13, Thm. 16.1.5] Consider the ideal* $B = \langle x_1, \dots, x_n \rangle \in \mathbb{K}[\boldsymbol{x}]$ *and a* $\mathbb{Z}$*-graded finitely generated* $\mathbb{K}[\boldsymbol{x}]$*-module* M *with respect to the standard grading, see Ex. 2.3.3. Then*

- *For every* $i \in \mathbb{N}$ *and* $d \in \mathbb{Z}$, *the* $\mathbb{K}$*-module* $H^i_B(M)_d$ *is finitely generated.*
- *There is a* $d_0 \in \mathbb{Z}$ *such that* $H^i_B(M)_d = 0$ *for every* $i \in \mathbb{N}$ *and* $d \geq d_0$.

The difference between the Hilbert function and the Hilbert polynomial is given by the local cohomologies.

Proposition 2.9.10 (Hilbert function and local cohomology)**.** *[BH98, Thm. 4.4.3] Consider the ideal $B = \langle x_1, \ldots, x_n \rangle \in \mathbb{K}[\boldsymbol{x}]$ and a quotient ring $M = \mathbb{K}[\boldsymbol{x}]/J$, where J is a homogeneous ideal and the Krull dimension of M is t. The Hilbert function of M is the sum of its Hilbert polynomial and the alternating sums of the dimensions of the local cohomologies as $\mathbb{K}$-vector spaces, that is, for all $d \in \mathbb{N}$ we have*

$$HF_M(d) = HP_M(d) + \sum_{i=0}^{t} (-1)^i \dim_{\mathbb{K}}(H^i_B(M)_d).$$

Local cohomology give us another way of defining the Castelnuovo-Mumford regularity for the $\mathbb{Z}$-graded ring $\mathbb{K}[\boldsymbol{x}]$.

Proposition 2.9.11 (Castelnuovo-Mumford regularity using local cohomology)**.** *[BS13, Thm. 16.3.7] Consider the ideal $B = \langle x_1, \ldots, x_n \rangle \in \mathbb{K}[\boldsymbol{x}]$ and a finitely generated $\mathbb{Z}$-graded $\mathbb{K}[\boldsymbol{x}]$-module generated M, see Ex. 2.3.2. Let a_i be the maximum degree at which $H^i_B(M)$ does not vanish, that is*

$$a_i(M) := \begin{cases} \max(\{j : (H^i_B(M))_j \neq 0\}) & \text{if } H^i_B(M) \neq 0, \\ -\infty & otherwise. \end{cases}$$

The Castelnuovo-Mumford regularity is

$$reg_{\mathfrak{CM}}(M) = \max_i (a_i(M) + i).$$

The previous propositions provides a different proof of Cor. 2.8.4.

2.10 Multihomogeneous polynomials and multiprojective varieties

In this section we discuss about systems of multihomogeneous polynomial equations and their relation to the multiprojective space, that is, the Cartesian product of projective spaces. Multihomogeneous systems arise when we consider the tensor product of homogeneous polynomials. Geometrically, the solution set of such systems belongs to the Cartesian product of projective spaces. They are the easiest examples, besides monomial ideals, of multigraded systems.

Definition 2.10.1 (Multihomogeneous polynomials)**.** *We consider q blocks of variables $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q$, where $\boldsymbol{x}_i = \{x_{i,0}, \ldots, x_{i,n_i}\}$, for each $1 \leq i \leq q$. We consider the polynomial algebra $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q] := \mathbb{K}[\boldsymbol{x}_1] \otimes \cdots \otimes \mathbb{K}[\boldsymbol{x}_q]$, that is multigraded with respect to $\mathbb{Z}^q$, thus*

$$\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q] = \sum_{(d_1,\ldots,d_q)\in\mathbb{Z}^q} \mathbb{K}[\boldsymbol{x}_1]_{d_1} \otimes \cdots \otimes \mathbb{K}[\boldsymbol{x}_q]_{d_q}.$$

Given $\boldsymbol{d} \in \mathbb{Z}^q$, we denote by $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]_{\boldsymbol{d}}$ the $\boldsymbol{d}$-graded part of $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$. Given a polynomial $f \in \mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]_{\boldsymbol{d}}$, we say that it has multidegree $\deg(f) = \boldsymbol{d}$.

When it is clear from context, we refer to the *multidegree* of a multihomogeneous polynomial as its *degree*.

Example 2.10.2 (Multilinear polynomial)**.** *Consider two blocks of variables* $\boldsymbol{x} = \{x_0, x_1\}$ *and* $\boldsymbol{y} = \{y_0, y_1\}$*. The polynomials* $f_1 := x_0\,y_0 + x_0\,y_1 - x_1\,y_0$ *and* $f_2 := x_0 + x_1$ *are multihomogeneous polynomials in* $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$*. Moreover,* $f_1 \in \mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]_{(1,1)}$ *and* $f_2 \in \mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]_{(1,0)}$*. When the multidegree of a multihomogeneous polynomial is such that every block of variables has degree at most* 1*, we say that it is a* multilinear polynomial*. Both* f_1 *and* f_2 *are multilinear polynomials.*

A multihomogeneous ideal is a $\mathbb{Z}^q$-graded ideal in $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$.

Definition 2.10.3 (Multihomogeneous ideal)**.** *An ideal in* $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$ *is multihomogeneous if it is homogeneous with respect to the grading* $\mathbb{Z}^q$ *(see Def. 2.3.5).*

Similarly to the case of homogeneous ideals, see Def. 2.3.7, if $I \in \mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$ is a multihomogeneous ideal, then $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]/I$ is a $\mathbb{Z}^q$-graded module and we can define its Hilbert function. Moreover, there is a multigraded version of the Hilbert polynomial, see Prop. 2.3.18.

In the following, given two vectors $\boldsymbol{d}, \boldsymbol{d_0} \in \mathbb{Z}^q$, we say that $\boldsymbol{d}$ is component-wise bigger than $\boldsymbol{d_0}$, and write it as $\boldsymbol{d} \geq \boldsymbol{d_0}$, if and only if $\boldsymbol{d} - \boldsymbol{d_0} \in \mathbb{N}^q$.

Proposition 2.10.4 (Hilbert Polynomial and Series)**.** *[Rém01, Thm. 2.10] Let* $I \subset \mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$ *be a multihomogeneous ideal. We define the multigraded Hilbert function of* $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]/I$ *as*

$$HF_{\mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]/I}(\boldsymbol{d}) = \dim_{\mathbb{K}}((\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]/I)_{\boldsymbol{d}}).$$

There is a unique multivariate polynomial $HP_{\mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]/I} \in \mathbb{Q}[t_1, \ldots, t_q]$ *and a* $\boldsymbol{d_0} \in \mathbb{Z}^q$ *such that for* $\boldsymbol{d} \geq \boldsymbol{d_0}$ *(component-wise) the multigraded Hilbert function of* $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]/I$ *agrees with the polynomial, that is*

$$HP_{\mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]/I}(\boldsymbol{d}) = HF_{\mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]/I}(\boldsymbol{d}).$$

We call $HP_{\mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]/I}$ *the multigraded Hilbert polynomial of* $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]/I$*.*

Proposition 2.10.5. *The Hilbert polynomial of* $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$ *is*

$$HP_{\mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]}(d_1, \ldots, d_q) = \prod_{i=1}^{q} \binom{d_i + n_i}{n_i}.$$

Example 2.10.6. *Consider two blocks of variables* $\boldsymbol{x} = \{x_0, x_1, x_2\}$ *and* $\boldsymbol{y} = \{y_0, y_1, y_2\}$*. The Hilbert polynomial of* $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}] = \mathbb{K}[\boldsymbol{x}] \otimes \mathbb{K}[\boldsymbol{y}]$ *is,*

$$HP_{\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]}(i, j) = \binom{i+2}{2}\binom{j+2}{2}.$$

Consider the multihomogeneous ideal

$$I := \langle -x_0^2\,y_0\,y_2 + x_0\,x_1\,y_0\,y_1, x_1^2\,y_1^2 - x_1\,x_2\,y_0\,y_1, -x_0\,x_1\,y_2^2 + x_1\,x_2\,y_2^2 \rangle,$$

generated by polynomials in $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]_{(2,2)} = \mathbb{K}[\boldsymbol{x}]_2 \otimes \mathbb{K}[\boldsymbol{y}]_2$*. The Hilbert function of* $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]/I$ *is*

$$HF_{\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]}(i, j) = \begin{cases} \frac{1}{2}\,i^2 + 2\,i\,j + j^2 + \frac{27}{2}\,i + 11\,j - 32 & i \geq 3 \text{ and } j \geq 3 \quad (\checkmark) \\ \frac{3}{2}\,i^2 + \frac{21}{2}\,i + 6 & i \geq 2 \text{ and } j = 2 \quad (\blacksquare) \\ \frac{3}{2}\,j^2 + \frac{21}{2}\,j + 6 & i = 2 \text{ and } j > 2 \quad (\bullet) \\ \frac{1}{4}\,(i+1)\,(i+2)\,(j+1)\,(j+2) & i \leq 1 \text{ or } j \leq 1 \quad (\circ) \end{cases}$$

The following table illustrates the multidegrees corresponding to the different formulas of the Hilbert function,

$$\begin{array}{c|cccccccc} j & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 6 & \circ & \circ & \bullet & \checkmark & \checkmark & \checkmark & \checkmark & \cdots \\ 5 & \circ & \circ & \bullet & \checkmark & \checkmark & \checkmark & \checkmark & \cdots \\ 4 & \circ & \circ & \bullet & \checkmark & \checkmark & \checkmark & \checkmark & \cdots \\ 3 & \circ & \circ & \bullet & \checkmark & \checkmark & \checkmark & \checkmark & \cdots \\ 2 & \circ & \circ & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \cdots \\ 1 & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \cdots \\ 0 & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \cdots \\ & 0 & 1 & 2 & 3 & 4 & 5 & 6 & i \end{array} \tag{2.3}$$

The Hilbert polynomial of $\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]/I$ is given by the case ($\checkmark$) and the vector $\boldsymbol{d_0}$ is $(3,3)$. Thus,

$$HP_{\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]}(i,j) = \frac{1}{2}\,i^2 + 2\,i\,j + j^2 + \frac{27}{2}\,i + 11\,j - 32.$$

Note that $HP_{\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]}(i,2) \neq HF_{\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]}(i,2)$, for any value of $i \in \mathbb{Z}$, no matter how big i is.

In Prop. 2.10.4 we mentioned that the Hilbert polynomial is unique and that there is $\boldsymbol{d_0} \in \mathbb{Z}^q$ such that for $\boldsymbol{d} \geq \boldsymbol{d_0}$ the Hilbert function agrees with the Hilbert polynomial. When we write that $\boldsymbol{d_0} \geq \boldsymbol{d}$, we mean that the inequality holds coordinate-wise. Using this definition, contrary to the case of the $\mathbb{Z}$-graded modules,

- with respect to the partial order $\boldsymbol{d} \geq \boldsymbol{d_0}$ if and only if $\boldsymbol{d} - \boldsymbol{d_0} \in \mathbb{N}^q$, there is no minimal vector $\boldsymbol{d_0} \in \mathbb{Z}^q$ such that the Hilbert polynomial and function agree, and
- the area at which the Hilbert polynomial and Hilbert function agree might not be bounded.

An example of the last case appears in Ex. 2.10.6. We will study further this relation in Sec. 2.11.

Geometrically speaking, the multihomogeneous ideals correspond to multiprojective varieties.

Definition 2.10.7 (Multiprojective variety)**.** *Given an multihomogeneous ideal $I \subset \mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]$, we define the multiprojective variety as*

$$\mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(I) := \{\boldsymbol{p} \in \mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q} : (\forall \textit{ multihomogeneous } f \in I)\ f(\boldsymbol{p}) = 0\}.$$

A multiprojective variety is a subvariety of the multiprojective space $\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}$.

The dimension of a multihomogeneous variety V corresponds to the supremum of all the integers $m \geq 0$ such that there is a chain $W_0 \subset W_1 \subset \cdots \subset W_m \subseteq V$ of different irreducible multiprojective varieties contained in V. In contrast to what happened with the affine and projective varieties, it is not possible to relate this dimension with the Krull dimension of its multihomogeneous coordinate ring.

The concept of degree (Def. 2.5.20) extends to multiprojective varieties but as a vector and not as a number. We are not going to enter into details and we refer to the interested reader to [vdW78]. For our purposes, we want to know the degree of a zero-dimensional multiprojective variety to estimate the number of solutions. Generically, we can compute this number as a coefficient of a polynomial.

Definition 2.10.8 (Square multihomogeneous system)**.** *Let $N = n_1 + \cdots + n_q$. A system of equations defined by the polynomials $f_1, \ldots, f_N \in \mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$ is a square multihomogeneous system if it is defined by N multihomogeneous polynomials.*

We can generalize the Bézout bound, see Cor. 2.7.15, to the case of a square multihomogeneous system.

Proposition 2.10.9 (Multihomogeneous Bézout bound)**.** *[Sha13, Example 4.9] Consider a square multihomogeneous system $\{f_1, \ldots, f_N\}$ of multidegrees $\boldsymbol{d_1}, \ldots, \boldsymbol{d}_N \in \mathbb{N}^q$. If the system has a finite number of solutions over $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$, then their number, counted with multiplicities, see [CLO06, Sec. 4.2], is the coefficient of the monomial $\prod_i t^{n_i}$ in the polynomial*

$$\prod_{j=1}^{N} \sum_{i=1}^{q} d_{j,i}\, t_i.$$

When we consider the number of solutions of $\{f_1, \ldots, f_N\}$ without taking into account their multiplicities, this coefficient is an upper bound. We refer to this coefficient as the multihomogeneous Bézout bound *and we will write it as* $\texttt{MHB}(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N)$.

The multihomogeneous Bézout bound is tight for systems with generic coefficients, see Sec. 2.13.

Proposition 2.10.10 ([DKS13, Thm. 1.11])**.** *Let $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N$ be multidegrees in $\mathbb{N}^q$. There is an open subset with respect to the Zariski topology $O \subset \mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]_{\boldsymbol{d}_1} \times \cdots \times \mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]_{\boldsymbol{d}_N}$ such that for every square multihomogeneous system $(f_1, \ldots, f_N) \in O$, the number of its solutions, counted with multiplicities, is exactly the* multihomogeneous Bézout bound, $\texttt{MHB}(\boldsymbol{d_1}, \ldots, \boldsymbol{d_N})$.

2.11 Multigraded Castelnuovo-Mumford regularity

The results of this section come mainly from [Bot11, Ch. 6], [MS04] and [ACG05].

Our interest in local cohomology emanates from the extension of the Castelnuovo-Mumford regularity in the context of multigraded algebras. We are interested in the multigraded Castelnuovo-Mumford regularity of modules related to zero-dimensional multiprojective varieties defined by square systems.

In the multigraded case, as we observed with the multigraded Hilbert Series, the regularity it is not a point, but is a non-bounded region in $\mathbb{Z}^q$ where some property holds. Following [MS04, Bot11, BC17], the multigraded Castelnuovo-Mumford regularity will be the region of multidegrees at which the local cohomology modules vanish.

Consider the $\mathbb{Z}^q$-multigraded algebra $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$ as in Sec. 2.10. Consider $N := n_1 + \cdots + n_q$ and let $\{\boldsymbol{e_1}, \ldots, \boldsymbol{e_q}\}$ be the standard (canonical) basis of the vector space $\mathbb{K}^q$. Let B be the ideal given by the intersections of all the monomials in each block, that is

$$B := \bigcap_{i=1}^{q} \langle x_{i,0}, \ldots, x_{i,n_i} \rangle = \left\langle \prod_{i=1}^{q} x_{i,j_i} : 0 \leq j_i \leq n_i \right\rangle.$$

There is a multigraded version of Prop. 2.9.10, that relates the local cohomology and the multigraded Hilbert polynomial (see Prop. 2.10.4).

Proposition 2.11.1. *[CN08, Prop. 2.4.3] Let M be a finitely generated and $\mathbb{Z}^q$ graded $\mathbb{K}[\boldsymbol{x}_1,\dots,\boldsymbol{x}_q]$-module. For each $\boldsymbol{m}\in\mathbb{N}^q$ and $i\in\mathbb{N}$, $H^i_B(M)_{\boldsymbol{m}}$ is a finite dimensional $\mathbb{K}$-vector space, and the multigraded Hilbert function of M is the Hilbert polynomial plus the alternating sums of the dimensions of the local cohomologies. That is, for each $\boldsymbol{m}\in\mathbb{N}^q$, it holds*

$$HF_M(\boldsymbol{m}) = HP_M(\boldsymbol{m}) + \sum_i (-1)^i \dim_{\mathbb{K}}(H^i_B(M)_{\boldsymbol{m}}).$$

In the following, we introduce the notion of *multigraded Castelnuovo-Mumford regularity* as defined in [MS04, Def. 1.1] and [BC17, Def. 4.2]. Recall that given two vectors $\boldsymbol{\gamma},\boldsymbol{m}\in\mathbb{Z}^q$, we say that $\boldsymbol{\gamma}$ is component-wise bigger than $\boldsymbol{m}$, and write it as $\boldsymbol{\gamma}\geq\boldsymbol{m}$, if and only if $\boldsymbol{\gamma}-\boldsymbol{m}\in\mathbb{N}^q$.

Definition 2.11.2 (Multigraded Castelnuovo-Mumford regularity)**.** *Let M be a finitely generated $\mathbb{Z}^q$-graded $\mathbb{K}[\boldsymbol{x}_1,\dots,\boldsymbol{x}_q]$-module. Given $\boldsymbol{m}\in\mathbb{Z}^q$, we say that M is $\boldsymbol{m}$-regular if for every $\boldsymbol{\gamma}\in\mathbb{Z}^q$ such that $\boldsymbol{\gamma}\geq\boldsymbol{m}$ (component-wise), it holds*

- $(H^0_B(M))_{\boldsymbol{\gamma}+\boldsymbol{e_k}} = 0$ *for every $1\leq k\leq q$, and*
- $(H^i_B(M))_{\boldsymbol{\gamma}+\sum_{k=1}^q \lambda_k \boldsymbol{e_k}} = 0$, *for every $i\geq 1$ and $\lambda_1,\dots,\lambda_q\in\mathbb{N}$ such that $\sum_{k=1}^q\lambda_k = i-1$.*

The regularity of M is the subset of $\mathbb{Z}^q$ where M is regular, that is

$$reg_{\mathfrak{CM}}(M) := \{\boldsymbol{m}\in\mathbb{Z}^q : M \text{ is } \boldsymbol{m}\text{-regular}\}.$$

Remark 2.11.3. *Note that if M is $\boldsymbol{m}$-regular, then for every $\boldsymbol{\gamma}\geq\boldsymbol{m}$, M is $\boldsymbol{\gamma}$-regular, that is $\boldsymbol{\gamma}\in reg_{\mathfrak{CM}}(M)$*

Using Prop. 2.11.1, Cor. 2.8.4 extends to the multigraded case.

Corollary 2.11.4. *[BC17, Cor. 4.27] Let M be a finitely generated $\mathbb{Z}^q$-graded $\mathbb{K}[\boldsymbol{x}_1,\dots,\boldsymbol{x}_q]$-module such that M is $\boldsymbol{m}$-regular. Then, for $\boldsymbol{m}\in reg_{\mathfrak{CM}}(M)$, the Hilbert polynomial and the Hilbert function agree, that is*

$$HF_M(\boldsymbol{m}) = HP_M(\boldsymbol{m}).$$

Our goal is to extend the Macaulay bound, see Prop. 2.8.5, to the case of multigraded algebras. This bound give us the Castelnuovo-Mumford regularity of a quotient ring related to a regular sequence. Before introducing a multigraded version of the Macaulay bound, we need to review the regular sequences. When we work over the $\mathbb{Z}$-graded algebra $\mathbb{K}[\boldsymbol{x}]$, we can consider homogeneous regular sequences. But, when we work over the $\mathbb{Z}^q$-multigraded algebra $\mathbb{K}[\boldsymbol{x}_1,\dots,\boldsymbol{x}_q]$, there might not be multihomogeneous regular sequences.

Example 2.11.5. *[SDC07, Sec. 3.1.2] Consider two blocks of variables $\boldsymbol{x}=\{x_0,x_1\}$ and $\boldsymbol{y}=\{y_0,y_1\}$, and the $\mathbb{Z}^2$-graded algebra $\mathbb{K}[\boldsymbol{x},\boldsymbol{y}] = \bigoplus_{(i,j)\in\mathbb{Z}^2}\mathbb{K}[\boldsymbol{x}]_i\otimes\mathbb{K}[\boldsymbol{y}]_j$. Consider three polynomials $f_1,f_2,f_3\in\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]$ of bidegrees strictly bigger than zero. If f_1,f_2,f_3 is a regular sequence, by Prop. 2.7.13, then the Krull dimension of $\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]/\langle f_1,f_2,f_3\rangle$ should be $4-3=1$. But this is not possible because, as $\langle f_1,f_2,f_3\rangle\subset\langle x_0\,y_0, x_1\,y_0, x_0\,y_1, x_1\,y_1\rangle$, it holds*

$$2 = \dim_{Krull}(\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]/\langle x_0\,y_0, x_1\,y_0, x_0\,y_1, x_1\,y_1\rangle) \leq \dim_{Krull}(\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]/\langle f_1,f_2,f_3\rangle).$$

Hence, there are not three multihomogeneous polynomials $f_1,f_2,f_3\in\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]$ with bidegrees strictly bigger than zero such that (f_1,f_2,f_3) is a regular sequence.

Definition 2.11.6 (Regular sequence outside B)**.** *Consider multihomogeneous polynomials $f_1, \dots, f_r \in \mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$. We say that $(f_1, \dots, f_r)$ is a* regular sequence outside B *if for every prime ideal $\mathfrak{p} \subset \mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$, such that $\mathfrak{p} \not\subset B$, $(f_1, \dots, f_r)$ form a regular sequence over $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]_{\mathfrak{p}}$, the localization of $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$ at $\mathfrak{p}$ Def. 2.9.4.*

Such sequences are related to the filter regular sequences [Tru98, Sec. 2] and sequences of "almost" nonzero divisors [MS04, Sec. 3], [SVT06, Sec. 2].

If $(f_1, \dots, f_r)$ is a *regular sequence outside* B, then the multihomogeneous variety $\mathbb{V}_{\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}}(f_1, \dots, f_r) \subset \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$ has dimension $\sum_{i=1}^{q} n_i - r$.

Consider multidegrees $\boldsymbol{d_1}, \dots, \boldsymbol{d_r} \in \mathbb{N}^q$, for $r \leq \sum_{i=1}^{q} n_i$ such that, for every i, $\boldsymbol{d_i}$ has all its coordinates strictly positive, that is, $\boldsymbol{d_i} - (1, \dots, 1) \in \mathbb{N}^q$. A generic sequence of multihomogeneous polynomials $(f_1, \dots, f_r)$ of degrees $\boldsymbol{d_1}, \dots, \boldsymbol{d_r}$, see Sec. 2.13, is a *regular sequence outside* B, see [ACG05, Sec. 4].

When we have a *regular sequence outside* B, its Koszul homologies are B*-torsion*. This proposition follows from considering the spectral sequence of the double complex given by the Koszul complex and the Čech complex of $f_1, \dots, f_r$ over B, when $f_1, \dots, f_r$ is a regular sequence outside B, see [ACG05, Sec. 4].

Proposition 2.11.7. *Let $(f_1, \dots, f_r)$ be a multihomogeneous regular sequence outside B. Consider the Koszul complex of $\mathcal{K}(f_1, \dots, f_r)$ and let $\mathcal{H}_i(f_1, \dots, f_r)$ be its i-th homology, see Def. 2.7.5. Then, for every $j > 0$, the j-th Koszul homology is B-torsion, that is, $H^0_B(\mathcal{H}_j(f_1, \dots, f_r)) = \mathcal{H}_j(f_1, \dots, f_r)$. In particular, if $i > 0$ and $j > 0$, it holds $H^i_B(\mathcal{H}_j(f_1, \dots, f_r)) = 0$.*

Given a $\mathbb{Z}^q$-graded $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$-module M, we define $\mathtt{sp}(M) := \{\boldsymbol{d} \in \mathbb{Z}^q : (M)_{\boldsymbol{d}} \neq 0\}$ as the set of multidegrees where the module is not zero. Note that, given a finitely generated $\mathbb{Z}^q$-graded $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$-module M, the multidegrees in $\mathtt{sp}(H^i_B(M))$ are related to the multidegrees in the complement of $reg_{\mathfrak{CM}}(M)$. Hence, we will deduce bounds for $reg_{\mathfrak{CM}}(M)$ in terms of of $\mathtt{sp}(H^i_B(M))$.

Consider $\alpha \subseteq \{1, \dots, q\}$. We define the $Q_\alpha \subset \mathbb{Z}^q$ as the set of integer points in the polyhedron defined by the points $(v_1, \dots, v_q) \in \mathbb{R}^q$ such that for every $i \leq q$,

$$\begin{cases} v_i \leq -n_i - 1 & \text{if } i \in \alpha, \\ v_i \geq 0 & \text{otherwise.} \end{cases}$$

Given $v \in \mathbb{Z}^q$ and a set $Q \subset \mathbb{Z}^q$, we define $Q + v := \{w + v \in \mathbb{R}^q : w \in Q\}$. If $Q = \emptyset$, then $Q + v := \emptyset$. For $\alpha \subset \{1, \dots, q\}$, consider the map $\mathtt{in?}_\alpha : \{1, \dots, q\} \to \{0, 1\}$ such that, for all $i \in \alpha$, $\mathtt{in?}_\alpha(i) = 0$, and, for all $i \notin \alpha$, $\mathtt{in?}_\alpha(i) = 1$. Then, we can think Q_α as

$$((-1)^{\mathtt{in?}_\alpha(1)}\mathbb{N} \times \cdots \times (-1)^{\mathtt{in?}_\alpha(q)}\mathbb{N}) - (\mathtt{in?}_\alpha(1)\,(n_1 + 1), \dots, \mathtt{in?}_\alpha(q)\,(n_q + 1)).$$

Let $N_\alpha := \sum_{i \in \alpha} n_i$.

Proposition 2.11.8. *[Bot11, Lem. 6.4.4 & Lem. 6.4.7] For each $i \geq 0$, the degrees at which the i-th local cohomology of $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$ over B does not vanishes are*

$$\mathtt{sp}(H^i_B(\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q])) = \bigcup_{\substack{\alpha \subset \{1, \dots, q\} \\ N_\alpha + 1 = i \\ \alpha \neq \emptyset}} Q_\alpha.$$

Example 2.11.9. *[Bot11, Ex. 6.4.11] Consider two blocks of variables* $\boldsymbol{x} = \{x_0, x_1\}$ *and* $\boldsymbol{y} = \{y_0, y_1, y_2, y_3\}$. *Let* $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}] := \mathbb{K}[\boldsymbol{x}] \otimes_{\mathbb{K}} \mathbb{K}[\boldsymbol{y}]$ *and consider* $B := \langle x_i\, y_j : x_i \in \boldsymbol{x}, y_j \in \boldsymbol{y} \rangle$. *By Prop. 2.11.8, it holds*

- $\mathtt{sp}(H^2_B(\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}])) = Q_{\{1\}} = (-\mathbb{N} \times \mathbb{N}) + (-2, 0)$.
- $\mathtt{sp}(H^4_B(\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}])) = Q_{\{2\}} = (\mathbb{N} \times -\mathbb{N}) + (0, -4)$.
- $\mathtt{sp}(H^5_B(\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}])) = Q_{\{1,2\}} = (-\mathbb{N} \times -\mathbb{N}) + (-2, -4)$.
- $\mathtt{sp}(H^i_B(\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}])) = \emptyset$, *for* $i \notin \{2, 4, 5\}$.

The gray area in Fig. 2.1 represents the multidegrees (d_x, d_y) *at which a local cohomology of* $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$ *does not vanish. Hence, the ring* $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$ *is* $(-1, -3)$*-regular (the shaded area does not intersect with the gray ones).*

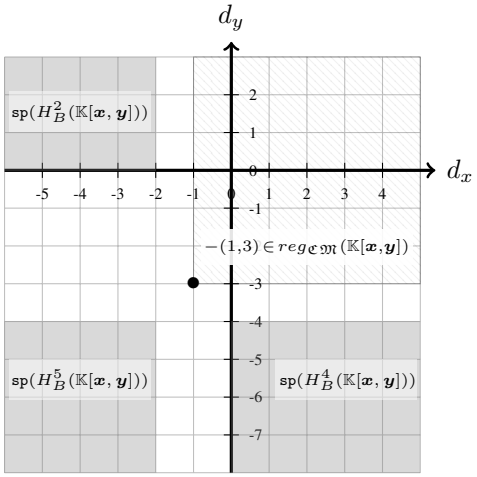


Figure 2.1: Degrees (d_x, d_y) at which a local cohomology of $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$ does not vanish. The ring $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$ is $(-1, -3)$-regular.

Given a multihomogeneous system $(f_1, \ldots, f_r)$, consider the set

$$\boldsymbol{\Sigma}(i) := \left\{ \sum_{j \in I} \deg(f_j) : I \subset \{1 \ldots r\}, \#I = i \right\} \tag{2.4}$$

containing the sums of the degrees of i (different) polynomials from the set $\{f_1, \ldots, f_r\}$.

Proposition 2.11.10 (Bounds for the multihomogeneous Castelnuovo-Mumford regularity)**.** *[MS04, Thm. 7.2] [Bot11, Rmk 6.4.10], [ACG05, Cor. 4.3] Let* $(f_1, \ldots, f_r)$ *be a multihomogeneous* regular sequence outside B. *For every* i, j, *it holds*

$$\mathtt{sp}(H^i_B(\mathcal{H}_j(f_1, \ldots, f_r))) \subset \bigcup_{t \in \mathbb{Z}} \bigcup_{v \in \boldsymbol{\Sigma}(t+j-i)} \left(\mathtt{sp}(H^t_B(\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q])) + v \right).$$

Note that $\mathbf{\Sigma}(t+j-i) \neq \emptyset \iff 0 \leq t+j-i \leq r$. *By Prop. 2.11.8, the previous equation is equivalent to*

$$\begin{aligned} \mathtt{sp}(H^i_B(\mathcal{H}_j(f_1,\ldots,f_r))) &\subset \bigcup_{t=0}^{r+j-i} \bigcup_{v\in\mathbf{\Sigma}(t+j-i)} \bigcup_{\substack{\alpha\subset\{1,\ldots,q\}\\ N_\alpha+1=t\\ \alpha\neq\emptyset}} (Q_\alpha + v) \\ &\subset \bigcup_{\substack{\alpha\subset\{1,\ldots,q\}\\ N_\alpha+1+j-i\leq r\\ \alpha\neq\emptyset}} \bigcup_{v\in\mathbf{\Sigma}(N_\alpha+1+j-i)} (Q_\alpha + v)\,. \end{aligned} \tag{2.5}$$

As $\mathcal{H}_0(f_1,\ldots,f_r) = \mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]/\langle f_1,\ldots,f_r\rangle$, when $(f_1,\ldots,f_r)$ is a multihomogeneous *regular sequence outside* B, we can compute bounds for the multigraded Castelnuovo-Mumford regularity of $\mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]/\langle f_1,\ldots,f_r\rangle$.

Example 2.11.11 (Continuation of Ex. 2.11.9)**.** *Let* $(f_1,\ldots,f_4)$ *be a multihomogeneous* regular sequence outside B *such that* $\deg(f_1) = (1,2)$ *and* $\deg(f_2) = \deg(f_3) = \deg(f_4) = (2,1)$. *We want to study when the local cohomology* $H^i_B(\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]/\langle f_1,\ldots,f_4\rangle)$ *vanishes. By Prop. 2.11.8, it holds*

$$\mathtt{sp}(H^i_B(\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]/\langle f_1,\ldots,f_4\rangle)) \subset \bigcup_{t\in\mathbb{Z}} \bigcup_{v\in\mathbf{\Sigma}(t-i)} \big(\mathtt{sp}(H^t_B(\mathbb{K}[\boldsymbol{x},\boldsymbol{y}])) + v\big)\,.$$

According to Prop. 2.11.10, see Ex. 2.11.9, it holds

$$\begin{aligned} \mathtt{sp}(H^i_B(\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]/\langle f_1,\ldots,f_4\rangle)) \subset & \bigcup_{v\in\mathbf{\Sigma}(2-i)} ((-\mathbb{N}\times\mathbb{N}) + (-2,0) + v) \cup \\ & \bigcup_{v\in\mathbf{\Sigma}(4-i)} ((\mathbb{N}\times-\mathbb{N}) + (0,-4) + v) \cup \\ & \bigcup_{v\in\mathbf{\Sigma}(5-i)} ((-\mathbb{N}\times-\mathbb{N}) + (-2,-4) + v)\,. \end{aligned}$$

Note that if $4 > s > 0$, *then* $\mathbf{\Sigma}(s) = \{(1,2) + (s-1)\,(2,1), s\,(2,1)\}$ *and* $\mathbf{\Sigma}(4) = \{(1,2) + 3\,(2,1)\}$. *Hence, for* $i = 0$,

$$\begin{aligned} \mathtt{sp}(H^0_B(\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]/\langle f_1,\ldots,f_4\rangle)) \subset\ & ((-\mathbb{N}\times\mathbb{N}) + (1,3)) \cup ((-\mathbb{N}\times\mathbb{N}) + (2,2)) \cup \\ & ((\mathbb{N}\times-\mathbb{N}) + (7,1)) \cup ((\mathbb{N}\times-\mathbb{N}) + (8,0)) \cup \\ & ((-\mathbb{N}\times-\mathbb{N}) + (7,2)) \cup ((-\mathbb{N}\times-\mathbb{N}) + (8,1))\,. \end{aligned}$$

The gray area in Fig. 2.2 shows the degrees (d_x, d_y) *at which the local cohomology of* $H^0_B(\mathbb{K}[\boldsymbol{x},\boldsymbol{y}]/\langle f_1,\ldots,f_4\rangle)$ *might not vanish.*

2.12 Semigroup algebras and toric geometry

The results of this section come mainly from [CLS11, Ch. 1], [MS05, Ch. 7] and [CLS11, Ch. 1,2,3,4].

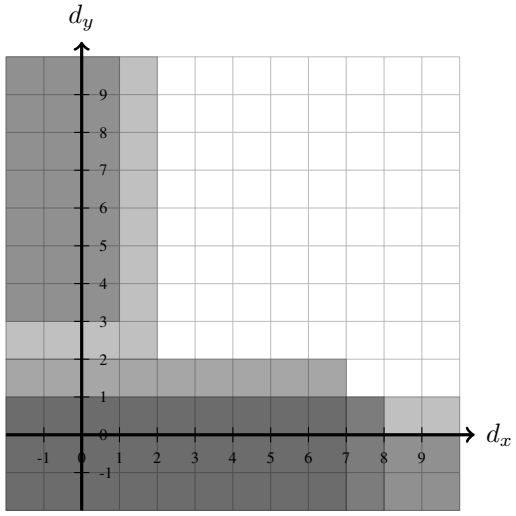


Figure 2.2: The white area represents the degrees (d_x, d_y) at which the local cohomology of $H^0_B(\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]/\langle f_1, \ldots, f_4\rangle)$ vanishes.

Definition 2.12.1 (Affine semigroup)**.** *An affine semigroup S is a finitely generated additive subsemigroup of $\mathbb{Z}^n$, for some $n \in \mathbb{N}$, such that it contains $\mathbf{0} \in \mathbb{Z}^n$.*

Given a finite set $\mathcal{A} \subset \mathbb{Z}^n$, we define the affine semigroup $\mathbb{N}\mathcal{A}$ as the subsemigroup of $\mathbb{Z}^n$ generated by $\mathcal{A}$, that is, $\mathbb{N}\mathcal{A} := \left\{\sum_{v\in A} \lambda_v v : (\forall v \in \mathcal{A})\, \lambda_v \in \mathbb{N}\right\}$.

Given a semigroup S, generated by a finite set $\mathcal{A}$, we consider the smallest subgroup of $\mathbb{Z}^n$ that contains S, that is, the subgroup $\mathbb{Z}\mathcal{A}$ generated by $\mathcal{A}$, where

$$\mathbb{Z}\mathcal{A} := \left\{ \sum_{v\in A} \lambda_v v : (\forall v \in \mathcal{A})\, \lambda_v \in \mathbb{Z} \right\}. \tag{2.6}$$

We consider special classes of affine semigroups, the pointed affine semigroups.

Definition 2.12.2 (Pointed affine semigroup)**.** *[MS05, Def 7.8] An affine semigroup S is pointed if it does not contain non-zero invertible elements, that is, for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in S \setminus \{\mathbf{0}\}$, $\boldsymbol{\alpha} + \boldsymbol{\beta} \neq \mathbf{0}$.*

In the following, when we refer to semigroups, we always mean pointed affine semigroups.

Definition 2.12.3 (Convex set and convex hull)**.** *A set $\Delta \subset \mathbb{R}^n$ is convex if every line segment connecting two elements of Δ also lies in Δ; that is, for every $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \Delta$ and $0 \leq \lambda \leq 1$ it holds $\lambda\boldsymbol{\alpha} + (1-\lambda)\boldsymbol{\beta} \in \Delta$. The convex hull of Δ is the unique minimal, with respect to inclusion, convex set that contains Δ.*

Definition 2.12.4 (Pointed rational polyhedral cones)**.** *A cone $\mathcal{C}$ is a convex subset of $\mathbb{R}^n$ such that $\mathbf{0} \in \mathcal{C}$ and for every $\boldsymbol{\alpha} \in \mathcal{C}$ and $\lambda > 0$, $\lambda\,\boldsymbol{\alpha} \in \mathcal{C}$. The* dimension of a cone *is the dimension of the vector space spanned by the cone. A cone is pointed (or strongly convex) if does not contain any line; that is, if $\mathbf{0} \neq \boldsymbol{\alpha} \in \mathcal{C}$, then $-\boldsymbol{\alpha} \notin \mathcal{C}$. A ray is a pointed cone of dimension one. A ray is rational if it contains a non-zero point of $\mathbb{Z}^n$. A rational polyhedral cone is the convex hull of a finite set of rational rays. For a*

set of points $\Delta \subset \mathbb{R}^n$, let $\mathcal{C}_\Delta$ be the cone generated by the elements in Δ. If Δ is (the convex hull of) a finite set of integer points, then $\mathcal{C}_\Delta$ is a rational polyhedral cone.

In the following, when we refer to a cone, we always mean a pointed rational polyhedral cone.

Cones and affine semigroups are related in the following way.

Proposition 2.12.5 (Gordan's Lemma)**.** *[MS05, Thm. 7.16] Let $\mathcal{C} \in \mathbb{R}^n$ be a pointed rational polyhedral cone and $G \subset \mathbb{Z}^n$ a subgroup of $\mathbb{Z}^n$. then $C \cap G$ is an affine semigroup.*

In particular, a rational polyhedral cone $\mathcal{C}$ defines the affine semigroup $\mathcal{C} \cap \mathbb{Z}^n$, which is pointed if and only if the cone is pointed.

Every affine semigroup has a unique minimal set of generators, known as its Hilbert basis.

Proposition 2.12.6 (Hilbert basis)**.** *[MS05, Prop. 7.15] Let S be a pointed affine semigroup. Then, there is a unique and finite minimal set $\mathcal{A} \subset S$ such that $\mathbb{N}\mathcal{A} = S$. We call the set $\mathcal{A}$ the Hilbert basis of S.*

Given a strongly convex rational polyhedral cone $\mathcal{C} \subset \mathbb{R}^n$, we define its Hilbert basis as the Hilbert basis of $\mathcal{C} \cap \mathbb{Z}^n$.

The main objects of this section are the semigroup algebras, which generalize the standard polynomial algebra $\mathbb{K}[\boldsymbol{x}]$.

Definition 2.12.7 (Semigroup algebra)**.** *Given an affine semigroup S (not necessarily pointed), the semigroup algebra $\mathbb{K}[S]$ is the $\mathbb{K}$-algebra generated by the monomials $\{\boldsymbol{X}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S\}$ such that $\boldsymbol{X}^{\boldsymbol{\alpha}} \cdot \boldsymbol{X}^{\boldsymbol{\beta}} = \boldsymbol{X}^{\boldsymbol{\alpha}+\boldsymbol{\beta}}$.*

Example 2.12.8 (Standard polynomial algebra)**.** *We can identify the standard polynomial algebra $\mathbb{K}[x_1, \ldots, x_n]$ with $\mathbb{K}[\mathbb{N}^n]$, where we identify the monomials $x_1^{\alpha_1} \cdots x_n^{\alpha_n} \in \mathbb{K}[x_1, \ldots, x_n]$ with $\boldsymbol{X}^{(\alpha_1,\ldots,\alpha_n)} \in \mathbb{K}[\mathbb{N}^n]$.*

Example 2.12.9 (Laurent polynomials)**.** *A Laurent polynomial is a finite $\mathbb{K}$-linear combination of monomials $\boldsymbol{X}^{\boldsymbol{\alpha}}$, where $\boldsymbol{\alpha} \in \mathbb{Z}^n$. In other words, a Laurent polynomial is a polynomial that can have monomials of negative degree; for example $2\,x^{-2} + x^3$. The Laurent polynomials form a $\mathbb{K}$-algebra, $\mathbb{K}[\mathbb{Z}^n]$, that corresponds to the semigroup algebra of $\mathbb{Z}^n$ generated by $\{x^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathbb{Z}\}$.*

Remark 2.12.10. *As S is a subsemigroup of $\mathbb{Z}^n$, the semigroup algebra $\mathbb{K}[S]$ is a subalgebra of the $\mathbb{K}$-algebra of the Laurent polynomials $\mathbb{K}[\mathbb{Z}^n]$.*

Proposition 2.12.11. *[MS05, Thm. 7.4] Given an affine semigroup S, the semigroup algebra $\mathbb{K}[S]$ is an integral domain, that is, it has no zero divisors (Def. 2.2.20).*

Definition 2.12.12 (Lattice ideal)**.** *[MS05, Thm. 7.4] Let $\mathcal{A} := \{a_1, \ldots, a_r\} \subset \mathbb{Z}^n$ and $e_1 \ldots e_r$ the canonical basis of $\mathbb{Z}^r$. Consider the group homomorphism $\rho : \mathbb{Z}^r \to \mathbb{Z}\mathcal{A}$ that sends e_i to a_i, for $1 \leq i \leq r$. We define the lattice ideal $I_\mathcal{A} \subset \mathbb{K}[\mathbb{N}^r]$ as,*

$$I_\mathcal{A} := \langle \boldsymbol{X}^\alpha - \boldsymbol{X}^\beta : \alpha, \beta \in \mathbb{N}^r \text{ and } \rho(\alpha - \beta) = 0 \rangle.$$

This ideal is prime.

We can think the semigroup algebras as quotient rings of the standard algebra $\mathbb{K}[x_1, \ldots, x_r] \cong \mathbb{K}[\mathbb{N}^r]$.

Proposition 2.12.13. *[MS05, Thm. 7.3] Consider the affine semigroup $\mathbb{N}\mathcal{A}$, where $\mathcal{A} := \{a_1, \ldots, a_r\} \subset \mathbb{Z}^n$. The semigroup algebra $\mathbb{K}[\mathbb{N}\mathcal{A}]$ is isomorphic to the quotient ring $\mathbb{K}[\mathbb{N}^r]/I_{\mathcal{A}}$.*

We consider semigroup algebras related to polytopes.

Definition 2.12.14 (Integer polytopes)**.** *A integer polytope $\Delta \subset \mathbb{R}^n$ is the convex hull of a finite set of (integer) points in $\mathbb{Z}^n$, that is,*

$$\Delta = \left\{ \sum_{i=1}^{r} \lambda_i \boldsymbol{\alpha}_i : (\forall i)\, \lambda_i \in \mathbb{R}_{\geq 0} \text{ and } \sum_i \lambda_i = 1 \right\}, \text{ for } \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r \in \mathbb{Z}^n.$$

Example 2.12.15 (Simplex)**.** *Let $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\} \subset \mathbb{R}^n$ be the standard basis of $\mathbb{R}^n$, that is, the basis where each $\boldsymbol{e}_i$ has one in the i-th position and zero elsewhere. The n-simplex $\delta \subset \mathbb{R}^n$ is the lattice polytope generated by $\{\mathbf{0}, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$.*

Definition 2.12.16 (Minkowski sum)**.** *The Minkowski sum of two integer polytopes Δ_1 and Δ_2 is*

$$\Delta_1 + \Delta_2 = \{\boldsymbol{\alpha} + \boldsymbol{\beta} : \boldsymbol{\alpha} \in \Delta_1, \boldsymbol{\beta} \in \Delta_2\}$$

For each polytope Δ and $k \in \mathbb{N}$, we denote by $k \cdot \Delta$ the dilation by k of Δ, that is, the Minkowski sum of k copies of Δ,

$$k \cdot \Delta := \overbrace{\Delta + \cdots + \Delta}^{k \text{ times}}.$$

Toric varieties relate semigroup algebras with the torus $(\mathbb{C}^*)^n$. A toric variety is an irreducible variety X that contains $(\mathbb{C}^*)^n$ as an open subset such that the action of $(\mathbb{C}^*)^n$ on itself extends to an algebraic action of $(\mathbb{C}^*)^n$ on X [CLS11, Def. 3.1.1]. Semigroup algebras correspond to the coordinate rings of the affine pieces of X.

Given an integer polytope Δ, we can define a projective complete normal irreducible toric variety X associated to it [CLS11, Sec. 2.3]. Likewise, given a polynomial system $(f_1, \ldots, f_m)$, we can define a projective toric variety X associated to the Minkowski sum of the corresponding Newton polytopes. We can homogenize these polynomials in a way that they belong to the total coordinate ring of X [CLS11, Sec. 5.4]. This homogenization is related to the facets of the polytopes.

Definition 2.12.17 (Newton polytope)**.** *Given a Laurent polynomial $f := \sum_{\boldsymbol{\alpha}} c_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\mathbb{Z}^n]$, see Ex. 2.12.9, its* support *is the set of monomials with non-zero coefficients, that is,*

$$Supp(f) := \{\boldsymbol{x}^{\boldsymbol{\alpha}} : c_{\boldsymbol{\alpha}} \neq 0\}.$$

The Newton polytope *of f is the integer polytope generated by the set of the exponents $\boldsymbol{\alpha}$ of the support of f; that is,*

$$\mathrm{NP}(f) := \textit{Convex Hull}(\{\boldsymbol{\alpha} \in \mathbb{Z}^n : \boldsymbol{x}^{\boldsymbol{\alpha}} \in Supp(f)\}).$$

When we have a zero-dimensional square system, the Bernstein-Kushnirenko-Khovanskii (BKK) theorem [Ber75, Kus76, Kho78] counts the number of solutions of the homogenization of the system

over X. This number (called BKK bound) bounds the (finite) number of solutions of a square system of sparse Laurent polynomials over the torus $(\mathbb{C}^*)^n$, where

$$\mathbb{C}^* := \mathbb{C} \setminus \{0\}. \tag{2.7}$$

The BKK bound is given in terms of the volumes of the polytopes, and more precisely, their mixed volume. It generalizes the Bézout bound (Cor. 2.7.15) and the multihomogeneous Bézout bound (Prop. 2.10.9).

Definition 2.12.18 (Mixed volume)**.** *Let $\Delta_1, \ldots, \Delta_n \in \mathbb{R}^n$ be integer polytopes. Their mixed volume, $\mathrm{MV}(\Delta_1, \ldots, \Delta_n)$, is the alternating sum of the number of integer points of the polytopes obtained by all possible Minkowski sums, that is*

$$\mathrm{MV}(\Delta_1, \ldots, \Delta_n) = (-1)^n + \sum_{k=1}^{n} (-1)^{n-k} \left(\sum_{\substack{I \subset \{1,\ldots,n\} \\ \#I = k}} \#\left((\Delta_{I_1} + \cdots + \Delta_{I_k}) \cap \mathbb{Z}^n \right) \right). \tag{2.8}$$

Theorem 2.12.19 (BKK bound)**.** *[CLO06, Thm 7.5.4] Let $f_1, \ldots, f_n$ be a system of polynomials with Newton polytopes $\Delta_1, \ldots, \Delta_n$ having a finite number of solutions over $(\mathbb{C}^*)^n$. The mixed volume $\mathrm{MV}(\Delta_1, \ldots, \Delta_n)$ upper bounds the number of solutions of the system over the torus $(\mathbb{C}^*)^n$. If the non-zero coefficients of the polynomials are generic (Sec. 2.13), then the bound is tight.*

2.13 Genericity

In this thesis, we will discuss many properties that hold in general, meaning that they fail only in degenerate situations. To formalize this notion, we use the Zariski topology.

Definition 2.13.1. *A property $P(u_1, \ldots, u_m)$ is a logic proposition involving m free variables $u_1, \ldots, u_m$.*

Definition 2.13.2 (Genericity)**.** *A property $P(u_1, \ldots, u_m)$ is true generically if there is a non-zero open subset of $\mathcal{O} \subset \mathbb{K}^m$ such that $P(c_1, \ldots, c_m)$ is true for every point $(c_1, \ldots, c_m) \in \mathcal{O}$.*

Example 2.13.3. *Consider the matrix $\left[\begin{smallmatrix} u_{1,1} & u_{1,2} \\ u_{2,1} & u_{2,2} \end{smallmatrix}\right]$. Generically, if we evaluate the matrix on $\mathbb{K}^4$ it is invertible. This holds because the matrix is singular if and only if its determinant vanishes. But the determinant is the non-zero polynomial $u_{1,1}\,u_{2,2} - u_{2,1}\,u_{1,2} \in K[u_{1,1}, u_{1,2}, u_{2,1}, u_{2,2}]$ and the points where it vanishes are $\mathbb{V}_{\mathbb{K}^n}(\langle u_{1,1}\,u_{2,2} - u_{2,1}\,u_{1,2} \rangle)$, which is a closed subset of $\mathbb{K}^4$.*

Observation 2.13.4. *Over the Zariski topology, the proper closed subsets of $\mathbb{K}^m$ have measure zero. Hence, a generic property holds for all the points of $\mathbb{K}^m$, with exception of a measure zero subset.*

Chapter 3

Resultants

Roughly speaking, the resultant of a system of $n+1$ polynomial equations in n unknowns is a polynomial in the coefficients of the input polynomials, which vanishes if and only if the system has a solution over a variety. It is also a tool to solve square systems [Stu02, CLO06]. In this chapter we will discuss two kind of resultants: the projective and the multiprojective one. Besides their theoretical aspects, we will focus on how we can efficiently compute them. Normally, we can compute resultants as a quotient of determinants of matrices [Cha93, DD01, D'A02, Bus06] and, in some cases, as the determinant of only one matrix. In the latter case, we say that we have a *determinantal formula*. Such a formula does not exist in general and it is an open problem to determine when it does. We will study the Weyman complex, which is a useful tool to derive determinantal formulas [WZ94, DE03, EM12, BMT17].

3.1 The determinant

We start this chapter by studying the resultant of a linear system, that is, a system of linear polynomials. Consider homogeneous linear polynomials (linear forms) $f_1, \dots, f_n \in \mathbb{K}[\boldsymbol{x}]_1$, where $f_i = \sum_{j=1}^{n} c_{i,j} x_j$, with $c_{i,j} \in \mathbb{K}$. The equations $f_1, \dots, f_n$ are linear independent if and only if $\mathbf{0} \in \mathbb{K}^n$ is the only solution of the system. The latter corresponds to the projective variety $\mathbb{V}_{\mathbb{P}^n}(\langle f_1, \dots, f_n \rangle) \subset \mathbb{P}^{n-1}$ being empty.

We can decide if the linear forms are linearly independent by constructing the matrix $(c_{i,j})_{1 \leq i,j \leq n}$ and by checking if its determinant is non-zero. The determinant of a matrix is a multilinear polynomial in the elements of the matrix, and in our case, it is multilinear in the coefficients of the linear forms. Hence, we can use it to parameterize all the square linear polynomial systems in $\mathbb{K}[\boldsymbol{x}]$ that have solutions different to $\mathbf{0} \in \mathbb{K}^n$.

To do so, we define the ring of parameters $\mathbb{K}[u_{i,j} : 1 \leq i, j \leq n]$ and consider a square polynomial system over this ring, say $\{F_1, \dots, F_n\} \subset \mathbb{K}[u_{i,j} : 1 \leq i, j \leq n][\boldsymbol{x}]$ where $F_i := \sum_{j=1}^{n} u_{i,j} x_j$. Let $\texttt{Det}$ be the determinant of the matrix $(u_{i,j})_{1 \leq i,j \leq n} \in \mathbb{K}[u_{i,j} : 1 \leq i, j \leq n]^{n \times n}$. $\texttt{Det}$ is a multilinear polynomial in $\mathbb{K}[u_{i,j} : 1 \leq i, j \leq n]$. Let $V \in \mathbb{K}^{n^2}$ be the affine variety associated to $\texttt{Det}$. For each $\boldsymbol{c} \in V$, we consider the morphism that specializes the coefficients of the polynomials at $\boldsymbol{c}$, that is the morphism $\mathfrak{sp}_{\boldsymbol{c}} : \mathbb{K}[u_{i,j} : 1 \leq i, j \leq n] \to \mathbb{K}$, where $\mathfrak{sp}(u_{i,j}) := c_{i,j}$. Then, we can use this morphism to describe every square linear system that vanishes over $\mathbb{P}^{n-1}$,

$$\{(\mathfrak{sp}_{\boldsymbol{c}}(F_1), \dots, \mathfrak{sp}_{\boldsymbol{c}}(F_n)) : \boldsymbol{c} \in V\} = \{(f_1, \dots, f_n) \in \mathbb{K}[\boldsymbol{x}]_1^n : \exists \boldsymbol{p} \in \mathbb{P}^n, f_1(\boldsymbol{p}) = \cdots = f_n(\boldsymbol{p}) = 0\}.$$

In conclusion, we can tell when an linear square system has a solution over $\mathbb{P}^{n-1}$ by checking when a polynomial vanishes. The resultant theory generalizes this condition beyond linear forms.

3.2 The projective resultant

The results of this sections come mainly from [Bus06, Ch. 2 & Ch. 3], [CLO06, Ch. 3], [Jou91], and [SS01, Ch. 15].

Consider the $\mathbb{Z}$-graded polynomial algebra $\mathbb{K}[\boldsymbol{x}]$. To simplify the notation, throughout this section we fix $d_1, \ldots, d_n$ to be elements in $\mathbb{Z}$.

Definition 3.2.1 (Generic system)**.** *[Bus06, Sec. 3.1.2] For fix $d_1, \ldots, d_n$, we consider the ring $\mathbb{Z}[\boldsymbol{u}] := \mathbb{Z}[u_{i,\boldsymbol{\alpha}} : 1 \leq i \leq n, \boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}]_{d_i}]$, where $\boldsymbol{u} := \{u_{i,\boldsymbol{\alpha}} : 1 \leq i \leq n, \boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}]_{d_i}\}$ is a set of fresh variables that parametrizes the coefficients of a list of n homogeneous polynomials of degrees $d_1, \ldots, d_n$, respectively.*

We define the polynomial ring $\mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}]$ as the ring of polynomials whose coefficients belong to the ring $\mathbb{Z}[\boldsymbol{u}]$, that is,

$$\mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}] := (\mathbb{Z}[\boldsymbol{u}])[x_1, \ldots, x_n].$$

If $\mathbb{K}[\boldsymbol{x}]$ is L-graded, then $\mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}]$ inherits this grading, that is, we consider $\mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}]$ as an L-graded ring, that is,

$$\mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}] = \bigoplus_{m \in L} (\mathbb{Z}[\boldsymbol{u}] \otimes_{\mathbb{Z}} \mathbb{K}[\boldsymbol{x}]_m).$$

The generic polynomial system *is the system $\boldsymbol{F} := \{F_1, \ldots, F_n\} \subset \mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}]$, where*

$$F_i := \sum_{\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}]_{d_i}} u_{i,\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}.$$

We warn the reader that the notion of generic systems and genericity (see Sec. 2.13), even though they are related, they are not the same and they should not be confused. Sometimes, generic systems are called systems of "universal" polynomials [CLO06, Sec. 3.2].

Definition 3.2.2 (Specialization morphism)**.** *The* specialization morphism *specializes the elements of $\mathbb{Z}[\boldsymbol{u}]$ to elements in $\mathbb{K}$. That is, for each $\boldsymbol{c} \in \mathbb{K}^{\#\boldsymbol{u}}$, we define $\mathtt{sp}_{\boldsymbol{c}} : \mathbb{Z}[\boldsymbol{u}] \to \mathbb{K}$ such that $\mathtt{sp}_{\boldsymbol{c}}(u_{i,\boldsymbol{\alpha}}) = c_{i,\boldsymbol{\alpha}}$.*

To simplify the notation, given an element $g \in \mathbb{Z}[\boldsymbol{u}]$ and a polynomial system $(f_1, \ldots, f_n)$ such that $(\mathtt{sp}_{\boldsymbol{c}}(F_1), \ldots, \mathtt{sp}_{\boldsymbol{c}}(F_n)) = (f_1, \ldots, f_n)$, for some $\boldsymbol{c} \in \mathbb{K}^{\#\boldsymbol{u}}$, we write $g(f_1, \ldots, f_n)$ to refer to the specialization $\mathtt{sp}_{\boldsymbol{c}}(g)$.

Proposition 3.2.3 (Projective resultant)**.** *[Bus06, Thm. 3.8] There is a unique irreducible polynomial $\mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$, called the* projective resultant*, that has the following properties:*

- $\mathbb{V}_{\mathbb{P}^{n-1}}(f_1, \ldots, f_n) \neq \emptyset$ *if and only if* $\mathtt{res}(f_1, \ldots, f_n) = 0$*, and*
- $\mathtt{res}(x_1^{d_1}, \ldots, x_n^{d_n}) = 1$.

By Cor. 2.5.19, a system of homogeneous polynomials given by $(f_1, \dots, f_r)$, with $r < n$, has always solutions over $\mathbb{P}^{n-1}$, so it makes no sense to define the resultant in this case. However, when $r > n$, we can ask ourselves when the overdetermined system has solutions. It turns out that we can define a mathematical object similar to the resultant, but it is not a single polynomial anymore. It consist of a set of polynomials, that we call the *resultant system*. We will not go in this direction, but we refer the interested reader to [vdW91, Sec. 16.5] or [Jou91]. The reason we talk about the resultant as a polynomial in the case when $n = r$ is because the resultant system spans a principal ideal, and so we can reduce it to only one polynomial, which we call the *resultant*. For a proof of this statement, see [Bus06, Sec. 3].

Proposition 3.2.4 (Degree of the resultant)**.** *Let $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ be a partition of the variables in $\mathbb{Z}[\boldsymbol{u}]$ where $\boldsymbol{u}_i := \{u_{i,\boldsymbol{\alpha}} : \boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}]_{d_i}\}$. The projective resultant $\mathtt{res}$ is a multihomogeneous polynomial with respect to this partition, see Sec. 2.10. The degree of $\mathtt{res}$ with respect to the block of variables $\boldsymbol{u}_i$ is the Bézout bound of a square system of polynomials with degrees $d_1, \dots, d_{i-1}, d_{i+1}, \dots, d_n$, that is, $\prod_{j \neq i} d_j$ (Cor. 2.7.15).*

3.2.1 Computation

In general, when we talk about computing the resultant of a system, we refer to the following two cases:

- We want to compute the polynomial $\mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$.
- Given a system $(f_1, \dots, f_n)$, we want to check if $\mathtt{res}(f_1, \dots, f_n)$ is zero.

Both questions have been addressed extensively. We will concentrate in methods that involve linear algebra. Either we will compute $\mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$ as a determinant (or a quotient of determinants) of a matrix, or we will check if $\mathtt{res}(f_1, \dots, f_n) \neq 0$, by computing the rank of a matrix. To do so, we will exploit the relation between the systems $(f_1, \dots, f_n)$ such that $\mathtt{res}(f_1, \dots, f_n) \neq 0$ and the regular sequences, see Sec. 2.7.

Proposition 3.2.5 (Macaulay resultant formula)**.** *[CLO06, Thm. 3.4.9] For each $1 \leq i \leq n$, we define the set*

$$P_i := \{\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}]_{\mathtt{MB}} \text{ s.t. } x_i^{d_i} | \boldsymbol{x}^{\boldsymbol{\alpha}} \text{ and } (\forall j < i)\ \boldsymbol{x}^{\boldsymbol{\alpha}} \notin P_j\},$$

where $\mathtt{MB}$ denotes the Macaulay bound for a system of polynomials of degrees $d_1, \dots, d_n$, that is, $\sum_{i \geq 1} d_i - n + 1$, see Prop. 2.8.5. The sets $P_1, \dots, P_n$ form a partition of the set of monomials in $\mathbb{K}[\boldsymbol{x}]_{\mathtt{MB}}$.

Consider a matrix $\mathcal{M} \in \mathbb{Z}[\boldsymbol{u}]^{\dim_{\mathbb{K}}(\mathbb{K}[\boldsymbol{x}]_{\mathtt{MB}}) \times \dim_{\mathbb{K}}(\mathbb{K}[\boldsymbol{x}]_{\mathtt{MB}})}$ whose the columns and rows are indexed by the monomials in $\mathbb{K}[\boldsymbol{x}]_{\mathtt{MB}}$. For each element in the matrix, the element in the row indexed by $\boldsymbol{x}^{\boldsymbol{\alpha}}$ and column indexed by $\boldsymbol{x}^{\boldsymbol{\beta}}$ corresponds to the coefficient of the monomial $\boldsymbol{x}^{\boldsymbol{\beta}}$ in the polynomial $\frac{\boldsymbol{x}^{\boldsymbol{\alpha}}}{x_i^{d_i}} F_i$, where $\boldsymbol{x}^{\boldsymbol{\alpha}} \in P_i$.

Let $\widetilde{\mathcal{M}}$ be a submatrix of $\mathcal{M}$ obtained from $\mathcal{M}$ by taking its rows and columns indexed by the monomials in the set

$$\{\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}]_{\mathtt{MB}} : \exists i \neq j \text{ such that } x_i^{d_i} | \boldsymbol{x}^{\boldsymbol{\alpha}} \text{ and } x_j^{d_j} | \boldsymbol{x}^{\boldsymbol{\alpha}}\}. \tag{3.1}$$

If the set described in Eq. (3.1) *is empty, we consider* $\widetilde{\mathcal{M}} = 1$.

Then, we can compute the projective resultant as a quotient of determinants as,

$$\texttt{res} = \frac{\det(\mathcal{M})}{\det(\widetilde{\mathcal{M}})}.$$

Example 3.2.6. *Consider a system of polynomials of degrees* $(d_1, d_2, d_3) = (1, 1, 2)$ *in* $\mathbb{K}[x, y, z]$. *The system of generic polynomials is as follows:*

$$\begin{cases} F_1 := u_{1,x}\, x + u_{1,y}\, y + u_{1,z}\, z, \\ F_2 := u_{2,x}\, x + u_{2,y}\, y + u_{2,z}\, z, \\ F_3 := u_{3,x^2}\, x^2 + u_{3,x\,y}\, x\, y + u_{3,x\,z}\, x\, z + u_{3,y^2}\, y^2 + u_{3,y\,z} y\, z + u_{3,z^2}\, z^2. \end{cases} \tag{3.2}$$

The Macaulay bound is $\texttt{MB} = (1+1+2)-3+1 = 2$. *To compute the resultant we consider the following partition of the set of monomials in* $\mathbb{K}[x, y, z]_2$,

$$\begin{cases} P_1 = \{x^2, x\, y, x\, z\}, \\ P_2 = \{y^2, y\, z\}, \text{ and} \\ P_3 = \{z^2\}. \end{cases} \tag{3.3}$$

Note that the only monomial in $\mathbb{K}[x, y, z]_2$ *related to* $\widetilde{\mathcal{M}}$ *is the monomial* xy *as* $x^{d_1}|x\, y$ *and* $y^{d_2}|x\, y$. *Hence, the matrices* $\mathcal{M}$ *and* $\widetilde{\mathcal{M}}$ *are as follows:*

$$\mathcal{M} = \begin{array}{c|cccccc} & x^2 & x\,y & x\,z & y^2 & y\,z & z^2 \\ \hline x^2 & u_{1,x} & u_{1,y} & u_{1,z} & & & \\ x\,y & & \boldsymbol{u_{1,x}} & & u_{1,y} & u_{1,z} & \\ x\,z & & & u_{1,x} & & u_{1,y} & u_{1,z} \\ y^2 & & u_{2,x} & & u_{2,y} & u_{2,z} & \\ y\,z & & & u_{2,x} & & u_{2,y} & u_{2,z} \\ z^2 & u_{3,x^2} & u_{3,x\,y} & u_{3,x\,z} & u_{3,y^2} & u_{3,y\,z} & u_{3,z^2} \end{array} \qquad \text{and} \qquad \widetilde{\mathcal{M}} = \begin{bmatrix} u_{1,x} \end{bmatrix}.$$

The resultant of the system is the quotient of the determinants of M *and* $\widetilde{M}$, *that is,*

$$\begin{aligned} \texttt{res} = \frac{\det(\mathcal{M})}{\det(\widetilde{\mathcal{M}})} = \frac{1}{u_{1,x}}\, u_{1,x}\, (&u_{1,x}^2\, u_{2,y}^2\, u_{3,z^2} - u_{1,x}^2\, u_{2,y}\, u_{2,z}\, u_{3,y,z} + u_{1,x}^2\, u_{2,z}^2\, u_{3,y^2} - \\ &2\, u_{1,x}\, u_{1,y}\, u_{2,x}\, u_{2,y}\, u_{3,z^2} + u_{1,x}\, u_{1,y}\, u_{2,x}\, u_{2,z}\, u_{3,y,z} + \\ &u_{1,x}\, u_{1,y}\, u_{2,y}\, u_{2,z}\, u_{3,x,z} - u_{1,x}\, u_{1,y}\, u_{2,z}^2\, u_{3,x,y} + \\ &u_{1,x}\, u_{1,z}\, u_{2,x}\, u_{2,y}\, u_{3,y,z} - 2\, u_{1,x}\, u_{1,z}\, u_{2,x}\, u_{2,z}\, u_{3,y^2} - \\ &u_{1,x}\, u_{1,z}\, u_{2,y}^2\, u_{3,x,z} + u_{1,x}\, u_{1,z}\, u_{2,y}\, u_{2,z}\, u_{3,x,y} + u_{1,y}^2\, u_{2,x}^2\, u_{3,z^2} - \\ &u_{1,y}^2\, u_{2,x}\, u_{2,z}\, u_{3,x,z} + u_{1,y}^2\, u_{2,z}^2\, u_{3,x^2} - u_{1,y}\, u_{1,z}\, u_{2,x}^2\, u_{3,y,z} + \\ &u_{1,y}\, u_{1,z}\, u_{2,x}\, u_{2,y}\, u_{3,x,z} + u_{1,y}\, u_{1,z}\, u_{2,x}\, u_{2,z}\, u_{3,x,y} - \\ &2\, u_{1,y}\, u_{1,z}\, u_{2,y}\, u_{2,z}\, u_{3,x^2} + u_{1,z}^2\, u_{2,x}^2\, u_{3,y^2} - u_{1,z}^2\, u_{2,x}\, u_{2,y}\, u_{3,x,y} + \\ &u_{1,z}^2\, u_{2,y}^2\, u_{3,x^2}). \end{aligned}$$

A classical method to compute resultants involves the study of the exactness of the Koszul complex (Def. 2.7.5). This method, that we called the Cayley method, allows us to compute resultants as the

determinant of Koszul complexes [GKZ08, Sec. 3.4.A][1]. This approach works for a general kind of resultants, which include the projective one. The next proposition relates the exactness of the Koszul complex with the projective resultant. For more details about this relation we refer the reader to [Cha93].

Proposition 3.2.7 (Resultant and the Koszul complex)**.** *[SS01, Thm. 15.4] Given n homogeneous polynomials $f_1, \dots, f_n \in \mathbb{K}[\boldsymbol{x}]$, the following three assertions are equivalent:*

- $\texttt{res}(f_1, \dots, f_n) \neq 0$*,*
- $(f_1, \dots, f_n)$ *is a regular sequence, and*
- *The Koszul complex $\mathcal{K}_\bullet(f_1, \dots, f_n)$ is exact.*

When $(f_1, \dots, f_n)$ is a regular sequence, its Castelnuovo-Mumford regularity (Def. 2.8.3) is the Macaulay bound, see Prop. 2.8.5. We can use this fact to restate the previous proposition in terms of a strand of the Koszul complex, see Def. 2.6.8. By doing so, we can decide if the resultant of a system of polynomial equations vanishes by studying the exactness of a complex of vector spaces.

Proposition 3.2.8 (Exactness of a strand of the Koszul complex)**.** *[GKZ08, Thm. 3.4.2] Given n homogeneous polynomials $f_1, \dots, f_n \in \mathbb{K}[\boldsymbol{x}]$, the following statements are equivalent:*

- $\texttt{res}(f_1, \dots, f_n) \neq 0$*, and*
- *Any strand of the Koszul complex $\mathcal{K}_\bullet(f_1, \dots, f_n)$ of degree $v \geq \texttt{MB} = \sum_i d_i - n + 1$ is exact.*

Using the previous proposition and Prop. 2.7.8, we can decide if $\texttt{res}(f_1, \dots, f_n) \neq 0$ by studying the first map of a strand of the Koszul complex.

Proposition 3.2.9. *[Mas16, Thm. 3.C] Given a system of homogeneous polynomials $f_1, \dots, f_n$, its resultant does not vanish, $\texttt{res}(f_1, \dots, f_n) \neq 0$, if and only if the first map of the strand of the Koszul complex of degree $\texttt{MB}$ (Prop. 2.8.5) is surjective, that is, if and only if the following map is surjective,*

$$\begin{array}{rccl} \delta_1 : & \bigoplus_{i=1}^n \mathbb{K}[\boldsymbol{x}]_{\texttt{MB}-d_i} & \to & \mathbb{K}[\boldsymbol{x}]_{\texttt{MB}} \\ & (g_1, \dots, g_n) & \mapsto & \delta_1(g_1, \dots, g_n) := \sum_i g_i f_i \end{array}$$

As a corollary of the previous proposition, we reduce the question of the vanishing of the resultant to the problem of computing the rank of a matrix.

Corollary 3.2.10. *With the same notation as in Prop. 3.2.9, let $\mathcal{M}$ be the matrix representing the map δ_1. The matrix $\mathcal{M}$ is full-rank if and only if $\texttt{res}(f_1, \dots, f_n) \neq 0$.*

Remark 3.2.11 (Sylvester map)**.** *We call the maps of the form $(g_1, \dots, g_k) \mapsto \sum_i g_i f_i$, as δ_1 in Prop. 3.2.9,* Sylvester maps*.*

Following Prop. 3.2.9, we can compute the resultant (which is a polynomial $\texttt{res} \in \mathbb{Z}[\boldsymbol{u}]$) as a factor of the determinant of a matrix with elements in $\mathbb{Z}[\boldsymbol{u}]$.

[1] The determinant of a complex is a generalization of the determinant of a matrix. The determinant of a complex vanishes if and only if the complex is not exact. For an introduction to the subject see [GKZ08, App. A].

Corollary 3.2.12. *Consider the map δ_1 to be the first map of the strand of the Koszul complex of the generic systems $(F_1, \ldots, F_n)$ at degree* $\mathtt{MB}$*, that is*

$$\begin{array}{rccl} \delta_1: & \bigoplus_{i=1}^n \mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}]_{\mathtt{MB}-d_i} & \to & \mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}]_{\mathtt{MB}} \\ & (g_1, \ldots, g_n) & \mapsto & \delta_1(g_1, \ldots, g_n) := \sum_i g_i\, f_i \end{array}$$

Let W be a free $\mathbb{Z}[\boldsymbol{u}]$-submodule of $\bigoplus_{i=1}^n \mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}]_{\mathtt{MB}-d_i}$ of rank equals $\mathtt{Rank}(\mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}]_{\mathtt{MB}}) = \binom{\mathtt{MB}+(n-1)}{n-1}$ such that $\delta_1(W) = \mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}]_{\mathtt{MB}}$. Consider bases for the free $\mathbb{Z}[\boldsymbol{u}]$-modules W and $\mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}]_{\mathtt{MB}}$ and consider the matrix $\mathcal{M} \in \mathbb{Z}[\boldsymbol{u}]^{\binom{\mathtt{MB}+(n-1)}{n-1}\times\binom{\mathtt{MB}+(n-1)}{n-1}}$ representing the restriction of δ_1 to W. Then, the resultant $\mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$ is a factor of $\det(\mathcal{M})$, that is, there is a polynomial $E \in \mathbb{Z}[\boldsymbol{u}]$, that we call the extraneous factor*, such that,*

$$\det(\mathcal{M}) = E \cdot \mathtt{res}.$$

Example 3.2.13. *The Macaulay resultant formula (Prop. 3.2.5) is a particular case of Cor. 3.2.12 where the extraneous factor is the determinant of a submatrix of $\mathcal{M}$.*

For some particular systems, we can compute its resultant as the determinant of a matrix, that is, we have no extraneous factors.

Definition 3.2.14 (Determinantal formula)**.** *We say that a matrix $\mathcal{M} \in \mathbb{Z}[\boldsymbol{u}]^{t\times t}$ is a determinantal formula for* $\mathtt{res}$ *if* $\mathtt{res} = t \cdot \det(\mathcal{M})$*, for $t \in \mathbb{Z} \setminus \{0\}$. The degree of the formula is the maximal degree of the elements in $\mathcal{M}$, as polynomials in $\mathbb{Z}[\boldsymbol{u}]$.*

Example 3.2.15. *The determinant of a matrix is a determinantal formula for the resultant of the system of linear forms defined by its columns, see Sec. 3.1.*

Example 3.2.16 (Sylvester matrix)**.** *Consider the* Macaulay resultant matrix *(Prop. 3.2.5) for a system of two generic homogeneous equations $f, g \in \mathbb{K}[f_0, f_1, \ldots, f_{d_f}, g_0, \ldots, g_{d_g}][x, y]$ of degrees d_f and d_g, where $f = \sum_{i=0}^{d_f} f_i x^i y^{d_f - i}$ and $g = \sum_{i=0}^{d_g} g_i x^i y^{d_g - i}$. To construct the Macaulay resultant matrix $\mathcal{M}$ we partition the set of monomials of degree $d_f + d_g - 1$ in two sets, P_x and P_y. These are*

$$\left\{ \begin{array}{rcll} P_x & = & \{x^{d_1} \cdot \boldsymbol{x}^\alpha : \boldsymbol{x}^\alpha \in \mathbb{K}[x, y]_{d_2 - 1}\} & \text{and} \\ P_y & = & \{x^{d_2} \cdot \boldsymbol{x}^\alpha : \boldsymbol{x}^\alpha \in \mathbb{K}[x, y]_{d_1 - 1}\}. & \end{array} \right.$$

The matrix $\mathcal{M}$ has row/column dimension $d_f + d_g$. For $d_f = 4$ and $d_g = 3$, it is

$$\mathcal{M} = \begin{array}{c|ccccccc} & x^6 & x^5\,y & x^4\,y^2 & x^3\,y^3 & x^2\,y^4 & x\,y^5 & y^6 \\ \hline x^2\,f & f_4 & f_3 & f_2 & f_1 & f_0 & & \\ x\,y\,f & & f_4 & f_3 & f_2 & f_1 & f_0 & \\ y^2\,f & & & f_4 & f_3 & f_2 & f_1 & f_0 \\ x^3\,g & g_3 & g_2 & g_1 & g_0 & & & \\ x^2\,y\,g & & g_3 & g_2 & g_1 & g_0 & & \\ x\,y^2\,g & & & g_3 & g_2 & g_1 & g_0 & \\ y^3\,g & & & & g_3 & g_2 & g_1 & g_0 \end{array}$$

Note that x^{d_1} does not divide any monomial in P_2 and y^{d_1} does not divide any monomial in P_2. Hence, the matrix $\widetilde{\mathcal{M}}$ equals 1 and so, the resultant of the system is the determinant of the matrix $\mathcal{M}$, that is, $\mathcal{M}$ is a determinantal formula for `res`*,*

$$\det(\mathcal{M}) = \texttt{res}.$$

This determinantal formula *is given by a* Sylvester map*. We call* Sylvester formulas *the formulas due to Sylvester maps. The degree of the Sylvester formulas is one, as the elements of the associated matrices are the coefficients of the polynomials.*

There other ways for computing the projective resultant that we do not cover in this chapter, for example, the ones involving *Morley forms* [Jou91], *Bézoutians* [EM98] and *Dixon matrices* [KSY94]. We refer to [EM99b] for a survey on these methods. We refer the reader to [SZ94, WZ94, DD01] for some works on the existence of determinantal formulas. In particular, [DD01] also introduces a complete framework to study different ways of computing the resultant, which generalizes most of the known methods.

3.3 The multiprojective resultant

The results of this sections come mainly from [Rém01, DKS13, DS15].

We can extend the definition of projective resultant to the multiprojective settings, see Sec. 2.10. The *multiprojective resultant* is a polynomial in the coefficients of a multihomogeneous polynomial system that vanishes if and only if the system has a solution over the multiprojective space. Contrary to the projective case, the multiprojective resultant is not defined in a unique, or rather uniform, way. Its definition depends on whether we ask the resultant to be an irreducible polynomial [GKZ08] or if we ask its degree to be related to the number of solutions of the system [Rém01, DKS13, DS15]. In many cases, these definitions are the same [DKS13, Rmk. 1.39]. In this section, we follow the second definition.

Following the notation of Sec. 2.10, we consider the $\mathbb{Z}^q$-graded polynomial algebra $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$. We identify the monomials of $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$ with vectors in $\mathbb{Z}^{n_1+1} \times \cdots \times \mathbb{Z}^{n_q+1}$ and, for each $\boldsymbol{\alpha} = (\boldsymbol{\alpha_1}, \dots, \boldsymbol{\alpha_q}) \in \mathbb{Z}^{n_1+1} \times \cdots \times \mathbb{Z}^{n_q+1}$, we set

$$\boldsymbol{x}^{\boldsymbol{\alpha}} := \prod_{i=1}^{q} {\boldsymbol{x_i}}^{\boldsymbol{\alpha_i}}.$$

For each multidegree $\boldsymbol{d} \in \mathbb{Z}^q$, we denote by $\mathcal{A}(\boldsymbol{d})$ the set of exponents of the monomials of multidegree $\boldsymbol{d}$, that is

$$\mathcal{A}(\boldsymbol{d}) = \{\boldsymbol{\alpha} \in \mathbb{Z}^{n_1+1} \times \cdots \times \mathbb{Z}^{n_q+1} : \boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]_{\boldsymbol{d}}\}.$$

The cardinality of $\mathcal{A}(\boldsymbol{d})$ is

$$\#\mathcal{A}(\boldsymbol{d}) = \prod_{i=1}^{q} \binom{d_i + n_i}{n_i}.$$

We fix $N = n_1 + \cdots + n_q$. In this section, we will work over the multiprojective space $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$. To simplify the notation, throughout this section we fix $\boldsymbol{d}_0, \dots, \boldsymbol{d}_N$ to be elements in $\mathbb{Z}^q$. Following Def. 3.2.1, we introduce generic multihomogeneous polynomials systems.

Definition 3.3.1 (Generic multihomogeneous system)**.** *Consider the set of variables* $\boldsymbol{u} := \{u_{k,\boldsymbol{\alpha}} : 0 \leq k \leq N \text{ and } \boldsymbol{\alpha} \in \mathcal{A}(\boldsymbol{d}_k)\}$ *and the ring* $\mathbb{Z}[\boldsymbol{u}]$*. A generic multihomogeneous polynomial system is the system* $\boldsymbol{F} := \{F_0, \ldots, F_N\} \subset \mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$*, where*

$$F_k := \sum\nolimits_{\boldsymbol{\alpha} \in \mathcal{A}(\boldsymbol{d}_k)} u_{k,\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}. \tag{3.4}$$

The generic multihomogeneous system $\boldsymbol{F}$ parameterizes every overdetermined multihomogeneous system with polynomials of multidegrees $\boldsymbol{d}_0, \boldsymbol{d}_1, \ldots, \boldsymbol{d}_N$, respectively. Following Def. 3.2.2, for each $\boldsymbol{c} = (c_{k,\boldsymbol{\alpha}})_{0 \leq k \leq N, \boldsymbol{\alpha} \in \mathcal{A}(\boldsymbol{d}_k)} \in \mathbb{P}^{\#\mathcal{A}(\boldsymbol{d}_0)-1} \times \cdots \times \mathbb{P}^{\#\mathcal{A}(\boldsymbol{d}_N)-1}$, the specialization of $\boldsymbol{F}$ at $\boldsymbol{c}$, that we write as $\boldsymbol{F}(\boldsymbol{c})$, is a multihomogeneous polynomial system in $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$, say $(f_0, \ldots, f_N)$, where

$$f_k := F_k(\boldsymbol{c}) = \sum\nolimits_{\boldsymbol{\alpha} \in \mathcal{A}(\boldsymbol{d}_k)} c_{k,\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}. \tag{3.5}$$

Let Ω be the algebraic variety containing the overdetermined multihomogeneous systems that have solutions over $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$ and their solutions. We call Ω the *incidence variety*.

$$\Omega = \left\{ (\boldsymbol{p}, \boldsymbol{c}) \in \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q} \times \mathbb{P}^{\#\mathcal{A}(\boldsymbol{d}_0)-1} \times \cdots \times \mathbb{P}^{\#\mathcal{A}(\boldsymbol{d}_N)-1} : (\forall k \in [N])\, F_k(\boldsymbol{c})(\boldsymbol{p}) = 0 \right\}.$$

Let π be the projection map,

$$\begin{array}{rccl} \pi : & \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q} \times \mathbb{P}^{\#\mathcal{A}(\boldsymbol{d}_0)-1} \times \cdots \times \mathbb{P}^{\#\mathcal{A}(\boldsymbol{d}_N)-1} & \to & \mathbb{P}^{\#\mathcal{A}(\boldsymbol{d}_0)-1} \times \cdots \times \mathbb{P}^{\#\mathcal{A}(\boldsymbol{d}_N)-1} \\ & (\boldsymbol{p}, \boldsymbol{c}) & \mapsto & \pi(\boldsymbol{p}, \boldsymbol{c}) = \boldsymbol{c} \end{array}. \tag{3.6}$$

We can think $\pi(\Omega)$ as the set of overdetermined multihomogeneous polynomial systems with solutions over $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$. This set is an irreducible hypersurface [GKZ08, Prop. 3.3.1]. Its defining ideal in $\mathbb{Z}[\boldsymbol{u}]$ is principal and it is generated by an irreducible polynomial $\texttt{elim} \in \mathbb{Z}[\boldsymbol{u}]$ [GKZ08, Prop 8.1.1]. In particular,

$$\text{The system } \boldsymbol{F}(\boldsymbol{c}) \text{ has a solution over } \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q} \iff \boldsymbol{c} \in \pi(\Omega) \iff \texttt{elim}(\boldsymbol{c}) = 0.$$

Following [Rém01, DKS13, DS15], we call $\texttt{elim} \in \mathbb{Z}[\boldsymbol{u}]$ the *eliminant*. We warn the reader that in [GKZ08], the polynomial $\texttt{elim} \in \mathbb{Z}[\boldsymbol{u}]$ is called the *resultant*. We reserve the word resultant for a power of elim. More precisely, the resultant res is a polynomial in $\mathbb{Z}[\boldsymbol{u}]$ such that

$$\texttt{res} = \pm \texttt{elim}^{\mathcal{D}},$$

where $\mathcal{D}$ is the degree of the restriction of π to the incidence variety Ω, see [DS15, Def. 3.1]. Consequently,

$$\text{The system } \boldsymbol{F}(\boldsymbol{c}) \text{ has a solution over } \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q} \iff \texttt{res}(\boldsymbol{c}) = 0. \tag{3.7}$$

Proposition 3.3.2 ([Rém01, Prop. 3.4])**.** *Let* $\boldsymbol{u}_k$ *be the blocks of variables in* $\boldsymbol{u}$ *related to the polynomial* F_k*, that is*

$$\boldsymbol{u}_k = \{\boldsymbol{u}_{k,\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{A}(\boldsymbol{d}_i)}.$$

The resultant $\texttt{res} \in \mathbb{Z}[\boldsymbol{u}]$ *is a multihomogeneous polynomial with respect to the blocks of variables* $\boldsymbol{u}_0, \dots, \boldsymbol{u}_N$*. The degree of* res *with respect to the variables* $\boldsymbol{u}_k$ *is the multihomogeneous Bézout bound (Prop. 2.10.9) of a square system with multidegrees* $\boldsymbol{d}_0, \dots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \dots, \boldsymbol{d}_N$*, that is*

$$\mathit{degree}(\texttt{res}, \boldsymbol{u}_k) = \texttt{MHB}(\boldsymbol{d}_0, \dots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \dots, \boldsymbol{d}_N).$$

The total degree of the resultant is

$$\mathit{degree}(\texttt{res}) = \sum_{k=0}^{N} \mathit{degree}(\texttt{res}, \boldsymbol{u}_k) = \sum_{k=0}^{N} \texttt{MHB}(\boldsymbol{d}_0, \dots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \dots, \boldsymbol{d}_N).$$

3.3.1 Computation

We can generalize many of the ideas from Sec. 3.2.1 to the multihomogeneous case. Nevertheless, in this section, we follow a different approach and introduce the Weyman complex, that also allows us to construct determinantal formulas (Def. 3.2.14) for the multiprojective resultant. The results of this section come mainly from [WZ94, Wey03].

The Weyman complex [Wey94, WZ94, Wey03] of an overdetermined multihomogeneous system $\boldsymbol{f} = (f_0, \dots, f_N)$ in $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$ is a bounded complex (see Sec. 2.6) that is exact if and only if the system $\boldsymbol{f}$ has no solutions over $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$. More precisely, the determinant of the Weyman complex of the multihomogeneous generic system $\boldsymbol{F}$ (see Def. 3.3.1) is well-defined and it is equal to the multihomogeneous resultant [Wey03, Prop. 9.1.3]. If the Weyman complex involves only two non-zero vector spaces, then the resultant of $\boldsymbol{F}$ is the determinant of the map between these spaces. Thus, in this case, there is a determinantal formula for the resultant.

Definition 3.3.3 (Weyman complex)**.** *Let* $\boldsymbol{F} = (F_0, \dots, F_N)$ *in* $\mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$ *be a generic multihomogeneous system having multidegrees* $\boldsymbol{d}_0, \dots, \boldsymbol{d}_N$*, respectively (see Def. 3.3.1). Given a* degree vector $\boldsymbol{m} \in \mathbb{Z}^q$*, the Weyman complex,* $K_\bullet(\boldsymbol{m})$*, of* $\boldsymbol{F}$ *is*

$$K_\bullet(\boldsymbol{m}) : 0 \to K_{N+1}(\boldsymbol{m}) \xrightarrow{\delta_{N+1}(\boldsymbol{m})} \cdots \to K_1(\boldsymbol{m}) \xrightarrow{\delta_1(\boldsymbol{m})} K_0(\boldsymbol{m}) \xrightarrow{\delta_0(\boldsymbol{m})} \cdots \to K_{-N}(\boldsymbol{m}) \to 0.$$

For each $v \in \{-N, \dots, N+1\}$ *each the* $\mathbb{Z}[\boldsymbol{u}]$*-module* $K_v(\boldsymbol{m})$ *is*

$$K_v(\boldsymbol{m}) := \bigoplus_{p=0}^{N+1} K_{v,p}(\boldsymbol{m}) \otimes_{\mathbb{Z}} \mathbb{Z}[\boldsymbol{u}], \text{ where } K_{v,p}(\boldsymbol{m}) := \bigoplus_{\substack{I \subset \{0,\dots,N\} \\ \#I = p}} H^{p-v}_{\mathcal{P}}(\boldsymbol{m} - \sum_{k \in I} \boldsymbol{d}_k) \otimes \bigwedge_{k \in I} e_k. \tag{3.8}$$

The term $H^{p-v}_{\mathcal{P}}(\boldsymbol{m} - \sum_{k\in I} \boldsymbol{d}_k)$ *is the* $(p-v)$*-th cohomology of* $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$ *with coefficients in the sheaf* $\mathcal{O}(\boldsymbol{m} - \sum_{k\in I} \boldsymbol{d}_k)$ *whose global sections are* $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]_{\boldsymbol{m} - \sum_{k\in I} \boldsymbol{d}_k}$ *(see [Har77, Sec. II.5]). The element* $\bigwedge_{k\in I} e_k$ *is the singleton* $\{e_{I_1} \wedge \cdots \wedge e_{I_p}\}$*, where* $I_1 < \cdots < I_p$ *are the elements of* I*,* $e_0, \dots, e_N$ *is the standard basis of* $\mathbb{K}^{N+1}$*, and* $\wedge$ *is the wedge (exterior) product (see Def. 2.2.33).*

For a multihomogeneous system $\boldsymbol{f} = (f_0, \dots, f_N)$ *in* $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$ *that is the specialization of* $\boldsymbol{F}$ *at* $\boldsymbol{c}$*, see Eq. (3.5), the Weyman complex* $K_\bullet(\boldsymbol{m}; \boldsymbol{f})$ *is the Weyman complex* $K_\bullet(\boldsymbol{m})$ *where we specialize each variable* $u_{k,\boldsymbol{\alpha}}$ *at* $c_{k,\boldsymbol{\alpha}} \in \mathbb{K}$*.*

Proposition 3.3.4. *[WZ94, Prop. 2.1] The vector spaces $K_v(\boldsymbol{m}, \boldsymbol{f})$ are independent of the specialization of the variables $\boldsymbol{u}$, in particular*

$$K_v(\boldsymbol{m}, \boldsymbol{f}) = \bigoplus_{p=0}^{N+1} K_{v,p}(\boldsymbol{m}).$$

Hence, the rank of $K_v(\boldsymbol{m})$ as a $\mathbb{Z}[\boldsymbol{u}]$-module equals the dimension of $K_v(\boldsymbol{m}, \boldsymbol{f})$ as a $\mathbb{K}$-vector space. The differentials $\delta_v(\boldsymbol{m}, \boldsymbol{f})$ depend on the coefficients of $\boldsymbol{f}$.

In this thesis, we will not introduce sheaf cohomology and instead we follow [WZ94] to rewrite the previous cohomologies in terms of polynomials. Nevertheless, it worths to mention that the sheaf cohomology is "naturally" isomorphic to the local cohomology (see Sec. 2.9) in the context of multiprojective varieties [Har77, Thm. III.4.5]. For a complete introduction to sheaf cohomology we refer the reader to [Har77, Ch. III].

Before simplifying the cohomologies in Eq. (3.8), we need to introduce some extra notation. For this, we exploit the relation between symmetric algebras and polynomials.

For each $m \in \mathbb{N}$, let $(\mathbb{K}^m)^*$ be the dual of the vector space $\mathbb{K}^m$. For each $1 \leq i \leq q$, we consider new sets of $n_i + 1$ variables

$$\boldsymbol{\partial x}_i := \{\partial x_{i,0}, \ldots, \partial x_{i,n_i}\}.$$

As we did in Ex. 2.2.36, where we identified the polynomial algebra $\mathbb{K}[\boldsymbol{x}_i]$ with the symmetric algebra of the vector space $\mathbb{K}^{n_i+1}$, we will identify the algebra $\mathbb{K}[\boldsymbol{\partial x}_i]$ with the symmetric algebra $S(\mathbb{K}^{n_i+1})^*$. In the following, we will write the $\mathbb{Z}$-graduated algebras $\mathbb{K}[\boldsymbol{x}_i]$ and $\mathbb{K}[\boldsymbol{\partial x}_i]$ as

$$\mathbb{K}[\boldsymbol{x}_i] \cong S\left(\mathbb{K}^{n_i+1}\right) = \bigoplus_{d\in\mathbb{Z}} S_i(d) \qquad \text{and} \qquad \mathbb{K}[\boldsymbol{\partial x}_i] \cong S\left((\mathbb{K}^{n_i+1})^*\right) = \bigoplus_{d\in\mathbb{Z}} S_i^*(d).$$

For each i, $S_i(d)$ corresponds to the $\mathbb{K}$-vector space of polynomials in $\mathbb{K}[\boldsymbol{x}_i]$ of degree d and $S_i^*(-d)$ to the $\mathbb{K}$-vector space of polynomials in $\mathbb{K}[\boldsymbol{\partial x}_i]$ of degree d. Note that if $d < 0$, then $S_i(d) = S_i^*(-d) = 0$.

As we did with the monomials in $\mathbb{K}[\boldsymbol{x}_i]$, we identify the monomial of $\mathbb{K}[\boldsymbol{\partial x}_i]$ with vectors in $\mathbb{Z}^{n_i+1}$. For each $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_{n_i}) \in \mathbb{Z}^{n_i+1}$ we set

$$\boldsymbol{\partial x}_i{}^{\boldsymbol{\alpha}} := \prod_{j=0}^{n_i} \partial x_{i,j}^{\alpha_j}.$$

When $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$ is a product of projective spaces, using Künneth's formula (Prop. 3.3.5), we can write the cohomologies in Eq. (3.8) as a product of cohomologies of projective spaces, that in turn we can identify with the symmetric algebras of $S(\mathbb{K}^{n_k+1})$ and $S((\mathbb{K}^{n_k+1})^*)$ for $1 \leq k \leq q$.

Proposition 3.3.5 (Künneth Formula)**.** *The cohomologies of the product of projective spaces in each $K_{v,p}(\boldsymbol{m})$ of Eq.* (3.8) *are the direct sum of the tensor product of the cohomologies of each of the q projective spaces related to each block of variables, that is*

$$H_{\mathcal{P}}^{p-v}\Big(\boldsymbol{m} - \sum_{k\in I} \boldsymbol{d}_k\Big) = \bigoplus_{r_1+\cdots+r_q=p-v} \bigotimes_{i=1}^{q} H_{\mathbb{P}^{n_i}}^{r_i}\Big(m_i - \sum_{k\in I} d_{k,i}\Big). \tag{3.9}$$

By combining Bott formula and Serre's duality, see [OSS80, Sec. 1.1], we can identify the cohomologies of the previous proposition with the rings $\mathbb{K}[\boldsymbol{x}_i]$ and $\mathbb{K}[\boldsymbol{\partial x}_i]$. Moreover, for each $p - v$, there is at most one set of values for $(r_1, \ldots, r_q)$ such that the right hand side of the previous equation does not vanish.

Proposition 3.3.6. *For each $1 \leq i \leq q$, $a \in \mathbb{Z}$, it holds*

- $H^0_{\mathbb{P}^{n_i}}(a) \cong S_i(a)$, *that is the $\mathbb{K}$-vector space of the polynomials of degree a in the polynomial algebra $\mathbb{K}[\boldsymbol{x}_i]$.*
- $H^{n_i}_{\mathbb{P}^{n_i}}(a) \cong S_i^*(a + n_i + 1)$, *that is the $\mathbb{K}$-vector space of the polynomials of degree $a + n_i + 1$ in the polynomial algebra $\mathbb{K}[\boldsymbol{\partial x}_i]$.*
- *If $r_i \notin \{0, n_i\}$, then $H^{r_i}_{\mathbb{P}^{n_i}}(a) \cong 0$.*

Corollary 3.3.7. *For each $1 \leq i \leq q$, if $H^{r_i}_{\mathbb{P}^{n_i}}(a) \neq 0$, then $r_i \in \{0, n_i\}$. Moreover,*

- *If $a > -n_i - 1$, then*
$$H^{r_i}_{\mathbb{P}^{n_i}}(a) \neq 0 \iff r_i = 0 \text{ and } a \geq 0.$$
- *If $a < 0$, then*
$$H^{r_i}_{\mathbb{P}^{n_i}}(a) \neq 0 \iff r_i = n_i \text{ and } a \leq -n_i - 1.$$

We obtain the dual complex of a complex by dualizing the vector spaces and the maps. The dual of the Weyman complex, is again a Weyman complex. By exploiting Serre's duality, we can construct the degree vectors of a dual Weyman complex from the degree vector of the primal.

Proposition 3.3.8. *[Wey03, Thm. 5.1.4] Let $\boldsymbol{m}$ and $\bar{\boldsymbol{m}}$ be any degree vectors such that $\boldsymbol{m} + \bar{\boldsymbol{m}} = \sum_i \boldsymbol{d}_i - (n_1 + 1, \ldots, n_q + 1)$. Then, $K_v(\boldsymbol{m}) \cong K_{1-v}(\bar{\boldsymbol{m}})^*$ for all $v \in \mathbb{Z}$ and $K_\bullet(\boldsymbol{m})$ is dual to $K_\bullet(\bar{\boldsymbol{m}})$.*

The maps $\delta_v(\boldsymbol{m})$ between the modules of the Weyman complex are complicated to describe, and so we will only present them in a particular (alas important) case. Our goal is to obtain determinantal formulas given by matrices whose elements are linear forms in the coefficients of the input polynomials, that is linear in $\boldsymbol{u}$, see Eq. (3.4). To exploit the Weyman complex for this task, by [Wey03, Prop. 5.2.4], we have to choose a degree vector $\boldsymbol{m}$ so that the Weyman complex reduces to

$$K_\bullet(\boldsymbol{m}) : 0 \to K_{v,p+v}(\boldsymbol{m}) \otimes \mathbb{K}[\boldsymbol{u}] \xrightarrow{\delta_v(\boldsymbol{m})} K_{v-1,p+v-1}(\boldsymbol{m}) \otimes \mathbb{K}[\boldsymbol{u}] \to 0\,, \tag{3.10}$$

where $p = \sum_{k \in I} n_k$, for some set $I \subset \{1, \ldots, q\}$. That is, for all $t \notin \{v - 1, v\}$, it holds $K_t(\boldsymbol{m}) = 0$, $K_v(\boldsymbol{m}) = K_{v,p+v}(\boldsymbol{m})$ and $K_{v-1}(\boldsymbol{m}) = K_{v-1,p+v-1}(\boldsymbol{m})$.

We will describe the map $\delta_v(\boldsymbol{m})$ through an auxiliary map μ that acts like multiplication. For this, we need to introduce some additional notation. Let R be a ring; for example $R = \mathbb{Z}[\boldsymbol{u}]$ or $R = \mathbb{K}$. For each $1 \leq i \leq q$, the polynomial ring $R[\boldsymbol{x}_i]$, respectively $R[\boldsymbol{\partial x}_i]$, is a free R-module with basis $\{\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{A}(d), d \in \mathbb{Z}\}$, respectively $\{\boldsymbol{\partial x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{A}(d), d \in \mathbb{Z}\}$. We define the bilinear map

$$\mu_{(i)} : R[\boldsymbol{x}_i] \times (R[\boldsymbol{x}_i] \oplus R[\boldsymbol{\partial x}_i]) \to R[\boldsymbol{x}_i] \oplus R[\boldsymbol{\partial x}_i], \tag{3.11}$$

which acts as follows: for each $d_1, d_2 \in \mathbb{Z}$, $\boldsymbol{\alpha} \in \mathcal{A}(d_1)$ and $\boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathcal{A}(d_2)$,we have

$$\mu_{(i)}(\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\gamma}}) = \boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\gamma}} \text{ and } \quad \mu_{(i)}(\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{\partial}\boldsymbol{x}^{\boldsymbol{\beta}}) = \begin{cases} \boldsymbol{\partial}\boldsymbol{x}^{\boldsymbol{\beta}-\boldsymbol{\alpha}} & \text{if } d_\alpha \leq d_\beta \text{ and } \boldsymbol{\beta}-\boldsymbol{\alpha} \in \mathcal{A}(d_1-d_2) \\ 0 & \text{otherwise} \end{cases}.$$

The map $\mu_{(i)}$ is graded in the following way, for $f \in S_i(d)$ it holds

$$\mu_{(i)}(f, S_i(D)) \subseteq S_i(D+d) \quad \text{and} \quad \mu_{(i)}(f, S_i^*(D)) \subseteq S_i^*(D+d).$$

Remark 3.3.9. *If we restrict the domain of $\mu_{(i)}$ to $R[\boldsymbol{x}_i] \times R[\boldsymbol{x}_i]$, then $\mu_{(i)}$ acts as multiplication, that is for every $f, g \in R[\boldsymbol{x}_i]$ it holds*

$$\mu_{(i)}(f, g) = f\, g.$$

Example 3.3.10. *Consider $x_0^2 + x_0\, x_1 \in \mathbb{K}[x_0, x_1]_2$ and $2\,\partial x_0^2\,\partial x_1 - \partial x_0\,\partial x_1^2 \in \mathbb{K}[\partial x_0, \partial x_1]_{-3}$. We want to compute $\mu_{(x)}(x_0^2 + x_0\, x_1, 2\,\partial x_0^2\,\partial x_1 - \partial x_0\,\partial x_1^2) \in \mathbb{K}[\partial x_0, \partial x_1]_{-1}$. As $\mu_{(x)}$ is bilinear, we can rewrite the previous equation as,*

$$\begin{array}{ccccccccc} 2\,\mu_{(x)}(x_0^2, \partial x_0^2\,\partial x_1) & + & 2\,\mu_{(x)}(x_0\, x_1, \partial x_0^2\,\partial x_1) & - & \mu_{(x)}(x_0^2, \partial x_0\,\partial x_1^2) & - & \mu_{(x)}(x_0\, x_1, \partial x_0\,\partial x_1^2) & = & \\ 2 \quad \partial x_1 & + & 2 \quad \partial x_0 & - & 0 & - & \partial x_1 & = & 2\,\partial x_0 + \partial x_1. \end{array}$$

We define the bilinear map

$$\mu : (\bigotimes_{i=1}^{q} R[\boldsymbol{x}_i]) \times \bigotimes_{i=1}^{q} (R[\boldsymbol{x}_i] \oplus R[\boldsymbol{\partial}\boldsymbol{x}_i]) \to \bigotimes_{i=1}^{q} (R[\boldsymbol{x}_i] \oplus R[\boldsymbol{\partial}\boldsymbol{x}_i]), \tag{3.12}$$

such that, if $(\bigotimes_{i=1}^{q} \boldsymbol{x}_i^{\boldsymbol{\alpha}_i}) \in (\bigotimes_{i=1}^{q} R[\boldsymbol{x}_i])$ and, for each $1 \leq i \leq q$, $g_i \in (R[\boldsymbol{x}_i] \oplus R[\boldsymbol{\partial}\boldsymbol{x}_i])$, the map acts on $(\bigotimes_{i=1}^{q} \boldsymbol{x}_i^{\boldsymbol{\alpha}_i}, \bigotimes_{i=1}^{q} g_i)$ as follows

$$\mu(\otimes_{i=1}^{q} \boldsymbol{x}_i^{\boldsymbol{\alpha}_i}, \bigotimes_{i=1}^{q} g_i) = \bigotimes_{i=1}^{q} \mu_{(i)}(\boldsymbol{x}_i^{\boldsymbol{\alpha}_i}, g_i).$$

Example 3.3.11. *Consider $g = (\partial x_{1,0}^2 \otimes x_{2,0}^2\, x_{2,1} - \partial x_{1,0}\,\partial x_{1,1} \otimes x_{2,1}^3) \in \mathbb{K}[\partial x_{1,0}, \partial x_{1,1}]_2 \otimes \mathbb{K}[x_{2,0}, x_{2,1}]_3$ and $f = x_{1,0}\, x_{2,1} + x_{1,1}\, x_{2,0} \in \mathbb{K}[x_{1,0}, x_{1,1}]_1 \otimes \mathbb{K}[x_{2,0}, x_{2,1}]_1$.*

Then $\mu_f(g) = \partial x_{1,0} \otimes (x_{2,0}^2\, x_{2,1}^2 + x_{2,0}\, x_{2,1}^3) + \partial x_{1,1} \otimes x_{2,1}^4 \in \mathbb{K}[\partial x_{1,0}, \partial x_{1,1}]_1 \otimes \mathbb{K}[x_{2,0}, x_{2,1}]_4$, as

$$\begin{array}{rcl} \mu(f,g) = & & \mu(x_{1,0}\, x_{2,1}, g) + \mu(x_{1,1}\, x_{2,0}, g) \\ = & \big(\mu(x_{1,0}\, x_{2,1}, \partial x_{1,0}^2 \otimes x_{2,0}^2\, x_{2,1}) & - \mu(x_{1,0}\, x_{2,1}, \partial x_{1,0}\,\partial x_{1,1} \otimes x_{2,1}^3)\big) + \mu(x_{1,1}\, x_{2,0}, g) \\ = & \big(\mu_{(1)}(x_{1,0}, \partial x_{1,0}^2) \otimes \mu_{(2)}(x_{2,1}, x_{2,0}^2\, x_{2,1}) & - \mu(x_{1,0}\, x_{2,1}, \partial x_{1,0}\,\partial x_{1,1} \otimes x_{2,1}^3)\big) + \mu(x_{1,1}\, x_{2,0}, g) \\ = & \big(\partial x_{1,0} \otimes x_{2,0}^2\, x_{2,1}^2 & - \mu(x_{1,0}\, x_{2,1}, \partial x_{1,0}\,\partial x_{1,1} \otimes x_{2,1}^3)\big) + \mu(x_{1,1}\, x_{2,0}, g) \\ & & \vdots \\ = & & (\partial x_{1,0} \otimes x_{2,0}^2\, x_{2,1}^2 - \partial x_{1,1} \otimes x_{2,1}^4) + (0 + \partial x_{1,0} \otimes x_{2,0}\, x_{2,1}^3) \end{array}$$

Remark 3.3.12. *If we restrict the domain of μ to $(\bigotimes_{i=1}^{q} R[\boldsymbol{x}_i]) \times (\bigotimes_{i=1}^{q} R[\boldsymbol{x}_i])$, then μ acts as multiplication, that is for multihomogeneous polynomials $f, g \in (\bigotimes_{i=1}^{q} R[\boldsymbol{x}_i])$, it holds*

$$\mu(f, g) = f\, g.$$

Given $f \in (\bigotimes_{i=1}^{q} R[\boldsymbol{x}_i])$, we define the linear map

$$\begin{array}{rccc} \mu_f : & \bigotimes_{i=1}^{q}(R[\boldsymbol{x}_i] \oplus R[\boldsymbol{\partial x_i}]) & \to & \bigotimes_{i=1}^{q}(R[\boldsymbol{x}_i] \oplus R[\boldsymbol{\partial x}_i]) \\ & g & \mapsto & \mu_f(g) = \mu(f, g). \end{array}$$

Using the isomorphisms of Prop. 3.3.5 and Prop. 3.3.6, for $\boldsymbol{d} \in \mathbb{N}^q$ and $f \in \mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]_{\boldsymbol{d}}$, if we restrict the map μ_f to $H^r_{\mathcal{P}}(\boldsymbol{m})$, for $r \in \mathbb{N}$, and then

$$\mu_f : H^r_{\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}}(\boldsymbol{m}) \to H^r_{\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}}(\boldsymbol{m} + \boldsymbol{d}).$$

Proposition 3.3.13. *[WZ94, Prop. 2.6] Consider a generic multihomogeneous system $\boldsymbol{F} \subset \mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$ with polynomials of multidegrees $\boldsymbol{d}_0, \ldots, \boldsymbol{d}_N$, respectively. Given a degree vector $\boldsymbol{m} \in \mathbb{Z}^q$, we consider the Weyman complex $K_\bullet(\boldsymbol{m})$. If there is $v \in \{-N+1, \ldots, N+1\}$ and $p \in \{0, \ldots, N+1\}$ such that*

$$K_v(\boldsymbol{m}) = K_{v,p}(\boldsymbol{m}) \otimes \mathbb{Z}[\boldsymbol{u}] \qquad \text{and} \qquad K_{v-1}(\boldsymbol{m}) = K_{v-1,p-1}(\boldsymbol{m}) \otimes \mathbb{Z}[\boldsymbol{u}],$$

then the map $\delta_v(\boldsymbol{m}) : K_v(\boldsymbol{m}) \to K_{v-1}(\boldsymbol{m})$ is

$$\delta_v(\boldsymbol{m}) = \sum_{k=0}^{N} \mu_{F_k} \otimes \Phi_k,$$

where $\mu_{F_k} \otimes \Phi_k$ denotes the tensor product of the maps μ_{F_k} and Φ_k from Def. 2.7.3.

Definition 3.3.14 (Koszul-type determinantal formula)**.** *With the notation of Prop. 3.3.13, when the Weyman complex reduces to*

$$K_\bullet(\boldsymbol{m}) : 0 \to K_{1,p+1}(\boldsymbol{m}) \otimes \mathbb{K}[\boldsymbol{u}] \xrightarrow{\delta_1(\boldsymbol{m})} K_{0,p}(\boldsymbol{m}) \otimes \mathbb{K}[\boldsymbol{u}] \to 0\,, \tag{3.13}$$

we say that the map $\delta_1(\boldsymbol{m})$ is a Koszul-type determinantal formula*.*

Example 3.3.15. *Consider the blocks of variables $\boldsymbol{x}_1 := \{x_{1,0}, x_{1,1}\}$ and $\boldsymbol{x}_2 := \{x_{2,0}, x_{2,1}\}$, and the systems $\boldsymbol{f} := (f_0, f_1, f_2)$ of multidegrees $\boldsymbol{d}_0 = \boldsymbol{d}_1 = \boldsymbol{d}_2 = (1, 1)$. That is,*

$$\begin{cases} f_0 = (a_{0,0}\, x_{1,0} + a_{1,0}\, x_{1,1})\, x_{2,0} + (a_{0,1}\, x_{1,0} + a_{1,1}\, x_{1,1})\, x_{2,1} \\ f_1 = (b_{0,0}\, x_{1,0} + b_{1,0}\, x_{1,1})\, x_{2,0} + (b_{0,1}\, x_{1,0} + b_{1,1}\, x_{1,1})\, x_{2,1} \\ f_2 = (c_{0,0}\, x_{1,0} + c_{1,0}\, x_{1,1})\, x_{2,0} + (c_{0,1}\, x_{1,0} + c_{1,1}\, x_{1,1})\, x_{2,1}. \end{cases} \tag{3.14}$$

As in [EMT16, Lem. 2.2], consider the degree vector $\boldsymbol{m} = (-1, 2)$. Then, the Weyman complex is

$$K_\bullet(\boldsymbol{m}, \boldsymbol{f}) : 0 \to K_{1,2}(\boldsymbol{m}, \boldsymbol{f}) \xrightarrow{\delta_1(\boldsymbol{m}, \boldsymbol{f})} K_{0,1}(\boldsymbol{m}, \boldsymbol{f}) \to 0,$$

where

$$\begin{cases} K_{1,2}(\boldsymbol{m}) = & S_1(0) \otimes S_2(1)^* \otimes \Big(\{e_0 \wedge e_1\} \oplus \{e_0 \wedge e_2\} \oplus \{e_1 \wedge e_2\}\Big) \\ K_{0,1}(\boldsymbol{m}) = & S_1(1) \otimes S_2(0)^* \otimes \Big(\{e_0\} \oplus \{e_1\} \oplus \{e_2\}\Big). \end{cases}$$

If we consider monomial bases for $K_{1,2}(\boldsymbol{m}, \boldsymbol{f})$ and $K_{0,1}(\boldsymbol{m}, \boldsymbol{f})$, then we can represent $\delta_1(\boldsymbol{m}, \boldsymbol{f})$ with the matrix that follows. Note that, the element $\partial 1 \in \mathbb{K}[\boldsymbol{\partial x}_1, \boldsymbol{\partial x}_1]$ corresponds to the dual of $1 \in \mathbb{K}[\boldsymbol{x}_1, \boldsymbol{x}_2]$.

	$x_{1,0} \otimes \partial 1 \otimes e_0$	$x_{1,1} \otimes \partial 1 \otimes e_0$	$x_{1,0} \otimes \partial 1 \otimes e_1$	$x_{1,1} \otimes \partial 1 \otimes e_1$	$x_{1,0} \otimes \partial 1 \otimes e_2$	$x_{1,1} \otimes \partial 1 \otimes e_2$
$1 \otimes \partial x_{2,0} \otimes (e_0 \wedge e_1)$	$-b_{0,0}$	$-b_{1,0}$	$a_{0,0}$	$a_{1,0}$	0	0
$1 \otimes \partial x_{2,1} \otimes (e_0 \wedge e_1)$	$-b_{0,1}$	$-b_{1,1}$	$a_{0,1}$	$a_{1,1}$	0	0
$1 \otimes \partial x_{2,0} \otimes (e_0 \wedge e_2)$	$-c_{0,0}$	$-c_{1,0}$	0	0	$a_{0,0}$	$a_{1,0}$
$1 \otimes \partial x_{2,1} \otimes (e_0 \wedge e_2)$	$-c_{0,1}$	$-c_{1,1}$	0	0	$a_{0,1}$	$a_{1,1}$
$1 \otimes \partial x_{2,0} \otimes (e_1 \wedge e_2)$	0	0	$-c_{0,0}$	$-c_{1,0}$	$b_{0,0}$	$b_{1,0}$
$1 \otimes \partial x_{2,1} \otimes (e_1 \wedge e_2)$	0	0	$-c_{0,1}$	$-c_{1,1}$	$b_{0,1}$	$b_{1,1}$

As we saw in the previous example, once we have fixed a basis for the map in Prop. 3.3.13, we can represent the Koszul-type determinantal formula by the determinant of a matrix. We refer to this matrix as a *Koszul resultant matrix*.

Corollary 3.3.16. *[Wey03, Prop. 5.2.4] Let $\boldsymbol{F}$ be a generic multihomogeneous system of polynomials with multidegrees $\boldsymbol{d}_0, \ldots, \boldsymbol{d}_N$, respectively. Let $\boldsymbol{m} \in \mathbb{Z}^q$ be a degree vector so that the Weyman complex $K_\bullet(\boldsymbol{m})$ becomes*

$$K_\bullet(\boldsymbol{m}) : 0 \to K_{v,p+v}(\boldsymbol{m}) \otimes \mathbb{K}[\boldsymbol{u}] \xrightarrow{\delta_v(\boldsymbol{m})} K_{v-1,p+v-1}(\boldsymbol{m}) \otimes \mathbb{K}[\boldsymbol{u}] \to 0\,.$$

Then, the map $\delta_v(\boldsymbol{m})$ of Prop. 3.3.13 is linear in the coefficients of $\boldsymbol{F}$ and so each element in any matrix representing $\delta_v(\boldsymbol{f})$ is a polynomial in $\mathbb{K}[\boldsymbol{u}]$ of degree one. Moreover, as the determinant of the complex is the resultant, the rank of both $K_v(\boldsymbol{m})$ and $K_{v+1}(\boldsymbol{m})$, as $\mathbb{K}[\boldsymbol{u}]$-modules, equals the degree of the resultant (Prop. 3.3.2), which is

$$\mathit{degree}(Res) = \sum_{k=0}^{N} \mathtt{MHB}(\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \ldots, \boldsymbol{d}_N).$$

Remark 3.3.17. *Under the assumptions of Prop. 3.3.13, if $p = 1$ and $v = 0$, the map $\delta_v(\boldsymbol{f})$ acts as a Sylvester map (Rem. 3.2.11), that is $(g_0, \ldots, g_N) \mapsto \sum_{k=0}^{N} g_k\, F_k$. In this case*

$$\delta_v(\boldsymbol{m})(g_0 \otimes e_0 + \cdots + g_N \otimes e_N) = \Big(\sum_{k=0}^{N} g_k\, F_k\Big) \otimes 1.$$

Determinantal formulas for the multiprojective resultant of unmixed systems, that is systems where the multidegree of each polynomial is the same, where extensively studied by several authors [SZ94, WZ94, CK04, DE03]. In [WZ94], the authors derive determinantal formulas using the Weyman complex. Moreover, they classify all the unmixed systems for which we can construct such formulas and describe the corresponding degree vectors. We will not introduce the notation needed to describe their result, but we refer the reader to the formulas of type 1 and 4 in [WZ94, Sec. 4] for further details.

Proposition 3.3.18. *[WZ94, Prop. 3.7] Consider a system $\boldsymbol{f} = (f_0, \ldots, f_N)$ in $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$ with multidegrees $\boldsymbol{d}_0, \ldots, \boldsymbol{d}_N$, respectively, such that $\boldsymbol{d}_0 = \cdots = \boldsymbol{d}_N = (d_1, \ldots, d_q)$. If there is a degree vector $\boldsymbol{m}$ such that Weyman complex give us a Koszul-type determinantal formula then, for each $1 \leq i \leq q$, $\min(d_i, n_i) = 1$.*

There are very few results about determinantal formulas for *mixed* multihomogeneous systems, that is, when the supports are not the same. We know such formulas for scaled multihomogeneous systems [EM12], that is when the supports are scaled copies of one of them, and for bivariate tensor-product polynomial systems [MT17, BMT17]. In Sec. 6.1, we use the Weyman complex to derive new formulas for families of mixed multilinear systems.

3.4 The sparse resultant

The sparse resultant generalizes the resultant to the context of toric varieties, see Sec. 2.12. In this section, we will just mention some references related to this object and we refer the reader to [CLO06, Ch. 7] for a general introduction to the subject and to [Emi05] for a more concise introduction with a focus on applications.

Theorem 3.4.1 (Sparse resultant)**.** *[CLO06, Thm. 7.6.2] Given $n+1$ full-dimensional polytopes $\Delta_0, \dots, \Delta_n$, we define the ring $\mathbb{Z}[\boldsymbol{u}]$ generated by the variables*

$$\{u_{i,\boldsymbol{\alpha}} : 0 \leq i \leq n \text{ and } \boldsymbol{\alpha} \in \Delta_i \cap \mathbb{Z}^n\},$$

and the generic system $(F_0, \dots, F_n)$ such that, for each i,

$$F_i := \sum_{\boldsymbol{\alpha} \in \Delta_i \cap \mathbb{Z}^n} u_{i,\boldsymbol{\alpha}} \boldsymbol{X}^{\boldsymbol{\alpha}}.$$

Let $\mathfrak{sp}$ be the specialization morphism, see Def. 3.2.2, and consider $F_i(\boldsymbol{c}) := \mathfrak{sp}_{\boldsymbol{c}}(F_i)$.

The sparse resultant is a polynomial, $\mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$, such that

$$\mathtt{res}(\boldsymbol{c}) \neq 0 \implies \left\{ \begin{array}{c} \text{The system } (F_0, \dots, F_n)(\boldsymbol{c}) \in \mathbb{K}[\mathbb{Z}^n]^{n+1} \\ \text{has no solutions over the torus } (\mathbb{C}^*)^n \end{array} \right\}.$$

Moreover, following the notation from the discussion in page 63, if $\mathtt{res}(\boldsymbol{c}) = 0$, then the homogenization of the system $(F_0, \dots, F_n)(\boldsymbol{c}) \in \mathbb{K}[\mathbb{Z}^n]^{n+1}$ has a solution over the projective toric variety X associated to $\Delta_0, \dots, \Delta_n$, see [CLO06, Sec. 7.3].

Proposition 3.4.2 (Degree of the sparse resultant)**.** *[CLO06, Ch. 7] With the same notation as in Thm. 3.4.1, for each i, consider the sets of variables*

$$\boldsymbol{u}_i := \{u_{i,\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \Delta_i \cap \mathbb{Z}^n\}.$$

Then, the resultant $\mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$, is a multihomogeneous polynomial with respect to the blocks of variables $\boldsymbol{u}_0, \dots, \boldsymbol{u}_n$. Its degree with respect to the block of variables $\boldsymbol{u}_i$ is the mixed volume (Def. 2.12.18) of $\Delta_0, \dots, \Delta_{i-1}, \Delta_{i+1}, \dots, \Delta_n$. That is,

$$\deg_{\boldsymbol{u}_i}(\mathtt{res}) = \mathrm{MV}(\Delta_0, \dots, \Delta_{i-1}, \Delta_{i+1}, \dots, \Delta_n).$$

A classical way of computing the sparse resultant is as a factor of the determinant of a matrix [Stu94, CE93, CE00]. This matrix is related to a Sylvester map, see Rem. 3.2.11. To construct the matrix, we compute a mixed subdivision of the Minkowski sum of the polytopes, $\Delta_0 + \cdots + \Delta_n$, see [CLO06, Sec. 7.6]. Hence, the complexity of this approach is related to the number of integer points in this Minkowski sum [Emi96]. There are also incremental constructions of the resultant, see [EC95] and [Emi05, Sec. 7.2.2], which in some cases results in smaller matrices, for example, for multihomogeneous systems [DE03]. We can also compute the sparse resultant as a quotient of determinants [D'A02]. Also, in some cases, we have determinantal formulas (Def. 3.2.14) for the sparse resultant, see for example [CDS98, CK00, Khe02, Wey03].

Other kinds of resultants Before finishing this chapter, we should mention that besides the three resultants presented before, resultant theory was extended to various areas, giving birth to different kind of resultants as, for example, resultants over unirational algebraic varieties [BEM00], residual resultants [BEM01, Bus01], determinantal resultants [Bus04], differential resultants [Rit32, Cha91], sparse differential resultants [LGY11] and multivariate subresultants [GV91, Cha95, Sza10].

Chapter 4

Gröbner basis

Gröbner bases are at the heart of most nonlinear algebra algorithms [BW98]. We use them to compute geometric and algebraic properties of ideals and modules. For example, Gröbner bases gives us a way to solve the *Ideal Membership Problem*, that is, to decide when a polynomial belong to certain ideal. Moreover, they allow us to do computations efficiently over quotient rings. It is also a tool to solve polynomial systems. In this chapter, we introduce some basic properties of Gröbner bases together with algorithms to efficiently compute them. Also, we will present bounds for the arithmetic complexity of these algorithms. The bounds rely on algebraic invariants of the ideals, as the Castelnuovo-Mumford regularity (Def. 2.8.3).

4.1 Gröbner basics[1]

The results from this section come mainly from [CLO15, Ch. 2].

A monomial is a product of variables in $\mathbb{K}[\boldsymbol{x}]$. We will write the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ as $\boldsymbol{x}^{\boldsymbol{\alpha}}$ where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ (see Ex. 2.12.8). The *degree* of a monomial is $\deg(\boldsymbol{x}^{\boldsymbol{\alpha}}) = \sum_i \alpha_i$. We say that a monomial $\boldsymbol{x}^{\boldsymbol{\alpha}}$ *divides* a second monomial $\boldsymbol{x}^{\boldsymbol{\beta}}$ if there is a third monomial $\boldsymbol{x}^{\boldsymbol{\gamma}}$ such that $\boldsymbol{x}^{\boldsymbol{\alpha}} \cdot \boldsymbol{x}^{\boldsymbol{\gamma}} = \boldsymbol{x}^{\boldsymbol{\beta}}$. In this case, we write $\boldsymbol{x}^{\boldsymbol{\alpha}} | \boldsymbol{x}^{\boldsymbol{\beta}}$. We can think the polynomials in $\mathbb{K}[\boldsymbol{x}]$ as finite $\mathbb{K}$-linear combinations of the monomials in $\mathbb{K}[\boldsymbol{x}]$. We consider a particular kind of total order for the monomials in $\mathbb{K}[\boldsymbol{x}]$, called *monomial ordering*.

Definition 4.1.1 (Monomial ordering)**.** *A monomial ordering $>$ is a total order for $\mathbb{N}^n$ such that*

- *It is a well-ordering for $\mathbb{N}^n$, that is, for every nonempty subset of $\mathbb{N}^n$, there is a minimal element with respect to $>$.*
- *Its compatible with the addition, that is, if $\boldsymbol{\alpha} > \boldsymbol{\beta}$, then $\forall \boldsymbol{\gamma} \in \mathbb{N}^n$, $\boldsymbol{\alpha}\,\boldsymbol{\gamma} > \boldsymbol{\beta}\,\boldsymbol{\gamma}$.*

We abuse of the notation and we write $\boldsymbol{x}^{\boldsymbol{\alpha}} > \boldsymbol{x}^{\boldsymbol{\beta}}$ to refer to $\boldsymbol{\alpha} > \boldsymbol{\beta}$.

[1] I took this wordplay from Svartz's PhD thesis [Sva14], which copied it from Sturmfels' book [Stu96]. As we say in Spanish: "A thief that steals from another thief has one hundred years of forgiveness" ("Ladrón que roba a ladrón tiene cien años de perdón").

Lemma 4.1.2. *[CLO15, Ex. 2.3.7] Let $>$ be any monomial ordering and consider two monomials $\boldsymbol{x}^{\boldsymbol{\alpha}}$ and $\boldsymbol{x}^{\boldsymbol{\beta}}$ such that $\boldsymbol{x}^{\boldsymbol{\beta}}$ divides $\boldsymbol{x}^{\boldsymbol{\alpha}}$. Then, $\boldsymbol{x}^{\boldsymbol{\alpha}} > \boldsymbol{x}^{\boldsymbol{\beta}}$.*

The two main monomial orders that appear in practice are Lex and GRevLex.

Definition 4.1.3. *Consider a monomial ordering $>$ such that $x_1 > x_2 > \cdots > x_n$.*

- *The order $>$ is a* lexicographical ordering *(*Lex*) if $\boldsymbol{x}^{\boldsymbol{\alpha}} > \boldsymbol{x}^{\boldsymbol{\beta}}$ if and only if, there is a $k \leq n$ such that,*

$$\begin{cases} (\forall i < k) & \alpha_i = \beta_i \\ & \alpha_k > \beta_k \end{cases}$$

  Roughly speaking, this order is equivalent to the one that we use in a dictionary.

  We write this order as $>_{\text{Lex}}$.

- *The order $>$ is a* graded reverse lexicographical monomial ordering *(*GRevLex*) if $\boldsymbol{x}^{\boldsymbol{\alpha}} > \boldsymbol{x}^{\boldsymbol{\beta}}$ if and only if,*

$$\begin{cases} \sum_i \alpha_i > \sum_i \beta_i & \text{or} \\ \sum_i \alpha_i = \sum_i \beta_i & \text{and } \exists k \text{ such that} \quad (\forall i > k) \quad \alpha_i = \beta_i \\ & \qquad\qquad\qquad\quad \text{and} \qquad \alpha_k < \beta_k \end{cases}$$

  that is, $\boldsymbol{x}^{\boldsymbol{\alpha}} > \boldsymbol{x}^{\boldsymbol{\beta}}$ if the degree of $\boldsymbol{x}^{\boldsymbol{\alpha}}$ is bigger than the one of $\boldsymbol{x}^{\boldsymbol{\beta}}$, or they have the same degree and the degree of the smallest variable (with respect to $>$) such that its degree in $\boldsymbol{x}^{\boldsymbol{\alpha}}$ is different from its degree in $\boldsymbol{x}^{\boldsymbol{\beta}}$, is smaller in $\boldsymbol{x}^{\boldsymbol{\alpha}}$ than in $\boldsymbol{x}^{\boldsymbol{\beta}}$.

  We write this order as $>_{\text{GRevLex}}$.

Example 4.1.4. *Consider the* Lexicographical order $>_{\text{Lex}}$ *on $\mathbb{K}[x, y]$, such that $x >_{\text{Lex}} y$. Then, $x^2 >_{\text{Lex}} x\,y$, $x >_{\text{Lex}} y^{100}$ and $x\,y^2 >_{\text{Lex}} x\,y$.*

Consider the Graded Reverse Lexicographical order $>_{\text{GRevLex}}$ *on $\mathbb{K}[x, y, z]$, such that $x >_{\text{GRevLex}} y >_{\text{GRevLex}} z$. Then, $x\,y^2 >_{\text{GRevLex}} x^2\,z$ and $y^{100} >_{\text{GRevLex}} x$.*

In whats follows, we say that a monomial *appears* in a polynomial, if the coefficient associated to the monomial in the polynomial is not zero.

Definition 4.1.5 (Leading monomial, coefficient and term)**.** *Consider $>$ a monomial ordering and a polynomial non-zero $f = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} c_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}]$.*

- *The leading monomial of f with respect to $>$, $\texttt{LM}_>(f)$, is the biggest monomial (with respect to $>$) that appears in f, that is,*

$$\texttt{LM}_>(f) := \max_{>} \{\boldsymbol{x}^{\boldsymbol{\alpha}} : c_{\boldsymbol{\alpha}} \neq 0\}.$$

- *The leading coefficient of f with respect to $>$, $\texttt{LC}_>(f)$, is the coefficient associated to the leading monomial of f, that is,*

$$\texttt{LC}_>(f) := c_{\boldsymbol{\alpha}} \text{ such that } \texttt{LM}_>(f) = \boldsymbol{x}^{\boldsymbol{\alpha}}.$$

- *The leading term of f is the product of the leading monomial and coefficient of f, that is,*

$$\texttt{LT}_>(f) := \texttt{LC}_>(f)\,\texttt{LM}_>(f).$$

We can extend the notion of leading monomials to ideals.

Definition 4.1.6 (Initial ideal)**.** *Let $>$ be a monomial ordering and $I \subset \mathbb{K}[\boldsymbol{x}]$ an ideal. The* initial ideal *of I is the ideal generated by all the leading monomials of the polynomials in I, that is,*

$$\mathtt{LM}_{>}(I) := \langle \{\mathtt{LM}_{>}(f) : f \in I\} \rangle.$$

The initial ideal is a monomial ideal, see Def. 2.3.10.

We use the notion of initial ideal to define *Gröbner bases*.

Definition 4.1.7 (Gröbner basis)**.** *Let $>$ be a monomial ordering and $I \subset \mathbb{K}[\boldsymbol{x}]$ an ideal. A Gröbner basis of I with respect to $>$ is a set $G \subset I$ such that*

$$\mathtt{LM}_{>}(I) = \langle \{\mathtt{LM}_{>}(g) : g \in G\} \rangle.$$

Equivalently, G is a Gröbner basis if, for every $f \in I$, there is a $g \in G$ such that $\mathtt{LM}_{>}(g)$ divides $\mathtt{LM}_{>}(f)$.

Proposition 4.1.8. *[CLO15, Cor. 2.5.6] A Gröbner basis G of I is also a basis of I, that is, $\langle G \rangle = I$.*

Gröbner bases are not unique, neither necessarily finite. To make them unique and finite, we define *reduced Gröbner bases*.

Definition 4.1.9 (Reduced Gröbner basis)**.** *We say that a Gröbner basis G is a* reduced Gröbner basis *if and only if,*

- *every polynomial in G is monic, that is, for all $g \in G$, $LC_{>}(g) = 1$, and*
- *for all $g \in G$, no monomial appearing in g belongs to $\mathtt{LM}_{>}(G \setminus \{g\})$.*

Proposition 4.1.10. *[CLO15, Thm. 2.7.5] Let $>$ be a monomial ordering and $I \subset \mathbb{K}[\boldsymbol{x}]$ an ideal. Then, there is an unique reduced Gröbner basis G for I with respect to $>$ and it is finite.*

We mention that Gröbner basis were generalized to modules, but we will not discuss this subject in the thesis; we refer the reader to [CLO06, Sec. 5.2] and references there in.

4.2 Normal forms, division algorithm and Buchberger

The results from this section come mainly from [CLO15, Ch. 2].

Proposition 4.2.1 (Normal form)**.** *[CLO15, Prop. 2.6.1] Let $>$ be a monomial ordering, $I \subset \mathbb{K}[\boldsymbol{x}]$ an ideal and G a Gröbner basis of I with respect to $>$. Consider $f \in \mathbb{K}[\boldsymbol{x}]$. Then, there are unique $g \in I$ and $r \in \mathbb{K}[\boldsymbol{x}]$, such that $f = g + r$ and no monomial appearing in r belongs to $\mathtt{LM}_{>}(I)$. The polynomial r is called the* normal form *of g with respect of I and $>$. We denote r by $\mathtt{NF}_{>,I}(f)$.*

We can check if a polynomial belongs to an ideal by regarding its normal form.

Corollary 4.2.2. *[CLO15, Cor. 2.6.2] Consider a polynomials $f, g \in \mathbb{K}[\boldsymbol{x}]$. Then, $\mathtt{NF}_{>,I}(f) = \mathtt{NF}_{>,I}(g)$ if and only if $f - g \in I$. In particular, $f \in I$ if and only if $\mathtt{NF}_{>,I}(f) = 0$, and $f - \mathtt{NF}_{>,I}(f) \in I$.*

Algorithm 1 Division algorithm

Input: A polynomial f, a list of polynomials $(g_1, \ldots, g_s)$ and a monomial ordering $>$.
Output: Polynomials r (remainder) and $q_1, \ldots, q_s$ (cofactors) such that $f = \sum_i q_i\, g_i + r$ and none of the monomials appearing in r are divisible by any $\mathtt{LM}_<(g_i)$, for $1 \leq i \leq s$ (or $r = 0$).

$r, q_1, \ldots, q_s \leftarrow 0, 0, \ldots, 0$
$p \leftarrow f$
while $p \neq 0$ **do**
  if $\exists g_i \in G$ such that $\mathtt{LM}_>(g_i)$ divides $\mathtt{LM}_>(p)$ **then**
    $j \leftarrow$ minimal i such that $\mathtt{LM}_>(g_i)$ divides $\mathtt{LM}_>(p)$
    $q_i \leftarrow q_i + \frac{\mathtt{LT}_>(p)}{\mathtt{LT}_>(g_j)}$
    $p \leftarrow p - \frac{\mathtt{LT}_>(p)}{\mathtt{LT}_>(g_j)} g_j$
  else
    $r \leftarrow r + \mathtt{LT}_>(p)$
    $p \leftarrow p - \mathtt{LT}_>(p)$
  end if
end while
return $r, q_1, \ldots, q_s$

There is an algorithmic way of computing normal forms which involves Gröbner basis and the division algorithm.

We call the polynomial r from Alg. 1 the *remainder* of the division of f by $(g_1, \ldots, g_s$ with respect to $>$.

Proposition 4.2.3. *[CLO15, Thm. 2.3.3] The division algorithm always terminates and it is correct.*

Even if the division algorithm always terminates, its remainder might depend on the order of the polynomials $(g_1, \ldots, g_r)$.

Example 4.2.4. *Consider a lexicographical order $>_{\mathrm{Lex}}$ such that $x >_{\mathrm{Lex}} y$. We consider the division of the polynomial xy by the list $(x\,y + y, x)$ with respect to $>_{\mathrm{Lex}}$. As $\mathtt{LM}_{>_{\mathrm{Lex}}}(x\,y + y)$ divides xy, we use $(x\,y + y)$ to reduce xy and obtain $-y$. As we can not reduce $-y$ any further, the remainder of the division is $r = -y$. We swap the elements in the list $(x\,y + y, x)$ and consider the division of the polynomial $x\,y$ by the list $(x, x\,y + y)$ with respect to $>_{\mathrm{Lex}}$. As x divides $x\,y$, the remainder of the division is $r = 0$.*

As we see in Ex. 4.2.4, the remainder of the division algorithm is sensitive to the order of the polynomials in the list $(g_1, \ldots, g_r)$. However, when $(g_1, \ldots, g_r)$ is a Gröbner basis, their order does not matter.

Proposition 4.2.5. *[CLO15, Prop. 2.6.1] If G is a Gröbner basis of I with respect to a monomial ordering $>$, then the remainder of the division of $f \in \mathbb{K}[\boldsymbol{x}]$ by G (sorted in any order) with respect to $>$, that is, the polynomial r in Alg. 1, equals $\mathtt{NF}_{>,I}(f)$, and so it is unique.*

We can use the division algorithm to compute Gröbner bases.

Proposition 4.2.6 (Least common multiple)**.** *[CLO15, Prop. 4.3.13] Given two monomials $\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\beta}} \in \mathbb{K}[\boldsymbol{x}]$, consider the intersection of the ideals generated by these monomials, $\langle \boldsymbol{x}^{\boldsymbol{\alpha}} \rangle \cap \langle \boldsymbol{x}^{\boldsymbol{\beta}} \rangle$. This ideal is a principal (Def. 2.2.8) monomial ideal, that is, there is a monomial, that we call the* least common multiple *of $\boldsymbol{x}^{\boldsymbol{\alpha}}$ and $\boldsymbol{x}^{\boldsymbol{\beta}}$ and denote by $\mathtt{LCM}(\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\beta}}) \in \mathbb{K}[\boldsymbol{x}]$, such that*

$$\langle \mathtt{LCM}(\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\beta}}) \rangle = \langle \boldsymbol{x}^{\boldsymbol{\alpha}} \rangle \cap \langle \boldsymbol{x}^{\boldsymbol{\beta}} \rangle.$$

We can compute the least common multiple of $\boldsymbol{x}^{\boldsymbol{\alpha}}$ and $\boldsymbol{x}^{\boldsymbol{\beta}}$ as

$$\mathtt{LCM}(\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\beta}}) := \boldsymbol{x}^{(\max(\alpha_1,\beta_1),\ldots,\max(\alpha_n,\beta_n))}.$$

Definition 4.2.7 (S-polynomial)**.** *Consider two polynomial $f, g \in \mathbb{K}[\boldsymbol{x}]$ and a monomial ordering $>$. The* S-polynomial *of f and g, $\mathtt{Spol}(f, g)$, is a particular polynomial combination of f and g defined by*

$$\mathtt{Spol}_>(f, g) := \frac{\mathtt{LT}_>(g)}{\mathtt{LCM}(\mathtt{LM}_>(f), \mathtt{LM}_>(g))}\, f - \frac{\mathtt{LT}_>(f)}{\mathtt{LCM}(\mathtt{LM}_>(f), \mathtt{LM}_>(g))}\, g.$$

An important property of the S-polynomial $\mathtt{Spol}_>(f, g)$ is that its leading monomial with respect to $>$ is smaller than $\mathtt{LCM}(\mathtt{LM}_>(f), \mathtt{LM}_>(g))$, see [CLO15, Ex. 2.6.7]. We use S-polynomials to compute Gröbner bases.

Algorithm 2 Buchberger's algorithm

Input: List of polynomials $(f_1, \ldots, f_s)$ and a monomial ordering $>$.
Output: Gröbner basis G of $\langle f_1, \ldots, f_s \rangle$ with respect to $>$.

$G \leftarrow \{f_1, \ldots, f_s\}$
$G' \leftarrow \{\}$
while $G \neq G'$ **do**
  $G' \leftarrow G$
  $L \leftarrow$ List (G') (we sort the elements of G' with respect to an arbitrary order).
  for each pair $f, g \in G'$ such that $f \neq g$ **do**
    $r \leftarrow$ Remainder of the division algorithm of $\mathtt{Spol}(f, g)$ with respect to L and $>$.
    if $r \neq 0$ **then**
      $G \leftarrow G \cup \{r\}$
    end if
  end for
end while
return G

Theorem 4.2.8 (Buchberger's Criterion)**.** *[CLO15, Thm. 2.6.6] Given an ideal I, a monomial order $>$ and a finite set $G \subset I$ such that for each pair of different polynomials $f, g \in G$ the remainder of the division algorithm of the $\mathtt{Spol}_>(f, g)$ with respect to G and $>$ is zero, then G is a Gröbner basis of I with respect to $>$.*

Corollary 4.2.9. *Buchberger's algorithm (Alg. 2) is correct.*

Example 4.2.10 (Continuation of Ex. 4.2.4)**.** *If we consider the list $(x\,y+y, x)$, then*

$$\texttt{Spol}_{>}(x\,y+y, x) = \frac{x}{x}\,(x\,y+y) - \frac{x\,y}{x}\,(x) = y.$$

By Buchberger's criterion, the set $G := \{x\,y+y, x, y\}$ is a Gröbner basis for the ideal $\langle x\,y+y, x\rangle$ with respect to $>$, because

$$\texttt{Spol}_{>}(x\,y+y, y) = \tfrac{y}{y}\,(x\,y+y) - \tfrac{x\,y}{y}\,(y) = y \in G, \text{ and}$$
$$\texttt{Spol}_{>}(x, y) = \tfrac{y}{x\,y}\,(x) - \tfrac{x}{x\,y}\,(y) = 0.$$

Proposition 4.2.11 (Termination of Buchberger's Algorithm)**.** *[CLO15, Thm. 2.7.2] As $\mathbb{K}[\boldsymbol{x}]$ is a Noetherian ring, see Prop. 2.2.4, Buchberger's algorithm terminates in a finite number of step.*

4.3 Properties of the monomial order $\mathrm{GRevLex}$

The results from this section come mainly from [CLO15, Ch. 8].

We study some properties of the monomial order $\mathrm{GRevLex}$ (Def. 4.1.3). We show how it allow us to transform an affine variety, that is, the solution set of an arbitrary polynomial system (Sec. 2.4), into a closely related projective variety, that is, the solution set of a system of homogeneous polynomials (Sec. 2.5).

Definition 4.3.1 (Homogenization)**.** *[CLO15, Prop. 8.4.2] Consider a polynomial $f \in \mathbb{K}[\boldsymbol{x}]$ of degree d. We define $f^h \in \mathbb{K}[\boldsymbol{x}][x_0]$ as the homogeneous polynomial $f^h := x_0^d\, f(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0})$. This polynomial has degree d.*

Given an ideal $I \subset \mathbb{K}[\boldsymbol{x}]$, we define the homogenization *of the ideal I, $I^h \subset \mathbb{K}[\boldsymbol{x}][x_0]$, as the homogeneous ideal in $\mathbb{K}[\boldsymbol{x}][x_0]$ generated by the homogenization of all the polynomials in I, that is, $I^h := \langle\{f^h : f \in I\}\rangle$. The ideal I^h is a homogeneous ideal, see Def. 2.3.5.*

We can write the homogeneous polynomials in I^h as the product of a power of x_0 and the homogenization of a polynomial in I.

Proposition 4.3.2. *[CLO15, Prop. 8.2.7.iv] Consider an ideal $I \subset \mathbb{K}[\boldsymbol{x}]$ and a homogeneous polynomial $\bar{f} \in I^h$. Let d be the maximal number such that x_0^d divides $\bar{f}$. Then, there is $f \in I$ such that*

$$\bar{f} = x_0^d \cdot f^h.$$

The homogenization of an ideal contains any ideal generated by the homogenization of a set of its generators.

Proposition 4.3.3. *Consider an ideal $\langle f_1, \ldots, f_r\rangle = I$ generated by r polynomials, then the homogenization of the ideal contains the ideal generated by the homogenization of its r generators, that is,*

$$\langle f_1^h, \ldots, f_r^h\rangle \subseteq I^h$$

However, in general, the homogenization of an ideal differs from an ideal generated by the homogenization of a set of its generators.

Example 4.3.4 (Continuation of Ex. 4.2.4)**.** *We consider $x\,y + y \in \mathbb{K}[x, y]$, whose homogenization in $\mathbb{K}[x, y, z]$ is $z^2(\frac{x}{z}\,\frac{y}{z} + \frac{y}{z}) = x\,y + y\,z$. The homogenization of $x \in \mathbb{K}[x, y, z]$ is $z(\frac{x}{z}) = x$. Note that the ideal $\langle\{x\,y + y\,z, x\}\rangle \subset \mathbb{K}[x, y, z]$, obtained by homogenizing the set of generators of $\{x\,y + y, x\}$, is different from the homogenization $\langle\{x\,y + y, x\}\rangle^h \subset \mathbb{K}[x, y, z]$. We already proved that $y \in \langle\{x\,y + y\,z, x\}\rangle$, and so, its homogenization, $z\,(\frac{y}{z}) = y$, belongs to $\langle\{x\,y+y, x\}\rangle^h$. However, $y \notin \langle\{x\,y+y\,z, x\}\rangle$ because every polynomial of degree one in $\langle\{x\,y + y\,z, x\}\rangle$ is divisible by x. Nevertheless, note that $y\,z = 1\,(x\,y + y\,z) - y\,(x)$ belongs to $\langle\{x\,y + y\,z, x\}\rangle \subset \mathbb{K}[x, y, z]$.*

To generalize our last observation in Ex. 4.3.4, we need to introduce the *dehomogenization homomorphism*.

Definition 4.3.5 (Dehomogenization)**.** *Given a polynomial $g \in \mathbb{K}[\boldsymbol{x}][x_0]$, its* dehomogenization*, $\chi(g)$, corresponds to the specialization of the variable x_0 in g to 1. The dehomogenization is a surjective ring homomorphism, $\chi : \mathbb{K}[\boldsymbol{x}][x_0] \to \mathbb{K}[\boldsymbol{x}]$, and, if $J \subset \mathbb{K}[\boldsymbol{x}][x_0]$ is an ideal, $\chi(J) \subset \mathbb{K}[\boldsymbol{x}]$ is an ideal too. If we restrict the domain of χ to a graded piece of $\mathbb{K}[\boldsymbol{x}][x_0]$, $\mathbb{K}[\boldsymbol{x}][x_0]_d$, then χ is an injective map.*

By Prop. 4.3.2, we can think about the dehomogenization as the "anti-homogenization".

Proposition 4.3.6. *Given an ideal $I \subset \mathbb{K}[\boldsymbol{x}]$, then*

$$I = \chi(I^h).$$

Moreover, the affine variety defined by the dehomogenization of a homogeneous ideal $J \subset \mathbb{K}[\boldsymbol{x}][x_0]$, $\mathbb{V}_{\mathbb{K}^n}(\chi(J))$, is isomorphic to the affine piece $\mathbb{V}_{\mathbb{P}^n}(J) \cap U_0$ of $\mathbb{V}_{\mathbb{P}^n}(J)$, see Prop. 2.5.11.

The dehomogenization of the intersection of two homogeneous ideals equals the intersection of the dehomogenization of the two ideals, see discussion after [CLO15, Prop. 8.2.7].

Proposition 4.3.7. *Given two homogeneous ideals $\bar{I}, \bar{J} \subset \mathbb{K}[\boldsymbol{x}][x_0]$, then,*

$$\chi(\bar{I}) \cap \chi(\bar{J}) = \chi(\bar{I} \cap \bar{J}).$$

In particular, given two ideals $I, J \subset \mathbb{K}[\boldsymbol{x}]$, then

$$I \cap J = \chi((I \cap J)^h) = \chi(I^h \cap J^h).$$

Whenever we have a principal ideal and we know its unique generator, we can compute easily its homogenization.

Lemma 4.3.8. *The homogenization of a principal ideal $\langle f\rangle \subset \mathbb{K}[\boldsymbol{x}]$, see Def. 2.2.8, equals the ideal generated by the homogenization of its generator f, that is, $\langle f\rangle^h = \langle f^h\rangle$.*

Corollary 4.3.9. *Consider the ideal $I = \langle f_1, \ldots, f_r\rangle \subset \mathbb{K}[\boldsymbol{x}]$. Then, the dehomogenization of I^h is equal to the dehomogenization of $\langle f_1^h, \ldots, f_r^h\rangle \subset \mathbb{K}[\boldsymbol{x}][x_0]$, that is $\chi(I^h) = \chi(\langle f_1^h, \ldots, f_r^h\rangle)$.*

We can homogenize an ideal by computing its Gröbner basis with respect to the monomial ordering GRevLex, which is a *graded monomial ordering*.

Definition 4.3.10 (Graded monomial ordering)**.** *We say that a monomial ordering $>$ is* graded *if, for every two monomials of different degrees, the biggest monomial with respect to $>$ corresponds to the one with bigger degree. That is, given two monomials $\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\beta}}$,*

$$\deg(\boldsymbol{x}^{\boldsymbol{\alpha}}) > \deg(\boldsymbol{x}^{\boldsymbol{\beta}}) \implies \boldsymbol{x}^{\boldsymbol{\alpha}} > \boldsymbol{x}^{\boldsymbol{\beta}}.$$

We can extend graded monomial orderings on $\mathbb{K}[\boldsymbol{x}]$ to graded monomial orderings on $\mathbb{K}[\boldsymbol{x}][x_0]$.

Proposition 4.3.11. *[CLO15, Ex. 8.4.4] Given a graded monomial orderings $>$ for $\mathbb{K}[\boldsymbol{x}]$, we can extend it to the graded monomial ordering $>_h$ on $\mathbb{K}[\boldsymbol{x}][x_0]$ such that for $\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\beta}} \in \mathbb{K}[\boldsymbol{x}][x_0]$ it holds,*

$$\boldsymbol{x}^{\boldsymbol{\alpha}} >_h \boldsymbol{x}^{\boldsymbol{\beta}} \iff \begin{cases} \deg(\boldsymbol{x}^{\boldsymbol{\alpha}}) > \deg(\boldsymbol{x}^{\boldsymbol{\beta}}), \text{ or} \\ \deg(\boldsymbol{x}^{\boldsymbol{\alpha}}) = \deg(\boldsymbol{x}^{\boldsymbol{\beta}}) \text{ and } \chi(\boldsymbol{x}^{\boldsymbol{\alpha}}) > \chi(\boldsymbol{x}^{\boldsymbol{\beta}}). \end{cases}$$

We can define the monomial ordering GRevLex using Prop. 4.3.11. In what follows, we write $\mathrm{GRevLex}(x_1 < \cdots < x_n)$ to refer to the GRevLex monomial ordering $<_{\mathrm{GRevLex}}$ such that $x_1 <_{\mathrm{GRevLex}} \cdots <_{\mathrm{GRevLex}} x_n$.

Example 4.3.12. *Consider the monomial ordering* $\mathrm{GRevLex}(x_1 < \cdots < x_n)$ *on* $\mathbb{K}[\boldsymbol{x}]$*. Then, its extension, as defined in Prop. 4.3.11, corresponds to the monomial ordering* $\mathrm{GRevLex}(x_0 < x_1 < \cdots < x_n)$ *on* $\mathbb{K}[\boldsymbol{x}][x_0]$.

The main property of the graded monomial orderings is that, when we homogenize a polynomial, they preserve its leading monomials, see [CLO15, Lem. 8.4.5]. That is, for every $f \in \mathbb{K}[\boldsymbol{x}]$,

$$\mathtt{LM}_{<}(f) = \mathtt{LM}_{<_h}(f^h).$$

This allow us to transform Gröbner bases from $\mathbb{K}[\boldsymbol{x}][x_0]$ to $\mathbb{K}[\boldsymbol{x}]$ and vice-versa.

Proposition 4.3.13. *[CLO15, Ex. 8.4.4] Consider an ideal $I \subset \mathbb{K}[\boldsymbol{x}]$ and let G by a Gröbner basis of I with respect to a monomial ordering $>$* $\mathrm{GRevLex}(x_1 < \cdots < x_n)$*. Let $G^h := \{g^h : g \in G\}$ be the homogenization of the polynomials in G. Then, G^h is a Gröbner basis for $I^h \subset \mathbb{K}[\boldsymbol{x}][x_0]$ with respect to* $\mathrm{GRevLex}(x_0 < x_1 < \cdots < x_n)$.

Proposition 4.3.14. *[FSS14, Prop. 3.5] Let $\bar{I} \subset \mathbb{K}[\boldsymbol{x}][x_0]$ be a homogeneous ideal and let $\bar{G}$ be a finite set of homogeneous polynomials forming a Gröbner basis of $\bar{I}$ with respect to* $\mathrm{GRevLex}(x_0 < x_1 < \cdots < x_n)$*. Let $G := \{\chi(\bar{g}) : \bar{g} \in \bar{G}\}$ be the dehomogenization of the elements in $\bar{G}$. Then, G is a Gröbner basis for the ideal $\chi(I) \subset \mathbb{K}[\boldsymbol{x}]$ with respect to* $\mathrm{GRevLex}(x_1 < \cdots < x_n)$.

We can extend the previous propositions to extended graded monomial orderings, see the proof of [CLO15, Thm. 8.4.4].

4.4 Computing Gröbner bases for homogeneous ideals

In the previous section we observed the following:

- By Cor. 4.3.9, given a finite set of generators of an ideal, $\langle f_1, \dots, f_r \rangle = J \subset \mathbb{K}[\boldsymbol{x}]$, we can construct a homogeneous ideal $I = \langle f_1^h, \dots, f_r^h \rangle \subset \mathbb{K}[\boldsymbol{x}][x_0]$ such that its dehomogenization is the original ideal, that is, $J = \chi(I)$.
- By Prop. 4.3.14, from a Gröbner basis G for a homogeneous ideal I with respect to $\mathrm{GRevLex}(x_0 < x_1 < \cdots < x_n)$, we can recover a Gröbner basis for $\chi(I)$ with respect to $\mathrm{GRevLex}(x_1 < \cdots < x_n)$ from the dehomogenization of G, that is, the set $\{\chi(g) : g \in G\}$.

Hence, to compute Gröbner bases with respect to $\mathrm{GRevLex}$, and more in general with respect to graded monomial orderings, we only need an algorithm that computes Gröbner bases for homogeneous ideals. In this section, we discuss such an algorithm.

Our presentation follows the linear algebra approach introduced by Lazard [Laz83] and further developed by Faugère [Fau99, Fau02]. We introduce simplified versions of the algorithms that are implemented in practice. We refer the reader to [CLO15, Ch. 10] and [EF17] and references there in for more details.

Definition 4.4.1 (d-Gröbner basis)**.** *Let $>$ be a graded monomial ordering, $I \subset \mathbb{K}[\boldsymbol{x}][x_0]$ a homogeneous ideal and consider $d \in \mathbb{N}$. A* d-Gröbner basis *for I with respect to $>$ is a finite set $G \subset I$ such that for every $f \in I$ of degree at most d, $\mathtt{LM}_>(f) \in \langle \{\mathtt{LM}_>(g) : g \in G\} \rangle$.*

As every ideal in $\mathbb{K}[\boldsymbol{x}][x_0]$ has a finite Gröbner bases (Prop. 4.1.10), there is a d such that the d-Gröbner basis is a Gröbner basis.

Theorem 4.4.2 (Degree of regularity)**.** *[Laz83, Sec. III.B]. For each ideal $I \subset \mathbb{K}[\boldsymbol{x}][x_0]$ and each graded monomial ordering $>$, there is a d such that the d-Gröbner basis of I with respect to $>$ is a Gröbner basis of I with respect to $>$. The* degree of regularity *of I with respect to $>$ is the minimal d with this property.*

Using Thm. 4.4.2, we reduce the computation of a Gröbner basis to the computation of triangular bases of $\mathbb{K}$-vector spaces. To do so, we introduce the *Macaulay matrix*, a generalization of Sylvester's matrix (Ex. 3.2.16).

Definition 4.4.3 (Macaulay matrix)**.** *Let $>$ be a monomial ordering, $\boldsymbol{f} = (f_1, \dots, f_r)$ a list of homogeneous polynomials in $\mathbb{K}[\boldsymbol{x}][x_0]$ of degrees $d_1, \dots, d_r$, respectively. The Macaulay matrix of $\boldsymbol{f}$ in degree $d \in \mathbb{N}$ is the matrix $\mathcal{M}_{>,d}(\boldsymbol{f}) \in \mathbb{K}^{\left(\sum_i \binom{n+d-d_i}{d-d_i}\right) \times \binom{n+d}{d}}$ where*

- *The matrix $\mathcal{M}_{>,d}(\boldsymbol{f})$ has $\binom{n+d}{d}$ columns. We index each column by a monomial in $\mathbb{K}[\boldsymbol{x}][x_0]_d$. We sort the columns in decreasing order with respect to the monomial ordering $>$.*
- *The matrix $\mathcal{M}_{>,d}(\boldsymbol{f})$ has $\sum_i \binom{n+d-d_i}{d-d_i}$ rows. We index each row by a pair $(i, \boldsymbol{x}^{\boldsymbol{\alpha}})$, where $i \in \{1, \dots, r\}$ and $\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}][x_0]_{d-d_i}$. We sort the rows in increasing order with respect to the following order for their indices,*

$$(i, \boldsymbol{x}^{\boldsymbol{\alpha}}) < (j, \boldsymbol{x}^{\boldsymbol{\beta}}) \iff (i < j) \text{ or } (i = j \text{ and } \boldsymbol{x}^{\boldsymbol{\alpha}} < \boldsymbol{x}^{\boldsymbol{\beta}}).$$

We refer to the pair $(i, \boldsymbol{x}^{\boldsymbol{\alpha}})$ as the signature *of the row.*

- *The element in the row with signature $(i, \boldsymbol{x}^{\boldsymbol{\alpha}})$ and column indexed by $\boldsymbol{x}^{\boldsymbol{\beta}}$ corresponds to the coefficient of the monomial $\boldsymbol{x}^{\boldsymbol{\beta}}$ in the polynomial $\boldsymbol{x}^{\boldsymbol{\alpha}}$ f_i.*

From a commutative algebra point of view, the matrix $\mathcal{M}_{>,d}(\boldsymbol{f})$ represents the Sylvester map (Rem. 3.2.11) $(g_1, \ldots, g_r) \mapsto \sum_i g_i\, f_i$, where each $g_i \in \mathbb{K}[\boldsymbol{x}][x_0]_{d-d_i}$. This map corresponds to the first map of the strand of degree d of the Koszul complex $\mathcal{K}(f_1, \ldots, f_r)$, see Def. 2.7.5. As we explain at the end of the section, this relation is a key ingredient to improve the efficiency of our algorithms to compute Gröbner bases.

From a linear algebra point of view, a linear combination of the rows of $\mathcal{M}_{>,d}(\boldsymbol{f})$ represents a polynomial f of degree d in the ideal generated by $\boldsymbol{f}$, that is, $f \in \langle \boldsymbol{f} \rangle_d$. More precisely, if we multiply a non-zero constant c by a row with signature $(i, \boldsymbol{x}^{\boldsymbol{\alpha}})$, we obtain a row whose elements correspond to the coefficients of $c\, \boldsymbol{x}^{\boldsymbol{\alpha}} f_i$. If we add two rows with signatures $(i, \boldsymbol{x}^{\boldsymbol{\alpha}})$ and $(j, \boldsymbol{x}^{\boldsymbol{\beta}})$ we obtain a row whose elements correspond to the coefficients of $\boldsymbol{x}^{\boldsymbol{\alpha}} f_i + \boldsymbol{x}^{\boldsymbol{\beta}} f_j$. Let $v(d) \in (\mathbb{K}[\boldsymbol{x}][x_0]_d)^{\binom{n+d}{d}\times 1}$ be a vector whose elements are the monomials in $\mathbb{K}[\boldsymbol{x}][x_0]_d$ sorted in decreasing order with respect to $<$. Consider a matrix $\widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f})$ corresponding to a Macaulay matrix $\mathcal{M}_{>,d}(\boldsymbol{f})$ on which we performed linear algebra operations on its rows. The elements of the vector $\widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f}) \cdot v(d)$ correspond to polynomials in $\langle \boldsymbol{f} \rangle_{\boldsymbol{d}}$. Hence, we say that the rows of $\widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f})$ *represent* polynomials in $\langle \boldsymbol{f} \rangle_d$.

Lazard's approach to compute a Gröbner basis of $\boldsymbol{f}$ with respect to $>$ [Laz83] is to compute it from the matrices corresponding to the *row echelon form* of $\mathcal{M}_{>,1}(\boldsymbol{f}), \ldots, \mathcal{M}_{>,d_{reg}}(\boldsymbol{f})$, say $\widetilde{\mathcal{M}}_{>,1}(\boldsymbol{f}), \ldots, \widetilde{\mathcal{M}}_{>,d_{reg}}(\boldsymbol{f})$, where d_{reg} is the degree of regularity of $\boldsymbol{f}$ (Thm. 4.4.2). As this strategy constructs the Gröbner bases incrementally by computing d-Gröbner bases, we say that is follows a *degree-by-degree* strategy. For each d, let P_d be the set of non-zero polynomials represented by the rows of $\widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f})$. As we sorted the columns of $\mathcal{M}_{>,d}(\boldsymbol{f})$ and $\widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f})$ in decreasing order with respect to $>$, the set P_d is a maximal subset of $\boldsymbol{f}_d$ whose elements have different leading monomials. Hence, the polynomials of degree d in a d-Gröbner basis of $\boldsymbol{f}$ are contained in this maximal subset. We are not going to introduce linear algebra concepts as (reduced) row echelon forms or Gaussian elimination; we refer the reader to [Mey08, Ch. 2].

Theorem 4.4.4. *[Laz83, Sec. III.B] Consider a homogeneous system $\boldsymbol{f}$ and let G_{d-1} be a $(d-1)$-Gröbner basis for $\boldsymbol{f}$ with respect to a graded monomial ordering $>$. Let $\widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f})$ be the reduced row echelon form of the Macaulay matrix $\mathcal{M}_{>,d}(\boldsymbol{f})$. Let P_d be the set of non-zero polynomials represented by the rows of $\widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f})$. Let G_d be the union of G_{d-1} and the polynomials in P_d whose leading monomials are not divisible by the leading monomials of the polynomials in G_{d-1}, that is,*

$$G_d := G_{d-1} \cup \{g \in P_d \text{ such that } (\forall h \in G_{d-1})\, \mathtt{LM}_>(h) \text{ does not divide } \mathtt{LM}_>(g)\}$$

Then, G_d is a d-Gröbner basis of $\boldsymbol{f}$ with respect to $>$.

Hence, for every ideal $I \subset \mathbb{K}[\boldsymbol{x}][x_0]$ and a monomial order $>$, if we know the degree of regularity of I with respect to $>$ (Thm. 4.4.2) we can compute its Gröbner basis using linear algebra.

Example 4.4.5. *We want to compute a Gröbner basis for a homogeneous ideal generated by $\boldsymbol{f} := \{x + y + 2\,z, x, 3\,x^2 + z\,y + z^2\} \in \mathbb{K}[x, y, z]$, with respect* $\mathrm{GRevLex}(x > y > z)$. *As we will explain in Prop. 4.5.4, the degree of regularity of the ideal $\langle \boldsymbol{f} \rangle$ with respect to this monomial ordering is* 2.

Algorithm 3 Lazard's algorithm

Input: A list of homogeneous polynomials $\boldsymbol{f} := (f_1, \ldots, f_r)$, a monomial ordering $>$, and a degree d.
Output: The set G is a d-Gröbner basis of $\langle \boldsymbol{f} \rangle$ with respect to $>$.
If d is the *degree of regularity*, then G is a Gröbner basis of $\langle \boldsymbol{f} \rangle$.

1: $G_0 \leftarrow \{\}$
2: **for** d **from** 1 **to** d **do**
3: $\quad \mathcal{M}_{>,d}(\boldsymbol{f}) \leftarrow$ Macaulay matrix for $\boldsymbol{f}$ with respect to $>$ at degree d (Def. 4.4.3).
4: $\quad \widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f}) \leftarrow$ Reduced row echelon form of $\mathcal{M}_{>,d}(\boldsymbol{f})$.
5: $\quad P_d \leftarrow$ Non-zero polynomials represented by the rows of $\widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f})$.
6: $\quad G_d \leftarrow G_{d-1} \cup \{g \in P_d : (\forall h \in G_{d-1})\, \mathtt{LM}_>(h) \text{ does not divide } \mathtt{LM}_>(g)\}$.
7: **end for**
8: **return** G_d.

We consider the Macaulay matrix $\mathcal{M}_{>,1}(\boldsymbol{f})$ and its reduced row echelon form $\widetilde{\mathcal{M}}_{>,1}(\boldsymbol{f})$,

$$\mathcal{M}_{>,1}(\boldsymbol{f}) = \begin{array}{ccc|c} x & y & z & \\ \hline 1 & 1 & 2 & (1,1) \\ 1 & 0 & 0 & (2,1) \end{array} \longrightarrow \widetilde{\mathcal{M}}_{>,1}(\boldsymbol{f}) = \begin{array}{ccc} x & y & z \\ \hline 1 & 0 & 0 \\ 0 & 1 & 2 \end{array} \tag{4.1}$$

Hence, the 1-Gröbner basis of $\langle \boldsymbol{f} \rangle$ is $G_1 := \{x, y + 2\,z\}$.

Then, we consider the Macaulay matrix $\mathcal{M}_{>,2}(\boldsymbol{f})$ and its reduced row echelon form $\widetilde{\mathcal{M}}_{>,2}(\boldsymbol{f})$,

$$\mathcal{M}_{>,2}(\boldsymbol{f}) = \begin{array}{cccccc|c} x^2 & x\,y & y^2 & x\,z & y\,z & z^2 & \\ \hline 0 & 0 & 0 & 1 & 1 & 2 & (1,z) \\ 0 & 1 & 1 & 0 & 2 & 0 & (1,y) \\ 1 & 1 & 0 & 1 & 0 & 0 & (1,x) \\ 0 & 0 & 0 & 1 & 0 & 0 & (2,z) \\ 0 & 1 & 0 & 0 & 0 & 0 & (2,y) \\ 1 & 0 & 0 & 0 & 0 & 0 & (2,x) \\ 3 & 0 & 0 & 0 & 1 & 1 & (3,1) \end{array}$$

$$\downarrow$$

$$\widetilde{\mathcal{M}}_{>,2}(\boldsymbol{f}) = \begin{array}{cccccc} x^2 & x\,y & y^2 & x\,z & y\,z & z^2 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \tag{4.2}$$

Hence, the 2-Gröbner basis of $\langle \boldsymbol{f} \rangle$ is $G_2 := \{x, y + 2\,z, z^2\}$. As 2 is the degree of regularity of $\langle \boldsymbol{f} \rangle$ with respect to $>$, then $G_2 := \{x, y + 2\,z, z^2\}$ is the Gröbner basis of $\langle \boldsymbol{f} \rangle$ with respect to $>$. Note that the last row of $\widetilde{\mathcal{M}}_{>,2}(\boldsymbol{f})$ equals zero, and so, the rows of $\mathcal{M}_{>,2}(\boldsymbol{f})$ are not linearly independent.

As we observed in the previous example, Lazard's algorithm (Alg. 3) has a computational disadvantage: many of the rows reduce to zero in the Gaussian elimination step. So, we perform redundant computations when we reduce these rows. This phenomena is not exclusive of Lazard's approach. For example, in Buchberger's algorithm (Alg. 2), we perform redundant computations when we execute the division algorithm to reduce a S-polynomial whose remainder is zero. In the context of Gröbner bases, when we have an algorithm which, as part of its intermediate steps, consider polynomials which, after performing some computations, reduce to zero we say that the algorithm performs *reductions to zero*.

Reductions to zero have a strong impact in the practical efficiency of the algorithms to compute Gröbner bases, and so it is really important to try to avoid as much as them as we can. There are several ways to predict reductions to zero, see for example [Buc79, GM88, Tra89, Fau02, Mor03, EF17]. In this thesis, we concentrate in the `F5` criterion [Fau02].

As the ring $\mathbb{K}[\boldsymbol{x}][x_0]$ is commutative, many reductions to zero appearing in our computations come from the identity $f\,g - g\,f = 0$. The idea of the `F5` criterion is to use the signatures of the rows of the Macaulay matrix to prevent these reductions. It turns out that, when our system is a regular sequence (Sec. 2.7), the `F5` criterion identifies every reduction to zero appearing in Lazard's algorithm (Alg. 3). The `F5` criterion is a particular, and the first, example of the so-called signature-based algorithms for computing Gröbner bases. We are only going to discuss a simplified version of `F5`, applied to the *degree-by-degree* strategy of Lazard's algorithm, known as `Matrix-F5` [BFS15]. In the language of signature-based algorithms, this algorithm uses a $<_{\text{d-pot}}$ module monomial ordering. We refer the reader to the chapter [CLO15, Ch. 10.4], the PhD thesis [Ede12] and the survey [EF17] for more information about this family of algorithms.

Proposition 4.4.6 (`Matrix-F5` criterion)**.** *[BFS15, Prop. 8] Consider a system $\boldsymbol{f} := \{f_1, \ldots, f_r\}$ given by homogeneous polynomials of degree $d_1, \ldots, d_r$, respectively. Let $(i, \boldsymbol{x}^{\boldsymbol{\alpha}})$ be the signature of a row in $\mathcal{M}_{>,d}(\boldsymbol{f})$ such that $\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathtt{LM}_{>}(\langle f_1, \ldots, f_{i-1}\rangle)$, then this row is a linear combination of the rows in $\mathcal{M}_{>,d}(\boldsymbol{f})$ with smaller signature, see Def. 4.4.3.*

Observation 4.4.7. *As the polynomials are homogeneous, we can check if $\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathtt{LM}_{>}(f_1, \ldots, f_{i-1})$ by computing a $(d - d_i)$-Gröbner basis of $\langle f_1, \ldots, f_{i-1}\rangle$.*

The `Matrix-F5` algorithm (Alg. 4) avoids every reduction to zero when the input list of homogeneous polynomials forms a regular sequence.

Proposition 4.4.8. *[BFS15, Prop. 9] If $(f_1, \ldots, f_r)$ is a homogeneous regular sequence, then the* `Matrix-F5` *algorithm (Alg. 4) performs no reductions to zero.*

The `F5` criterion and the Koszul complex As we mentioned before, each Macaulay matrix $\mathcal{M}_{>,d}(\boldsymbol{f})$ represents a Sylvester map, see Rem. 3.2.11, and so the first map δ_1 of the strand of the Koszul complex $\mathcal{K}(\boldsymbol{f})_d$. Each row in the Macaulay matrix can be thought as an element in the image of the map δ_1. More specifically, each row obtained from the matrix by performing linear algebra operations on its rows represents the image of δ_1 at an element of $\bigoplus_i \mathbb{K}[\boldsymbol{x}][x_0]_{d-d_i}$. Hence, when we have a reduction to zero, that is, a linear combination of rows resulting in zero, we are considering an element $\boldsymbol{g} \in \bigoplus_i \mathbb{K}[\boldsymbol{x}][x_0]_{d-d_i}$ such that $\delta_1(\boldsymbol{g}) = 0$. Therefore, the reductions to zero appearing in the Gaussian elimination step of the previous algorithms correspond to elements in the first module of syzygies of the map δ_1. As the Koszul complex is a complex, the image of the map δ_2 of $\mathcal{K}(\boldsymbol{f})_d$ belongs to the

Algorithm 4 `Matrix-F5`

Input: A list of homogeneous polynomials $\boldsymbol{f} := (f_1, \ldots, f_r)$ of degrees $d_1, \ldots, d_r$, respectively, a graded monomial ordering $>$, and the degree of regularity d_{reg} of $\langle \boldsymbol{f} \rangle$ with respect to $>$ (Thm. 4.4.2).

Output: The set G is a Gröbner basis of $\langle \boldsymbol{f} \rangle$ with respect to $>$.

$G_{0,0}, \ldots, G_{0,d_{reg}} \leftarrow \{\}, \ldots, \{\}$.
for d **from** 1 **to** d_{reg} **do**
  for i **from** 1 **to** r **do**
    $\mathcal{M}_{>,d}(f_1, \ldots, f_i) \leftarrow$ Macaulay matrix for $(f_1, \ldots, f_i)$ with respect to $>$ at degree d, see Def. 4.4.3.
    $\mathcal{M}^{F5}_{>,d}(f_1, \ldots, f_i) \leftarrow$ Empty matrix.
    for each $(j, \boldsymbol{x}^{\boldsymbol{\alpha}})$ signature of $\mathcal{M}_{>,d}(f_1, \ldots, f_i)$ **do**
      if $d_j > d$ **or** for all $g \in G_{j-1,d-d_j}$, $\mathtt{LM}_>(g)$ does not divide $\boldsymbol{x}^{\boldsymbol{\alpha}}$ **then**
        Add to $\mathcal{M}^{F5}_{>,d}(f_1, \ldots, f_i)$ the row from $\mathcal{M}_{>,d}(f_1, \ldots, f_i)$ with signature $(j, \boldsymbol{x}^{\boldsymbol{\alpha}})$.
      end if
    end for
    $\widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f}) \leftarrow$ Reduced row echelon form of $\mathcal{M}^{F5}_{>,d}(\boldsymbol{f})$.
    $P_{i,d} \leftarrow$ Non-zero polynomials in the rows of $\widetilde{\mathcal{M}}_{>,d}(\boldsymbol{f})$.
    $G_{i,d} \leftarrow G_{i,d-1} \cup \{g \in P_{i,d} : (\forall h \in G_{i,d-1})\, \mathtt{LM}_>(h)$ does not divide $\mathtt{LM}_>(g)\}$.
  end for
end for
return $G_{r,d_{reg}}$.

kernel of δ_1. The syzygies generated by δ_2 are called *trivial syzygies* or *Koszul syzygies*. It turns out that, when there are non-trivial syzygies, the `F5` criterion fails to predict all the syzygies. Hence, the `F5` criterion avoids every reduction to zero only when the Koszul complex of the system is exact, and so, by Prop. 2.7.8, only when the system is a regular sequence.

4.5 Complexity

In this section we study the arithmetic complexity of computing Gröbner bases. Our complexity model is the standard arithmetical model, see [GG13, Ch. 2]. We relate the complexity of computing a Gröbner basis for a homogeneous ideal with respect to to a $\mathrm{GRevLex}$ monomial ordering with the Castelnuovo-Mumford regularity (Def. 2.8.3) of its initial ideal (Def. 4.1.6). Under some assumptions, this regularity is the same as the one of the original ideal.

To bound the complexity of computing Gröbner bases, we study the `Matrix-F5` algorithm (Alg. 4). There are two aspects to consider,

- The size and structure of the matrices on which we have to perform Gaussian elimination.
- The maximal degree up to which we have to compute these matrices, that is, the *degree of regularity* (Thm. 4.4.2).

In this section we will only focus on the second aspect. About the first aspect, we will just consider it superficially. We do not consider the structure of the Macaulay matrices and we bound the arithmetic complexity of the Gaussian elimination step by a general worst-case bound, see [Sto00]. Nevertheless, we emphasize that this aspect is extremely important from theoretical and practical reasons:

- Complexity-wise, in [BFS15] the authors proved that asymptotically, under regularity assumptions, the complexity of (a slightly modified version of) `Matrix-F5` (Alg. 4) is lower than what we present in this section.
- From a practical point of view, the library `FGb` [Fau10], Maple's state-of-the-art software to compute Gröbner bases, uses a dedicated library to exploit the structure of the matrices in the Gaussian elimination step, see [BEF$^+$16].

When we know the degree of regularity of a homogeneous ideal with respect to a graded monomial ordering, we can estimate the complexity of Lazard's algorithm (Alg. 3).

Proposition 4.5.1. *[BFS15, Prop. 1] Consider a homogeneous system $\boldsymbol{f} := (f_1, \ldots, f_r)$ and a graded monomial ordering $>$. Let d_{reg} be the degree of regularity of $\langle \boldsymbol{f} \rangle \subset \mathbb{K}[\boldsymbol{x}][x_0]$ with respect to $>$, see Thm. 4.4.2. Then, the complexity of computing a Gröbner basis of $\boldsymbol{f}$ with respect to $>$ using Lazard's algorithm (Alg. 3), or* `Matrix-F5` *(Alg. 4), is upper bounded by $\mathcal{O}(r\, d_{reg} \binom{n+d_{reg}}{n}^{\omega})$ arithmetic operations, where $\omega < 2.38$ is the* exponent of matrix multiplication *[GG13, Sec. 12.1].*

Let d_{reg} be the degree of regularity of an ideal I with respect to a graded monomial ordering $>$. Let $G_{d_{reg}-1}$, respectively $G_{d_{reg}}$, be the $(d_{reg}-1)$-Gröbner basis, respectively d_{reg}-Gröbner basis, of I.

By definition of degree of regularity (Thm. 4.4.2), $G_{d_{reg}}$ is a Gröbner basis of I with respect to $>$, but $G_{d_{reg}-1}$ it is not. By definition of Gröbner basis (Def. 4.1.7), it holds,

$$\mathtt{LM}_{>}(I) \neq \langle \{\mathtt{LM}_{>}(g) : g \in G_{d_{reg}-1}\}\rangle \quad \text{but} \quad \mathtt{LM}_{>}(I) = \langle \{\mathtt{LM}_{>}(g) : g \in G_{d_{reg}}\}\rangle.$$

Hence, the degree of regularity of I with respect to $>$ is the maximal degree of a minimal set of generators of the ideal $\mathtt{LM}_{>}(I)$, and so, it is bounded by the Castelnuovo-Mumford regularity of the initial ideal $\mathtt{LM}_{>}(I)$ (Def. 2.8.3).

Proposition 4.5.2. *[BM92, Sec. 3] Consider a homogeneous system $\boldsymbol{f} := (f_1, \ldots, f_r)$ and a graded monomial ordering $>$. Let d_{reg} be the degree of regularity of $\langle \boldsymbol{f}\rangle \subset \mathbb{K}[\boldsymbol{x}][x_0]$ with respect to $>$ (Thm. 4.4.2). Then,*

$$d_{reg} \leq reg_{\mathfrak{CM}}(\mathtt{LM}_{>}(I)).$$

Corollary 4.5.3. *The complexity of computing a Gröbner bases for a homogeneous system $\boldsymbol{f} := (f_1, \ldots, f_r)$ with respect to a graded monomial ordering $>$ is upper bounded by*

$$\mathcal{O}\left(r\left(reg_{\mathfrak{CM}}(\mathtt{LM}_{>}(I))\right)\binom{n + (reg_{\mathfrak{CM}}(\mathtt{LM}_{>}(I)))}{n}^{\omega}\right) \text{ arithmetic operations,}$$

where $\omega < 2.38$ is the exponent of matrix multiplication *[GG13, Sec. 12.1].*

For a further discussion about the bounds on the degree of regularity, we refer to [BM92, Sec. 3] and [Cha03].

When we consider Gröbner bases with respect to GRevLex monomial orderings, under some assumptions, we can replace, in Cor. 4.5.3, the Castelnuovo-Mumford regularity of the initial ideal by the regularity of the ideal.

Proposition 4.5.4 (Zero-dimensional case)**.** *[Cha03, Cor. 3] Consider a homogeneous ideal $I \subset \mathbb{K}[\boldsymbol{x}][x_0]$ such that $\dim(\mathbb{V}_{\mathbb{P}^n}(I)) = 0$, that is, $\dim_{Krull}(\mathbb{K}[\boldsymbol{x}][x_0]/I) = 1$, see Prop. 2.5.17. Then, the Castelnuovo-Mumford regularity of the initial ideal of I with respect to a* GRevLex *monomial ordering is the same as the regularity of the ideal, that is,*

$$reg_{\mathfrak{CM}}(\mathtt{LM}_{>_{\mathrm{GRevLex}}}(I)) = reg_{\mathfrak{CM}}(I).$$

Corollary 4.5.5. *The complexity of computing a Gröbner bases for a homogeneous system $\boldsymbol{f} := (f_1, \ldots, f_r)$ with respect to a graded monomial ordering $>$ such that $\dim(\mathbb{V}_{\mathbb{P}^n}(\langle \boldsymbol{f}\rangle)) = 0$ is upper bounded by*

$$\mathcal{O}\left(r\left(reg_{\mathfrak{CM}}(I)\right)\binom{n + (reg_{\mathfrak{CM}}(I))}{n}^{\omega}\right) \text{ arithmetic operations.}$$

In particular, when we have a square system defined by a regular sequence, it describes a zero-dimensional projective variety Prop. 2.7.13. In this case, we can bound the complexity of computing a Gröbner basis with respect to GRevLex using the Macaulay bound, see Prop. 2.8.5.

Corollary 4.5.6 (Complexity of computing a Gröbner basis for square regular sequence)**.** *Consider a homogeneous square system $\boldsymbol{f} := (f_1, \ldots, f_n)$ in $\mathbb{K}[\boldsymbol{x}][x_0]$ given by regular sequence of degrees $d_1, \ldots, d_n$, respectively. Then, the complexity of computing a Gröbner basis for $\boldsymbol{f}$ with respect to* GRevLex *is upper bounded by*

$$\mathcal{O}\left(n \cdot \left(\sum_i d_i - n + 1\right) \binom{(\sum_i d_i) + 1}{n}^{\omega}\right) \text{ arithmetic operations.}$$

When the variety defined by our ideal is not zero dimensional, we need some extra assumptions to extend the previous result. These assumptions are related to *generic coordinates*, and their relation to GRevLex was studied by Galligo [Gal74, Gal79], Giusti [Giu84], and Bayer and Stillman [BS87b, BS87a], among others.

Definition 4.5.7 (Linear change of coordinates)**.** *A* linear change of coordinates *is an 0-graded automorphism $A : \mathbb{K}[\boldsymbol{x}][x_0] \to \mathbb{K}[\boldsymbol{x}][x_0]$, that is, an invertible morphism from $\mathbb{K}[\boldsymbol{x}][x_0]$ to itself.*

Proposition 4.5.8. *[BS87a, Prop. 2.11] Consider a homogeneous system $\boldsymbol{f} := (f_1, \ldots, f_r)$ and consider a generic linear change of coordinates $A : \mathbb{K}[\boldsymbol{x}][x_0] \to \mathbb{K}[\boldsymbol{x}][x_0]$, see Sec. 2.13. Consider the system $\bar{\boldsymbol{f}} := (f_1 \circ A, \ldots, f_r \circ A)$. Then, the Castelnuovo-Mumford regularity of the initial ideal of $\langle \bar{f} \rangle$ with respect to a* GRevLex *monomial ordering is the same as the regularity of the ideal, that is,*

$$reg_{\mathfrak{GM}}(\mathtt{LM}_{>_{\mathrm{GRevLex}}}(\langle \bar{f} \rangle)) = reg_{\mathfrak{GM}}(\langle \bar{f} \rangle).$$

4.6 Gröbner bases over semigroup algebras

We recall some definitions related to Gröbner bases over semigroup algebras from [FSS14]. Let S be a pointed affine semigroup (Def. 2.12.1) and consider its associated pointed semigroup algebra $\mathbb{K}[S]$ (Def. 2.12.7).

Definition 4.6.1 (Monomial order)**.** *A monomial order for a pointed semigroup algebra $\mathbb{K}[S]$, say $<$, is a total order for the monomials in $\mathbb{K}[S]$ such that:*

- *For any $\boldsymbol{\alpha} \in S \setminus \{\mathbf{0}\}$, it holds $\boldsymbol{X}^{\mathbf{0}} < \boldsymbol{X}^{\boldsymbol{\alpha}}$.*
- *For every $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \in S$, if $\boldsymbol{X}^{\boldsymbol{\alpha}} < \boldsymbol{X}^{\boldsymbol{\beta}}$ then $\boldsymbol{X}^{\boldsymbol{\alpha}+\boldsymbol{\gamma}} < \boldsymbol{X}^{\boldsymbol{\beta}+\boldsymbol{\gamma}}$.*

Observation 4.6.2. *Monomial orders always exist for pointed affine semigroups. To construct them, first we embed any pointed affine semigroup of dimension n in a pointed rational cone $\mathcal{C} \subset \mathbb{R}^n$ (Def. 2.12.4). Then, we choose n linearly independent forms $l_1, \ldots, l_n$ from the dual cone of $\mathcal{C}$, which is*

$$\{l : \mathbb{R}^n \to \mathbb{R} \mid \forall \boldsymbol{\alpha} \in \mathbb{C}, l(\boldsymbol{\alpha}) \geq 0\}.$$

We define the monomial order $<$ so that $\boldsymbol{X}^{\boldsymbol{\alpha}} < \boldsymbol{X}^{\boldsymbol{\beta}}$ if and only if there is a $k \leq n$ such that for all $i < k$ it holds $l_i(\boldsymbol{\alpha}) = l_i(\boldsymbol{\beta})$ and $l_k(\boldsymbol{\alpha}) < l_k(\boldsymbol{\beta})$.

If our semigroup algebra is not pointed, we can not define monomial orders.

Definition 4.6.3 (Leading monomial)**.** *Given a monomial order $<$ for a pointed affine semigroup algebra $\mathbb{K}[S]$ and a polynomial $f \in \mathbb{K}[S]$, its leading monomial, $\mathtt{LM}_<(f)$, is the biggest monomial appearing in f with respect to the monomial order $<$.*

Note that the exponent of the leading monomial of f always corresponds to a vertex of the Newton polytope of f (Def. 2.12.17).

Definition 4.6.4 (Gröbner basis)**.** *Let $\mathbb{K}[S]$ be a pointed affine semigroup algebra and consider a monomial order $<$ for $\mathbb{K}[S]$. For an ideal $I \subset \mathbb{K}[S]$, a set $G \subset I$ is a Gröbner basis of I if*

$$\langle \{\mathtt{LM}_<(g) : g \in G\}\rangle = \langle \{\mathtt{LM}_<(f) : f \in I\}\rangle.$$

In other words, if for every $f \in I$, there is $g \in G$ and $\boldsymbol{X}^\alpha \in \mathbb{K}[S]$ such that $\mathtt{LM}_<(f) = \boldsymbol{X}^\alpha \, \mathtt{LM}_<(g)$.

Proposition 4.6.5. *As S is finitely generated, the algebra $\mathbb{K}[S]$ is a Noetherian ring [Gil84, Thm. 7.7]. Hence, for any monomial order and any ideal in $\mathbb{K}[S]$, there is a finite Gröbner basis.*

Chapter 5

Solving polynomial systems

In this chapter, we discuss different strategies to solve affine systems of polynomial equations having a finite set of solutions, that is, systems that define a zero-dimensional variety. We present two different strategies:

- Solving systems by extending the solutions of a univariate polynomial.
- Solving systems by inverting a monomial map.

For the first strategy, we compute a Gröbner basis with respect to a lexicographical monomial ordering (Lex). For the second one, we compute the eigenvectors of a matrix associated to the system, this is the multiplication map. The results of this chapter come mainly from [CLO06, Ch. 2].

We will not discuss other techniques to solve polynomial systems, and we refer the reader to [DE05] and references there in.

5.1 Lexicographical monomial orderings (Lex)

In this section, we study some properties of the Gröbner bases related to lexicographical monomial orderings (Lex), see Def. 4.1.3. These properties relate Lex to elimination theory. The results of this section come mainly from [CLO15, Ch. 3] and [CLO06, Sec. 2.1].

Proposition 5.1.1 (Elimination property of Lex)**.** *[CLO15, Thm. 3.1.2] Let $>_{\mathrm{Lex}}$ be a lexicographical monomial ordering such that $x_1 >_{\mathrm{Lex}} \cdots >_{\mathrm{Lex}} x_n$ and G a Gröbner basis of an ideal I with respect to this order. Then, for each $1 \leq k \leq n$, $G \cap \mathbb{K}[x_k, \ldots, x_n]$ is a Gröbner basis of the ideal $I \cap \mathbb{K}[x_k, \ldots, x_n]$ with respect to same order.*

Theorem 5.1.2 (Extension theorem)**.** *[CLO15, Thm. 3.1.3] Consider the ideal $I := \langle f_1, \ldots, f_r \rangle \subset \mathbb{K}[\boldsymbol{x}]$ and $I \cap \mathbb{K}[x_2, \ldots, x_n]$. For each $1 \leq i \leq r$, we write f_i as,*

$$f_i = c_i(x_2, \ldots, x_n) x_1^{N_i} + \text{ terms in which } x_1 \text{ has degree} < N_i.$$

where $N_i > 0$ and $c_i \in \mathbb{K}[x_2, \ldots, x_n]$ is non-zero.

Suppose that there is a partial solution of $I \cap \mathbb{K}[x_2, \ldots, x_n]$, that is, a point $(a_2, \ldots, a_n) \in \mathbb{V}_{\mathbb{K}^{n-1}}(I \cap \mathbb{K}[x_2, \ldots, x_n])$. If $(a_2, \ldots, a_n) \notin \mathbb{V}_{\mathbb{K}^{n-1}}(\langle c_1, \ldots, c_r \rangle)$, that is, if all the c_i do not vanish simultaneously at the point, then there is an $a_1 \in \mathbb{K}$ such that $(a_1, a_2, \ldots, a_n) \in \mathbb{V}_{\mathbb{K}^n}(I)$, that is, there is a solution for the original system.

Observation 5.1.3. *For each partial solution $(a_2, \ldots, a_n) \in \mathbb{K}^{n-1}$ that extends to a solution $(a_1, \ldots, a_n) \in \mathbb{K}^n$, we can recover the value a_1 as the solution of the system of univariate polynomials.*

$$\begin{cases} f_1(x_1, a_2, \ldots, a_n) = 0 \\ \qquad\qquad \vdots \\ f_r(x_1, a_2, \ldots, a_n) = 0 \end{cases}$$

Solving a system of univariate polynomials is equivalent to solve the univariate polynomial given by the Greatest Common Divisor *of the generators of the system (recall that $\mathbb{K}[x_1]$ is a principal ideal, see Def. 2.2.8). Hence, the extension theorem tell us that, when we have a partial solution such that the coefficients $c_i(x_2, \ldots, x_n)$ do not vanish all at the same time, then the* Greatest Common Divisor *of the partial evaluation of the generators of the system is not* 1.

When we have a zero-dimensional system, the Gröbner basis of the ideal contains a univariate polynomial. Hence, we can solve this polynomial and, recursively, use the extension theorem to recover a solution of the original system.

Proposition 5.1.4. *Consider the variety $V = \mathbb{V}_{\mathbb{K}^n}(f_1, \ldots, f_r)$ such that V is zero-dimensional. The reduced Gröbner basis of I with respect to a lexicographical order, such that $x_1 > \cdots > x_n$, can be written as a disjoint union of sets $G_1, \ldots, G_n$ such that, for $i < n$, $G_i \subset \mathbb{K}[x_i, \ldots, x_n]$ and $G_i \not\subset \mathbb{K}[x_{i-1}, \ldots, x_n]$, and $G_n = \{g_n\} \subset \mathbb{K}[x_n]$.*

Example 5.1.5. *We want to solve the system of (f_1, f_2), where*

$$\begin{cases} f_1(x_1, x_2) := x_1^2 + x_2^2 - 2 \\ f_2(x_1, x_2) := 2\,x_1\,x_2 - 1. \end{cases}$$

The Gröbner basis of $\langle f_1, f_2 \rangle$ with respect to $>_{\text{Lex}}$, such that $x_1 >_{\text{Lex}} x_2$, is $\{4\,{x_2}^4 - 8\,{x_2}^2 + 1, 2\,{x_2}^3 - 4\,x_2 + x_1\}$. The univariate polynomial $4\,{x_2}^4 - 8\,{x_2}^2 + 1$ has four solutions:

$$a_2^{(1)} = \frac{1+\sqrt{3}}{2}, \quad a_2^{(2)} = \frac{1-\sqrt{3}}{2}, \quad a_2^{(3)} = \frac{-1+\sqrt{3}}{2}, \quad a_2^{(4)} = \frac{-1-\sqrt{3}}{2}$$

As the coefficient of x_1 in $2\,{x_2}^3 - 4\,x_2 + x_1$ is a constant, we can extend all four solutions. For example, the solution $a_2^{(1)} \in \mathbb{V}_{\mathbb{K}}(4\,{x_2}^4 - 8\,{x_2}^2 + 1)$ extends to a solution $(a_1^{(1)}, a_2^{(1)}) \in \mathbb{V}_{\mathbb{K}^2}(f_1, f_2)$ where $(a_1^{(1)}, a_2^{(1)}) = (\frac{-1+\sqrt{3}}{2}, \frac{1+\sqrt{3}}{2})$.

In the previous example, we could apply the Extension theorem because the coefficients of the monomials $x_1^{N_i}$ in $c_i(x_2)$, see Ex. 5.1.5, do not vanish all at the same time, as some of them are constants. In particular, this ideal is in *shape position*.

Definition 5.1.6 (Shape position)**.** *We say that an ideal is in* shape position *when its reduced Gröbner basis with respect to a lexicographical monomial ordering* $>_{\text{Lex}}$*, such that* $x_1 >_{\text{Lex}} \cdots >_{\text{Lex}} x_n$*, is*

$$\left\{ \begin{array}{c} g_1(x_n) - x_1 \\ \vdots \\ g_{n-1}(x_n) - x_{n-1} \\ g_n(x_n) \end{array} \right\}.$$

Theorem 5.1.7. *Let* $(f_1, \ldots, f_r)$ *be a radical zero-dimensional system. Consider* A *to be a generic linear change of coordinates, see Def. 4.5.7 and Sec. 2.13. Then, the reduced Gröbner basis of* $(f_1 \circ A, \ldots, f_r \circ A)$ *with respect to the lexicographical monomial ordering* $>_{\text{Lex}}$*, such that* $x_1 >_{\text{Lex}} \cdots >_{\text{Lex}} x_n$*, is in* shape position*.*

For a proof of a stronger version of the previous theorem, see [BMMT94].

We can solve an ideal in *shape position* by solving a univariate polynomial $g_n \in \mathbb{K}[x_n]$. The solution set of an ideal I in shape position is given by the parameterization

$$\{(g_1(\boldsymbol{p}), \ldots, g_{n-1}(\boldsymbol{p}), \boldsymbol{p}) : \boldsymbol{p} \in \mathbb{V}_{\mathbb{K}}(g_n)\}.$$

5.2 Quotient rings and multiplication maps

Following Prop. 2.2.19, if I is an ideal such that $\mathbb{K}[\boldsymbol{x}]/I$ is zero-dimensional, then the quotient ring $\mathbb{K}[\boldsymbol{x}]/I$ is a finite dimensional $\mathbb{K}$-vector space. In this section, we use this finite dimensional $\mathbb{K}$-vector space to solve polynomial systems using linear algebra techniques, in particular eigenvalues and eigenvectors computations.

In what follows, we consider a fixed ideal $I \subset \mathbb{K}[\boldsymbol{x}]$ such that the quotient ring $\mathbb{K}[\boldsymbol{x}]/I$ is zero-dimensional. Let D be the dimension, as a finite dimensional $\mathbb{K}$-vector space, of $\mathbb{K}[\boldsymbol{x}]/I$, that is,

$$D := \dim_{\mathbb{K}}(\mathbb{K}[\boldsymbol{x}]/I).$$

Proposition 5.2.1 (Monomial basis)**.** *[CLO06, Pg. 38] There is a set of monomials* $\{\boldsymbol{x}^{\boldsymbol{\alpha_1}}, \ldots, \boldsymbol{x}^{\boldsymbol{\alpha_D}}\}$ *that is a basis of the* $\mathbb{K}$*-vector space* $\mathbb{K}[\boldsymbol{x}]/I$*. This set is not unique and we call it* monomial basis*.*

Definition 5.2.2 (Multiplication map)**.** *For each* $f \in \mathbb{K}[\boldsymbol{x}]$*, we define the multiplication map* $*_f : \mathbb{K}[\boldsymbol{x}]/I \to \mathbb{K}[\boldsymbol{x}]/I$ *as the map that multiplies* $g \in \mathbb{K}[\boldsymbol{x}]/I$ *by* f*. If for a fixed monomial basis* $\boldsymbol{x}^{\boldsymbol{\alpha_1}}, \ldots, \boldsymbol{x}^{\boldsymbol{\alpha_D}}$ *of* $\mathbb{K}[\boldsymbol{x}]/I$*, let* $\mathrm{M}_f \in \mathbb{K}^{D \times D}$ *be the matrix that represents* $*_f$ *in this basis.*

In the following, we fix the monomial basis of $\mathbb{K}[\boldsymbol{x}]/I$, say $\{\boldsymbol{x}^{\boldsymbol{\alpha_1}}, \ldots, \boldsymbol{x}^{\boldsymbol{\alpha_D}}\}$.

5.2.1 Eigenvalues and eigenvectors of the multiplication maps

We can recover information of the solutions in $\mathbb{V}_{\mathbb{K}^n}(I)$ by performing eigenvalue and eigenvector computations with multiplication maps. Our presentation follows [Cox05, Sec. 2.1] and [CLO06, Sec. 2.4].

Proposition 5.2.3 (Eigenvalue criterion)**.** *[Cox05, Thm. 2.1.4] The eigenvalues of the matrix M_f are the evaluations of f at the points of $\mathbb{V}_{\mathbb{K}^n}(I)$. That is,*

$$\lambda \text{ is a eigenvalue of } \mathrm{M}_f \iff \lambda = f(\alpha), \text{ for } \alpha \in \mathbb{V}_{\mathbb{K}^n}(I).$$

This criterion was proposed for the first time in [Laz81]. We generalize it in Sec. 6.2.

Proposition 5.2.4 (Eigenvector criterion)**.** *[Cox05, Thm. 2.1.4] For each $\boldsymbol{p} \in \mathbb{V}_{\mathbb{K}^n}(I)$, the vector $(\boldsymbol{x}^{\boldsymbol{\alpha_1}}(\boldsymbol{p}), \ldots, \boldsymbol{x}^{\boldsymbol{\alpha_D}}(\boldsymbol{p})) \in \mathbb{K}^D$, that is, the vector of evaluations of the monomial basis at the solution $\boldsymbol{p}$, is a right eigenvector of M_f, associated to the eigenvalue $f(\boldsymbol{p})$.*

This criterion was proposed for the first time in [AS88]. We generalize it in Sec. 6.3.

When the matrix M_f is *non-derogatory*, that is, when the dimension of each eigenspace of M_f is 1, we can compute each eigenvector and recover the solutions of the system. For this, we need to invert a monomial map. If all the solutions of I are curvilinear [Cox05, Def. 2.1.9], for example this happens when I is radical, then, for generic choices of f, M_f is non-derogatory, see [Cox05, Thm. 2.1.11]. If this is not the case, then we have to apply more sophisticated techniques to recover the solution; see the discussion at the end of [Cox05, Sec. 2.1.3] and references there in.

5.2.2 Computing Gröbner bases using multiplication maps (FGLM)

Besides using the multiplication maps to solve a system by computing eigenvalues and eigenvectors, we can use them to compute Gröbner bases. We present the FGLM algorithm [FGLM93], which allow us to perform such a computation. We refer the reader to [CLO06, Sec. 2.3] for a more detailed presentation.

We can use multiplication maps to test if a polynomial belongs to an ideal.

Proposition 5.2.5. *[CLO06, Prop. 2.4.1] The map from $\mathbb{K}[\boldsymbol{x}]/I$ to $\mathrm{M}_f \in \mathbb{K}^{D\times D}$ that maps $f \in \mathbb{K}[\boldsymbol{x}]/I \mapsto \mathrm{M}_f \in \mathbb{K}^{D\times D}$ is an injective linear map. Hence, for every $f \in \mathbb{K}[\boldsymbol{x}]$, $f \in I \iff \mathrm{M}_f = 0$.*

Given the multiplication maps $\mathrm{M}_{x_1}, \ldots, \mathrm{M}_{x_n}$, we can compute any multiplication map M_f.

Proposition 5.2.6. *[CLO06, Prop. 2.4.2] Consider $f, g \in \mathbb{K}[\boldsymbol{x}]$ and $\lambda \in \mathbb{K}$, then*

$$\mathrm{M}_\lambda = \lambda\, Id, \quad \mathrm{M}_{f+g} = \mathrm{M}_f + \mathrm{M}_g \quad \text{and} \quad \mathrm{M}_{g\,f} = \mathrm{M}_g\, \mathrm{M}_f.$$

In particular,

$$\mathrm{M}_f = f(\mathrm{M}_{x_1}, \ldots, \mathrm{M}_{x_n}).$$

The idea of the FGLM algorithm is to use the multiplication maps to construct incrementally polynomials in the ideal. By doing this incremental construction following a monomial ordering, we obtain a Gröbner basis. We emphasize that this algorithm terminates because the quotient ring is zero-dimensional.

Theorem 5.2.7. *[CLO06, Sec. 2.3] Let $\mathrm{M}_{x_1}, \ldots, \mathrm{M}_{x_n}$ be the multiplication maps of the quotient ring $\mathbb{K}[\boldsymbol{x}]/I$ with respect to a monomial basis $\{\boldsymbol{x}^{\boldsymbol{\alpha_1}}, \ldots, \boldsymbol{x}^{\boldsymbol{\alpha_D}}\}$. Let $e \in \mathbb{K}^D$ be a vector representing the element $1 \in \mathbb{K}[\boldsymbol{x}]/I$. Then, for any monomial ordering $>$,* `FGLM` *(Alg. 5) always terminates and returns a Gröbner basis for I with respect to $>$.*

Algorithm 5 FGLM [FGLM93]

Input: Multiplication maps $\mathrm{M}_{x_1}, \ldots, \mathrm{M}_{x_n}$, an element $e \in \mathbb{K}^D$ representing $1 \in \mathbb{K}[\boldsymbol{x}]/I$ and a monomial ordering $>$.
Output: Gröbner basis of I with respect to $>$.

```
L ← {x^α ∈ K[x]}.
B ← {(e, 1)} ⊂ K^D × K[x].
G ← ∅.
while L ≠ ∅ do
  x^α ← minimal element in L with respect to >.
  if ∃f ∈ K[x] s.t. (M_{x^α} e, f) belongs to the K-linear space generated by B then
    G ← G ∪ {x^α − f}.
    L ← L \ {x^β ∈ K[x] : x^α divides x^β}.
  else
    B ← B ∪ {(M_{x^α} e, x^α)}.
  end if
end while
return G.
```

The complexity of FGLM is polynomial in the number of solutions.

Proposition 5.2.8. *[FGLM93, Prop. 3.1] The arithmetic complexity of* FGLM *(Alg. 5) is upper bounded by* $\mathcal{O}(n\,D^3)$ *operations.*

The previous complexity bound can be improved [FGHR13] and, in practice, there are many efficient implementations of variants of this algorithm, see [FM17].

5.3 Computing multiplication maps

We can compute multiplication maps using Gröbner bases or, under genericity assumptions, using the formulas to compute resultants from Sec. 3.2.1.

5.3.1 Multiplication maps through Gröbner bases

We can use the Gröbner basis of I to compute a monomial basis of the quotient ring $\mathbb{K}[\boldsymbol{x}]/I$. For this, we use normal forms, see Sec. 4.2. The normal form $NF_{>,I}(f)$ equals zero if and only if $f \in I$ (Cor. 4.2.2). The normal form of a monomial not contained in the initial ideal $\mathtt{LM}_>(I)$ (Def. 4.1.6) equals to itself, that is,

$$\text{For all } \boldsymbol{x}^{\boldsymbol{\alpha}} \notin \mathtt{LM}_>(I) \text{ it holds } NF_{>,I}(\boldsymbol{x}^{\boldsymbol{\alpha}}) = \boldsymbol{x}^{\boldsymbol{\alpha}}.$$

Hence, the monomials not belonging to the initial ideal of I with respect to $>$ are linearly independent over $\mathbb{K}[\boldsymbol{x}]/I$. Moreover, every polynomial $f \in \mathbb{K}[\boldsymbol{x}]$ is equivalent to a polynomial $g \in \mathbb{K}[\boldsymbol{x}]$ such that every monomials appearing in g does not belong to $\mathtt{LM}_>(I)$ (Prop. 4.2.1 and Cor. 4.2.2). Hence, the monomials not belonging to $\mathtt{LM}_>(I)$ form a monomial basis of $\mathbb{K}[\boldsymbol{x}]/I$. This monomial basis is called *standard basis*.

Proposition 5.3.1 (Standard basis)**.** *Let $I \subset \mathbb{K}[\boldsymbol{x}]$ be an ideal and $>$ a monomial ordering, then the set of monomials $B := \{\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}] : \boldsymbol{x}^{\boldsymbol{\alpha}} \notin \mathtt{LM}_{>}(I)\}$ form a monomial basis of $\mathbb{K}[\boldsymbol{x}]/I$. We say that the set B is a* standard basis *of $\mathbb{K}[\boldsymbol{x}]/I$.*

For every $f \in \mathbb{K}[\boldsymbol{x}]$, it holds $NF_{>,I}(f) \in \langle B\rangle_{\mathbb{K}}$, that is, the normal form of f belongs to the linear span of B. Hence, the coordinates of $f \in \mathbb{K}[\boldsymbol{x}]/I$ with respect to the standard basis B correspond to the coefficients of $NF_{>,I}(f)$.

We can use Gröbner bases to compute multiplication maps with respect to a standard basis.

Corollary 5.3.2. *Let $I \subset \mathbb{K}[\boldsymbol{x}]$ be an ideal such that $\mathbb{K}[\boldsymbol{x}]/I$ is zero-dimensional. Consider a monomial ordering $>$ and the standard basis $B := \{\boldsymbol{x}^{\boldsymbol{\alpha_1}}, \dots \boldsymbol{x}^{\boldsymbol{\alpha_D}}\}$ defined in Prop. 5.3.1. For each $f \in \mathbb{K}[\boldsymbol{x}]$, let M_f be the multiplication map of f in $\mathbb{K}[\boldsymbol{x}]/I$ with respect to the monomial basis B. Then, the element of M_f in the position (i, j) corresponds to the coefficient of $\boldsymbol{x}^{\boldsymbol{\alpha_i}}$ in the polynomial $NF_{>,I}(\boldsymbol{x}^{\boldsymbol{\alpha_j}} f)$,*

Example 5.3.3 (Continuation of Ex. 5.1.5)**.** *The set $G := \{4\,{x_2}^4 - 8\,{x_2}^2 + 1, 2\,{x_2}^3 - 4\,x_2 + x_1\}$ is the reduced Gröbner basis of the ideal I, with respect to $>_{\mathrm{Lex}}$, such that $x_1 >_{\mathrm{Lex}} x_2$. The monomials not belonging to the initial ideal of I with respect to $>_{\mathrm{Lex}}$ are $\{1, x_2, x_2^2, x_2^3\}$. They form a standard basis of $\mathbb{K}[\boldsymbol{x}]/I$.*

To compute M_{x_1}, we need to compute,

$$\begin{cases} NF_{>,I}(1 \cdot x_1) & = -2\,x_2^3 + 4\,x_2 \\ NF_{>,I}(x_2 \cdot x_1) & = \frac{1}{2} \\ NF_{>,I}(x_2^2 \cdot x_1) & = \frac{1}{2}\,x_2 \\ NF_{>,I}(x_2^3 \cdot x_1) & = \frac{1}{2}\,x_2^2 \end{cases}$$

Hence,

$$\mathrm{M}_{x_1} = \left[\begin{array}{c|cccc} 1 & 0 & \frac{1}{2} & 0 & 0 \\ x_2 & 4 & 0 & \frac{1}{2} & 0 \\ x_2^2 & 0 & 0 & 0 & \frac{1}{2} \\ x_2^3 & -2 & 0 & 0 & 0 \end{array}\right].$$

Note that $a_1^{(1)} = \frac{-1+\sqrt{3}}{2}$ is an eigenvalue of M_{x_1} and its associated eigenvector is

$$(1, a_2^{(1)}, (a_2^{(1)})^2, (a_2^{(1)})^3,) = \left(1, \frac{1+\sqrt{3}}{2}, \frac{(1+\sqrt{3})^2}{4}, \frac{(1+\sqrt{3})^3}{8}\right).$$

Change of ordering

The original motivation of the `FGLM` algorithm (Alg. 5) is to perform change of orderings for Gröbner bases, that is, to use an already computed Gröbner basis with respect to a monomial order $>_1$, to compute a Gröbner basis for the same ideal with respect to $>_2$. We can do so by using the Gröbner basis with respect to $>_1$ to compute multiplication maps as we detailed previously. In general, we follow this approach when it is easier to compute a Gröbner basis with respect to $>_1$ and then change the ordering, than computing directly the Gröbner basis with respect to $>_2$. In practice, we follow this approach to

solve efficiently polynomial systems. We compute a Gröbner basis with respect to a $\mathrm{GRevLex}$ ordering, whose complexity we know is low, see discussion in Sec. 4.5, and then we use the `FGLM` algorithm to recover a lexicographical Gröbner basis. Hence, using this approach, we can bound the complexity of solving generic square systems.

Proposition 5.3.4 (Complexity of solving generic square systems)**.** *Consider a system of affine polynomials given by $(f_1, \ldots, f_n)$ in $\mathbb{K}[\boldsymbol{x}]$ of degrees $d_1, \ldots, d_n$, respectively. Assume that*

- *$\mathbb{K}[\boldsymbol{x}]/\langle f_1, \ldots, f_n\rangle$ is zero-dimensional,*
- *the Bézout bound (Cor. 2.7.15) is tight, that is, $\mathbb{V}_{\mathbb{K}^n} f_1, \ldots, f_n$ contains different $\prod_i d_i$ points, and*
- *the sequence $(f_1^h, \ldots, f_n^h)$, that is, the homogenization of the polynomials over $\mathbb{K}[\boldsymbol{x}][x_0]$ (Def. 4.3.1), is a regular sequence.*

Then, we can compute a Gröbner basis for $\langle f_1, \ldots, f_n\rangle$ with respect to $>_{\mathrm{Lex}}$ in

$$\mathcal{O}\left(n \cdot \left(\sum_i d_i - n + 1\right) \binom{(\sum_i d_i) + 1}{n}^{\omega} + n \left(\prod_{i=1}^{r} d_i\right)^3\right) \text{ arithmetic operations.}$$

5.3.2 Multiplication maps through the Macaulay resultant formula

In this section, we present how to use the Macaulay resultant formula (Prop. 3.2.5) to compute multiplication maps for ideals generated by generic square systems, see Sec. 2.13. The results of this section come mainly from [CLO06, Sec. 3.6]. We mention that the techniques discussed in this section were successfully extended to the sparse case, see [ER94, PS96, Emi96].

We follow the same notation as in Sec. 3.2, but instead of working over $\mathbb{K}[\boldsymbol{x}]$, we work over $\mathbb{K}[\boldsymbol{x}][x_0]$. We fix degrees $d_0, d_1, \ldots, d_n$ and, mutatis mutandis from Def. 3.2.1, we introduce the *generic polynomial system* $\boldsymbol{F} := \{F_0, F_1, \ldots, F_n\} \subset \mathbb{Z}[\boldsymbol{u}][\boldsymbol{x}][x_0]$, where

$$F_i := \sum_{\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}][x_0]_{d_i}} u_{i,\alpha} \boldsymbol{x}^{\boldsymbol{\alpha}}.$$

We consider the Macaulay resultant matrix from Prop. 3.2.5 but we reorder the variables in the construction. We consider $x_1 > \cdots > x_n > x_0$. Hence, the sets $P_1, \ldots, P_n, P_0$ required for the construction of the matrix are as follows:

$$P_i := \begin{cases} \{\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}][x_0]_{\mathtt{MB}} \text{ s.t. } x_i^{d_i} | \boldsymbol{x}^{\boldsymbol{\alpha}} \text{ and } (\forall 1 \leq j < i)\ \boldsymbol{x}^{\boldsymbol{\alpha}} \notin P_j\} & \text{for } i > 0, \\ \{\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{K}[\boldsymbol{x}][x_0]_{\mathtt{MB}} \text{ s.t. } x_0^{d_0} | \boldsymbol{x}^{\boldsymbol{\alpha}} \text{ and } (\forall j > 0)\ \boldsymbol{x}^{\boldsymbol{\alpha}} \notin P_j\} & \text{for } i = 0, \end{cases}$$

where $\mathtt{MB} := \sum_{i=0}^{n} d_i - (n+1) + 1$ is the Macaulay bound, see Prop. 2.8.5.

We consider the Macaulay resultant matrix $\mathcal{M} \in \mathbb{Z}[\boldsymbol{u}]^{\dim_{\mathbb{K}}(\mathbb{K}[\boldsymbol{x}][x_0]_{\mathtt{MB}}) \times \dim_{\mathbb{K}}(\mathbb{K}[\boldsymbol{x}][x_0]_{\mathtt{MB}})}$ from Prop. 3.2.5 related to the partition $P_1, \ldots, P_n, P_0$, but we will consider a particular order for some of its columns and rows.

We fix some monomial ordering $>$ for $\mathbb{K}[\boldsymbol{x}][x_0]$ and write the Macaulay resultant matrix as

$$\mathcal{M} = \begin{bmatrix} \mathcal{M}_{1,1} & \mathcal{M}_{1,2} \\ \mathcal{M}_{2,1} & \mathcal{M}_{2,2} \end{bmatrix},$$

where

- Both the columns of $\begin{bmatrix} \mathcal{M}_{1,1} \\ \mathcal{M}_{2,1} \end{bmatrix}$ and the rows of $[\, \mathcal{M}_{1,1} \ \mathcal{M}_{1,2} \,]$ are indexed by the monomials in $P_1 \cup \cdots \cup P_n$ sorted in decreasing order with respect to $>$.
- Both the columns of $\begin{bmatrix} \mathcal{M}_{1,2} \\ \mathcal{M}_{2,2} \end{bmatrix}$ and the rows of $[\, \mathcal{M}_{2,1} \ \mathcal{M}_{2,2} \,]$ are indexed by the monomials in P_0 sorted in decreasing order with respect to $>$.
- For each element in the matrix $\mathcal{M}$, the element in the row indexed by $\boldsymbol{x}^{\boldsymbol{\alpha}}$ and column indexed by $\boldsymbol{x}^{\boldsymbol{\beta}}$ corresponds to the coefficient of the monomial $\boldsymbol{x}^{\boldsymbol{\beta}}$ in the polynomial $\frac{\boldsymbol{x}^{\boldsymbol{\alpha}}}{x_i^{d_i}} F_i$, where $\boldsymbol{x}^{\boldsymbol{\alpha}} \in P_i$.

Given an homogeneous system $\boldsymbol{f_0} := (f_1, \ldots, f_n, f_0)$ and a matrix $M \in \mathbb{Z}[\boldsymbol{u}]^{\mathcal{K} \times \mathcal{K}}$, we define $M(\boldsymbol{f_0}) \in \mathbb{K}^{\mathcal{K} \times \mathcal{K}}$ as the matrix that we obtain by specializing the elements of $\mathbb{Z}[\boldsymbol{u}]$ in the matrix M to the coefficients of $\boldsymbol{f_0}$, see Def. 3.2.2.

Consider a zero-dimensional affine system $\boldsymbol{f} := (f_1, \ldots, f_n)$ in $\mathbb{K}[\boldsymbol{x}]$ defined by polynomials of degrees $d_1, \ldots, d_n$. Let $\boldsymbol{f}^h := (f_1^h, \ldots, f_n^h)$ be the system obtained by homogenizing each f_i in $\mathbb{K}[\boldsymbol{x}][x_0]$, see Def. 4.3.1. Consider an affine polynomial $f_0 \in \mathbb{K}[\boldsymbol{x}]$ of degree d_0 and let $f_0^h \in \mathbb{K}[\boldsymbol{x}][x_0]_{d_0}$ be its homogenization. Then, we have the homogeneous system $\boldsymbol{f_0^h} := (f_1^h, \ldots, f_n^h, f_0^h)$. If $\mathcal{M}_{1,1}(\boldsymbol{f_0^h})$ is invertible, then we can read the multiplication map of f_0 in $\mathbb{K}[\boldsymbol{x}]/\langle f_1, \ldots, f_n \rangle$ from the Schur complement of $\mathcal{M}(\boldsymbol{f_0^h})$.

Definition 5.3.5 (Schur complement)**.** *Given a block matrix $M := \left[\begin{smallmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{smallmatrix}\right] \in \mathbb{K}^{\mathcal{K} \times \mathcal{K}}$ such that $M_{1,1}$ is invertible, its Schur complement is the matrix*

$$M_{2,2} - M_{2,1} \cdot (M_{1,1})^{-1} \cdot M_{1,2}.$$

Theorem 5.3.6. *[CLO06, Thm. 3.6.7] Consider a zero-dimensional affine system $\boldsymbol{f} := (f_1, \ldots, f_n)$ in $\mathbb{K}[\boldsymbol{x}]$ defined by polynomials of degrees $d_1, \ldots, d_n$, respectively. Let $\boldsymbol{f}^h := (f_1^h, \ldots, f_n^h)$ be the system obtained by homogenizing each f_i in $\mathbb{K}[\boldsymbol{x}][x_0]$, see Def. 4.3.1. Consider an affine polynomial $f_0 \in \mathbb{K}[\boldsymbol{x}]$ of degree d_0 and let $f_0^h \in \mathbb{K}[\boldsymbol{x}][x_0]_{d_0}$ be its homogenization. We consider the homogeneous system $\boldsymbol{f_0^h} := (f_1^h, \ldots, f_n^h, f_0^h)$. If $\mathcal{M}_{1,1}(\boldsymbol{f_0^h})$ is invertible, then*

- *The set B given by the dehomogenization of P_0, see Def. 4.3.5, is a monomial basis of $\mathbb{K}[\boldsymbol{x}]/\langle f_1, \ldots, f_n \rangle$, where*
$$B := \{\chi(\boldsymbol{x}^{\boldsymbol{\alpha}}) : \boldsymbol{x}^{\boldsymbol{\alpha}} \in P_0\}.$$
- *The Schur complement of $\mathcal{M}(\boldsymbol{f_0^h})$, that is,*
$$\mathcal{M}_{2,2}(\boldsymbol{f_0^h}) - \mathcal{M}_{2,1}(\boldsymbol{f_0^h}) \cdot (\mathcal{M}_{1,1}(\boldsymbol{f_0^h}))^{-1} \cdot \mathcal{M}_{1,2}(\boldsymbol{f_0^h}),$$
is the matrix of the multiplication map of f_0 in $\mathbb{K}[\boldsymbol{x}]/\langle f_1, \ldots, f_n \rangle$, with respect to the monomial basis B.

Observation 5.3.7. *Note that the specialization $\mathcal{M}_{1,1}(\boldsymbol{f_0^h})$ does not depend on the value of the polynomial $f_0^h \in \mathbb{K}[\boldsymbol{x}][x_0]_{d_0}$. The existence of an f_0 such that the $\mathcal{M}_{1,1}(\boldsymbol{f_0^h})$ is invertible is independent of f_0. Moreover, if there is an f_0 such that $\mathcal{M}_{1,1}(\boldsymbol{f_0^h})$ is invertible, then we can compute any multiplication map with respect to a polynomial f_0 of degree d_0.*

Consider a square system $(f_1, \ldots, f_n)$. It is problematic to use Thm. 5.3.6 to compute the multiplication maps for $\mathbb{K}[\boldsymbol{x}]/\langle f_1, \ldots, f_n\rangle$ in the following cases,

- when the system $(f_1^h, \ldots, f_n^h)$ has more solutions over $\mathbb{P}^n$ than $(f_1, \ldots, f_n)$ over $\mathbb{K}^n$, or
- when the specialization of the *extraneous factor* $\widetilde{\mathcal{M}}$ at $(\boldsymbol{f_0^h})$, $\widetilde{\mathcal{M}}_{1,1}(\boldsymbol{f_0^h})$, is not invertible,

In these cases, the matrix $\mathcal{M}_{1,1}(\boldsymbol{f_0^h})$ will not be invertible and so we cannot use this method to compute the multiplication maps. We refer the reader to the discussion at the end of [CLO06, Sec. 3.6], and references there in, for some techniques to avoid these problems.

Part II

Contributions

Chapter 6

Determinantal formulas and algorithms to solve mixed multilinear systems

Effective computation of resultants is a central problem in elimination theory and polynomial system solving, see Chapter 3. Commonly, we compute the resultant as a quotient of determinants of matrices, see Sec. 3.2.1. We say that there exists a *determinantal formula* when we can express it as the determinant of a matrix whose elements are the coefficients of the input polynomials, see Def. 3.2.14.

In Chapter 6, we consider mixed multilinear polynomial systems. On the one hand, this is the simplest case of mixed multihomogeneous systems where no determinantal formula was known. On the other hand, multilinear polynomial systems are common in applications, for example in the *Multiparameter Eigenvalue Problem* related to mathematical physics [Atk72, Vol88].

In the first part of the chapter, Section 6.1, we study determinantal formulas for the multiprojective resultant of *mixed multilinear polynomial systems*. Following [WZ94], we use the Weyman complex, see Def. 3.3.3, to introduce determinantal formulas for two kinds of mixed multilinear systems. If $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_A$ and $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_B$ are $(A+B)$ different blocks of variables, then we consider the following mixed multilinear systems $(f_1, \ldots, f_n)$:

1. **Star multilinear systems**: For each polynomial f_k there is $1 \leq j_k \leq B$ such that

$$f_k \in \mathbb{K}[\boldsymbol{X}_1]_1 \otimes \cdots \otimes \mathbb{K}[\boldsymbol{X}_A]_1 \otimes \mathbb{K}[\boldsymbol{Y}_{j_k}]_1.$$

2. **Bipartite bilinear systems**: For each polynomial f_k there are $1 \leq i_k \leq A$ and $1 \leq j_k \leq B$ such that

$$f_k \in \mathbb{K}[\boldsymbol{X}_{i_k}]_1 \otimes \mathbb{K}[\boldsymbol{Y}_{j_k}]_1.$$

We add an additional polynomial f_0, linear or multilinear, and show that the resultant of the new system is the determinant of a *Koszul resultant matrix* (related to the maps in the Koszul complex, Prop. 3.3.13). As the size of the matrix, that is, the degree of the resultant, depends on the multidegree of f_0, we derive determinantal formulas for different choices of f_0. We relate the size of the matrices to the number of solutions of $(f_1, \ldots, f_n)$ (Sections 6.1.2 and 6.1.3). For example, in Lem. 6.1.12 we prove that, if we consider a square *star multilinear system* $(f_1, \ldots, f_n)$ having Υ solutions and we introduce a multilinear

$f_0 \in \mathbb{K}[X_1]_1 \otimes \cdots \otimes \mathbb{K}[X_A]_1$, then the size of the matrix associated to the determinantal formula of $(f_0, f_1, \ldots, f_n)$ is

$$\Upsilon \cdot \left(\sum_{i=1}^{A} \#X_i - A + 1 \right).$$

Moreover, Υ is bigger than $(\sum_i \#X_i - A + 1)$ and so, the size of the formula is *polynomial in the number of solutions*.

In the second part of the chapter, Section 6.3, we exploit the structure of these *Koszul resultant matrices* to solve square *star multilinear systems* and *bipartite bilinear systems*. For that, we generalize the classical eigenvalue criterion for multiplication maps [Laz81], see Prop. 5.2.3. Our extension applies to a general class of matrices (Def. 6.2.1), including the Sylvester and Koszul resultant matrices. Moreover, it relies only on the degree and structure of the associated formula. We prove that if the matrix $\left[\begin{smallmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{smallmatrix}\right]$ corresponds to a determinantal formula for $(f_0, f_1, \ldots, f_n)$ such that the diagonal of the matrix $M_{2,2}$ corresponds to the coefficient of the monomial $\boldsymbol{x}^{\boldsymbol{\theta}}$ in f_0, this coefficient only appears in this diagonal, and the system $(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1, \ldots, f_n)$ has no solutions, then the matrix $M_{1,1}$ is invertible and eigenvalues of the Schur complement of $M_{2,2}$, $M_{2,2}^c := M_{2,2} - M_{2,1}\, M_{1,1}^{-1}\, M_{1,2}$, correspond to the evaluation of $\frac{f_0}{\boldsymbol{X}^{\boldsymbol{\theta}}}$ at every solution of $(f_1, \ldots, f_n)$. That is,

$$\lambda \text{ is an eigenvalue of } M_{2,2}^c \iff \exists\, \boldsymbol{\alpha} \text{ such that } \begin{cases} f_1(\boldsymbol{\alpha}) = \cdots = f_n(\boldsymbol{\alpha}) = 0 \text{ and} \\ \lambda = \frac{f_0}{\boldsymbol{X}^{\boldsymbol{\theta}}}(\boldsymbol{\alpha}) \end{cases}.$$

In the third part of the chapter, Section 6.3, we extend the classical eigenvector criterion [AS88], see Prop. 5.2.4, to the case of *Koszul resultant matrices* for *2-bilinear systems*. These systems correspond to the star multilinear systems where $A = 1$ and $B = 2$.

Finally, in Section 6.4, we merge all the tools that we introduce in the chapter to propose a new algorithm to solve the *Multiparameter Eigenvalue Problem*. The complexity of our approach is *polynomial in the number of solutions*.

6.1 Determinantal formulas for some mixed multilinear systems

6.1.1 Mixed multilinear systems of interest

In this section we present the mixed multilinear polynomial systems for which we develop determinantal formulas. We split the blocks of variables in two groups. For this and to simplify the presentation, we change somewhat the notation we use for the polynomial systems in Sec. 2.10 and Sec. 3.3.

For $s \in \mathbb{N}$, let $[s] := \{1, \dots, s\}$. We replace the blocks of variables $\boldsymbol{x}_i$ by $\boldsymbol{X}_i$ or $\boldsymbol{Y}_j$ and the constants n_i, that correspond to the cardinalities of the blocks, by α_i or β_j. Let $A, B \in \mathbb{N}$ and $q = A + B$. Let $\bar{\boldsymbol{X}}$ be the set of A blocks of variables $\{\boldsymbol{X}_1, \dots, \boldsymbol{X}_A\}$. For each $i \in [A]$, $\boldsymbol{X}_i := \{x_{i,0}, \dots, x_{i,\alpha_i}\}$; so the number of variables in each block is $\alpha_i \in \mathbb{N}$. We also consider the polynomial algebra $\mathbb{K}[\boldsymbol{X}_i] = \bigoplus_{d\in\mathbb{Z}} S_{\boldsymbol{X}_i}(d)$, where $S_{\boldsymbol{X}_i}(d)$ is the $\mathbb{K}$-vector space of polynomials of degree d in $\mathbb{K}[\boldsymbol{X}_i]$.

Similarly, $\bar{\boldsymbol{Y}}$ is the set of B blocks of variables $\{\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_B\}$. For each $j \in [B]$, $\boldsymbol{Y}_j := \{y_{j,0}, \dots, y_{j,\beta_j}\}$; hence the number of variables in each block is $\beta_j \in \mathbb{N}$. Moreover, $\mathbb{K}[\boldsymbol{Y}_j] = \bigoplus_{d\in\mathbb{Z}} S_{\boldsymbol{Y}_j}(d)$, where $S_{\boldsymbol{Y}_j}(d)$ is the $\mathbb{K}$-vector space of polynomials of degree d in $\mathbb{K}[\boldsymbol{Y}_j]$.

Consider the $\mathbb{Z}^q$-multigraded algebra $\mathbb{K}[\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}]$, given by

$$\mathbb{K}[\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}] := \bigoplus_{(d_{\boldsymbol{X}_1},\dots,d_{\boldsymbol{X}_A},d_{\boldsymbol{Y}_1},\dots,d_{\boldsymbol{Y}_B})\in\mathbb{Z}^q} S_{\boldsymbol{X}_1}(d_{\boldsymbol{X}_1}) \otimes \cdots \otimes S_{\boldsymbol{X}_A}(d_{\boldsymbol{X}_A}) \otimes S_{\boldsymbol{Y}_1}(d_{\boldsymbol{Y}_1}) \otimes \cdots \otimes S_{\boldsymbol{Y}_B}(d_{\boldsymbol{Y}_B}).$$

For a multihomogeneous polynomial f of multidegree $\boldsymbol{d} \in \mathbb{Z}^q$, we denote by $d_{\boldsymbol{X}_i}$, respectively $d_{\boldsymbol{Y}_j}$, the degree of f with respect to the block of variables $\boldsymbol{X}_i$, respectively $\boldsymbol{Y}_j$.

Let $N = \sum_{i=1}^{A} \alpha_i + \sum_{j=1}^{B} \beta_j$. We say that a polynomial system is *square* if it has N equations and *overdetermined* if it has $N + 1$. We work in the multiprojective space

$$\mathcal{P} := \mathbb{P}^{\alpha_1} \times \cdots \times \mathbb{P}^{\alpha_A} \times \mathbb{P}^{\beta_1} \times \cdots \times \mathbb{P}^{\beta_B}.$$

For each group of indices $1 \le i_1 < \cdots < i_r \le A$ and $1 \le j_1 < \cdots < j_s \le B$, we denote by $\mathbb{K}[\boldsymbol{X}_{i_1}, \dots, \boldsymbol{X}_{i_r}, \boldsymbol{Y}_{j_1}, \dots, \boldsymbol{Y}_{j_s}]_{\mathbf{1}}$ the set of multilinear polynomials in $\mathbb{K}[\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}]$ with multidegree $(d_{\boldsymbol{X}_1}, \dots, d_{\boldsymbol{X}_A}, d_{\boldsymbol{Y}_1}, \dots, d_{\boldsymbol{Y}_B})$, where

$$d_{\boldsymbol{X}_l} = \begin{cases} 1 & \text{if } l \in \{i_1, \dots, i_r\} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad d_{\boldsymbol{Y}_l} = \begin{cases} 1 & \text{if } l \in \{j_1, \dots, j_s\} \\ 0 & \text{otherwise.} \end{cases}.$$

Definition 6.1.1 (Star multilinear system)**.** *A square multihomogeneous system $\boldsymbol{f} = (f_1, \dots, f_N)$ in $\mathbb{K}[\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}]$ with multidegrees $\boldsymbol{d}_1, \dots, \boldsymbol{d}_N \in \mathbb{Z}^q$, respectively, is a* Star multilinear systems *if for every $k \in [N]$, there is $j_k \in [B]$ such that*

$$f_k \in \mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A, \boldsymbol{Y}_{j_k}]_{\mathbf{1}}.$$

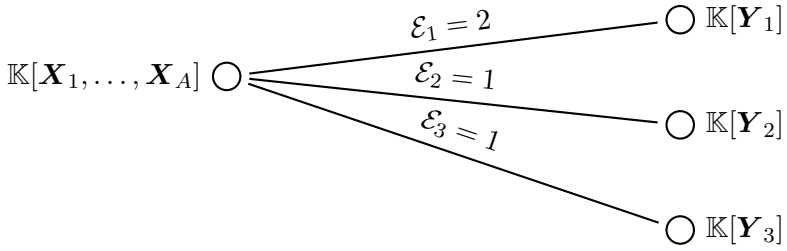
For each $j \in [B]$, we denote by $\mathcal{E}_j$ the number of polynomials in $\boldsymbol{f}$ that belong to $\mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A, \boldsymbol{Y}_j]_{\mathbf{1}}$.

We use the term *star* because we can represent such systems using a star graph. The vertices of the graph are the algebras $\mathbb{K}[\boldsymbol{Y}_1], \dots, \mathbb{K}[\boldsymbol{Y}_B]$ and $\mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A]$. For each $\boldsymbol{d}_k$ there is an edge between the vertices $\mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A]$ and $\boldsymbol{Y}_j$ whenever $d_{k,Y_j} = 1$. The graph is a star because every vertex is connected to $\mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A]$ and there is no edge between two vertices $\mathbb{K}[\boldsymbol{Y}_{j_1}]$ and $\mathbb{K}[\boldsymbol{Y}_{j_2}]$.

Example 6.1.2. *Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{Y}_3$ be five blocks of variables. Consider the multihomogeneous system $(f_1, f_2, f_3, f_4) \subset \mathbb{K}[\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}]$ with the following (pattern of) multidegrees*

$$\begin{array}{r|cc|ccc|} & (& d_{k,X_1}, & d_{k,X_2}, & d_{k,Y_1}, & d_{k,Y_2}, & d_{k,Y_3} &) \\ \boldsymbol{d}_1 = & (& 1, & 1, & 1, & 0, & 0 &) \\ \boldsymbol{d}_2 = & (& 1, & 1, & 1, & 0, & 0 &) \\ \boldsymbol{d}_3 = & (& 1, & 1, & 0, & 1, & 0 &) \\ \boldsymbol{d}_4 = & (& 1, & 1, & 0, & 0, & 1 &) \end{array}$$

It is a star multilinear system where $\mathcal{E}_1 = 2$, $\mathcal{E}_2 = 1$, and $\mathcal{E}_3 = 1$. The corresponding star graph is the following:



Definition 6.1.3 (Bipartite bilinear system)**.** *A square multihomogeneous system $\boldsymbol{f} = (f_1, \ldots, f_N)$ in $\mathbb{K}[\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}]$ with multidegrees $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N \in \mathbb{Z}^q$, respectively, is a* bipartite bilinear system *if for every $k \in [N]$, there are $i_k \in [A]$ and $j_k \in [B]$ such that*

$$f_k \in \mathbb{K}[\boldsymbol{X}_{i_k}, \boldsymbol{Y}_{j_k}]_{\mathbf{1}}.$$

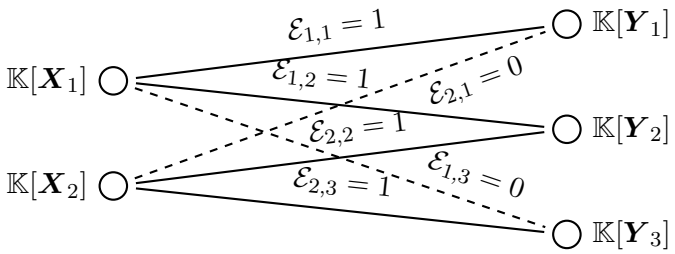
For each $i \in [A]$ and $j \in [B]$, let $\mathcal{E}_{i,j}$ be the number of polynomials in $\mathbb{K}[\boldsymbol{X}_i, \boldsymbol{Y}_j]_{\mathbf{1}}$.

We use the term *bipartite* because we can represent such systems using a bipartite graph. The vertices of the graph are the algebras $\mathbb{K}[\boldsymbol{X}_1], \ldots, \mathbb{K}[\boldsymbol{X}_A]$ and $\mathbb{K}[\boldsymbol{Y}_1], \ldots, \mathbb{K}[\boldsymbol{Y}_B]$. For each $\boldsymbol{d}_k$ there is an edge between the vertices $\mathbb{K}[\boldsymbol{X}_i]$ and $\mathbb{K}[\boldsymbol{Y}_j]$ whenever $d_{k,X_i} = d_{k,Y_j} = 1$. The graph is bipartite because we can partition the vertices to two sets, $\{\mathbb{K}[\boldsymbol{X}_1], \ldots, \mathbb{K}[\boldsymbol{X}_A]\}$ and $\{\mathbb{K}[\boldsymbol{Y}_1], \ldots, \mathbb{K}[\boldsymbol{Y}_B]\}$ such that there is no edge between vertices belonging to the same set.

Example 6.1.4. *Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{Y}_3$ be five blocks of variables. Consider the multihomogeneous system $(f_1, f_2, f_3, f_4) \subset \mathbb{K}[\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}]$ with multidegrees*

$$\begin{array}{r|cc|ccc|} & (& d_{i,X_1}, & d_{i,X_2}, & d_{i,Y_1}, & d_{i,Y_2}, & d_{i,Y_3} &) \\ \boldsymbol{d}_1 = & (& 1, & 0, & 1, & 0, & 0 &) \\ \boldsymbol{d}_2 = & (& 1, & 0, & 0, & 1, & 0 &) \\ \boldsymbol{d}_3 = & (& 0, & 1, & 0, & 1, & 0 &) \\ \boldsymbol{d}_4 = & (& 0, & 1, & 0, & 0, & 1 &) \end{array}$$

This system is a bipartite bilinear system where $\mathcal{E}_{1,1} = 1$, $\mathcal{E}_{1,2} = 1$, $\mathcal{E}_{1,3} = 0$, $\mathcal{E}_{2,1} = 0$, $\mathcal{E}_{2,2} = 1$ and $\mathcal{E}_{2,3} = 1$. The corresponding bipartite graph is the following:



Remark 6.1.5. *For each square bipartite bilinear system, it holds* $N = \sum_{i=1}^{A}\sum_{j=1}^{B}\mathcal{E}_{i,j}$. *Moreover, if the system has a finite number of solutions, then for each* $i \in \{1,\dots,A\}$, *it holds* $\sum_{j=1}^{B}\mathcal{E}_{i,j} \geq \alpha_i$ *and for each* $j \in \{1,\dots,B\}$ *it holds* $\sum_{i=1}^{A}\mathcal{E}_{i,j} \geq \beta_j$; *see Prop. 2.10.9.*

6.1.2 Determinantal formulas for star multilinear systems

We consider four different kinds of overdetermined multihomogeneous systems, related to *star multilinear systems* (Def. 6.1.1) and we construct determinantal formulas for each of them. These formulas are Koszul- and Sylvester-type determinantal formulas (Def. 3.3.14). We study overdetermined polynomial systems $(f_0, f_1, \dots, f_N)$ in $\mathbb{K}[\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}]$ where $(f_1, \dots, f_N)$ is a square star multilinear system and f_0 is a multilinear polynomial.

We consider different types of polynomials f_0. The obvious choice for f_0 is to have the same structure as one of the polynomials $f_1, \dots, f_N$; still we also choose f_0 to have a different support. This leads to resultants of smaller degrees and so to matrices of smaller size. The following f_0 lead to determinantal formulas:

- **Case 1:** $f_0 \in \mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A]_{\mathbf{1}}$.
- **Case 2:** $f_0 \in \mathbb{K}[\boldsymbol{Y}_j]_{\mathbf{1}}$, for any $j \in \{1,\dots,B\}$.
- **Case 3:** $f_0 \in \mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A, \boldsymbol{Y}_j]_{\mathbf{1}}$, for any $j \in \{1,\dots,B\}$,
- **Case 4:** $f_0 \in \mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A, \boldsymbol{Y}_{j_1}, \boldsymbol{Y}_{j_2}]_{\mathbf{1}}$, for any $j_1, j_2 \in \{1,\dots,B\}$ such that $j_1 \neq j_2$.

We can interpret the various multidegrees of f_0, $\boldsymbol{d}_0 = (d_{0,X_1}, \dots, d_{0,X_A}, d_{0,Y_1}, \dots, d_{0,Y_B})$, that lead to determinantal formulas as solutions of the following system of inequalities:

$$\begin{cases} (\forall\, 1 \leq i \leq A) & 0 \leq d_{0,\boldsymbol{X}_i} \leq 1, \\ (\forall\, 1 \leq j \leq B) & 0 \leq d_{0,\boldsymbol{Y}_j} \leq 1, \\ (\forall\, 1 \leq i_1 < i_2 \leq A) & d_{0,\boldsymbol{X}_{i_1}} = d_{0,\boldsymbol{X}_{i_2}}, \text{ and} \\ & \sum_{j=1}^{B} d_{0,\boldsymbol{Y}_j} \leq 1 + d_{0,\boldsymbol{X}_1}. \end{cases} \tag{6.1}$$

Definition 6.1.6. *Consider a partition of* $\{1,\dots,B\}$ *consisting of two sets* P *and* D *and a constant* $c \in \mathbb{N}$. *We say that the triplet* (P, D, c) *is* determinantal data *in the following cases:*

- *When* f_0 *corresponds to* ***cases 1 or 3****: if it holds,* $0 \leq c \leq A$.

- *When f_0 corresponds to **case 2**: if it holds,*

$$\begin{cases} c = 0 & \text{if } \sum_{j\in P} d_{0,\boldsymbol{Y}_j} = 0, \text{ or} \\ c = A & \text{if } \sum_{j\in D} d_{0,\boldsymbol{Y}_j} = 0. \end{cases}$$

- *When f_0 corresponds to **case 4**: if it holds,*

$$\begin{cases} 0 \le c \le A, \\ \sum_{j\in P} d_{0,\boldsymbol{Y}_j} \le 1, \text{ and} \\ \sum_{j\in D} d_{0,\boldsymbol{Y}_j} \le 1. \end{cases}$$

Equivalently, we say that the triplet (P, D, c) is determinantal data *for the multidegree $\boldsymbol{d}_0$ if the following conditions are true:*

$$\begin{cases} \sum_{j\in P} d_{0,\boldsymbol{Y}_j} \le 1 & \\ \sum_{j\in D} d_{0,\boldsymbol{Y}_j} \le 1 & \\ 0 \le c \le A & \text{if } (\forall\, i \in [A]) \text{ it holds } d_{0,\boldsymbol{X}_i} = 1, \\ c = 0 & \text{if } (\forall\, i \in [A]) \text{ it holds } d_{0,\boldsymbol{X}_i} = 0 \text{ and } \sum_{j\in P} d_{0,\boldsymbol{Y}_j} = 0, \\ c = A & \text{if } (\forall\, i \in [A]) \text{ it holds } d_{0,\boldsymbol{X}_i} = 0 \text{ and } \sum_{j\in D} d_{0,\boldsymbol{Y}_j} = 0. \end{cases} \tag{6.2}$$

Consider the set $\{0, \ldots, N\}$ that corresponds to generic polynomials $\boldsymbol{F} = (F_0, \ldots, F_N)$ (Def. 3.3.1). For each tuple $s_0, \ldots, s_B \in \mathbb{N}$, let $\mathcal{I}_{s_0,s_1,\ldots,s_B}$ be the set of all the subsets of $\{0, \ldots, N\}$, such that

- For $1 \le j \le B$, the index s_j indicates that we consider s_j polynomials from $\boldsymbol{F}$ that belong to $\mathbb{Z}[\boldsymbol{u}][\boldsymbol{X}_1, \ldots, \boldsymbol{X}_A, \boldsymbol{Y}_j]_{\mathbf{1}}$.
- In addition, if $s_0 = 1$, then 0 belongs to all the sets in $\mathcal{I}_{s_0,s_1,\ldots,s_B}$.

That is,

$$\begin{aligned} \mathcal{I}_{s_0,s_1,\ldots,s_B} := \Big\{ & I : I \subset \{0, \ldots, n\}, (0 \in I \Leftrightarrow s_0 = 1) \text{ and} \\ & (\forall 1 \le j \le B)\ s_j = \#\{k \in I : F_k \in \mathbb{Z}[\boldsymbol{u}][\boldsymbol{X}_1, \ldots, \boldsymbol{X}_A, \boldsymbol{Y}_j]_{\mathbf{1}}\} \Big\}. \end{aligned} \tag{6.3}$$

Notice that if $I, J \in \mathcal{I}_{s_0,s_1,\ldots,s_B}$, then I and J have the same cardinality and $\sum_{k\in I} \boldsymbol{d}_k = \sum_{k\in J} \boldsymbol{d}_k$, as they correspond to subsets of polynomials of $\boldsymbol{F}$ with the same supports.

The following lemma exploits the sets $\mathcal{I}_{s_0,s_1,\ldots,s_B}$ to rewrite the cohomologies of Eq. (3.8).

Lemma 6.1.7. *Consider a generic overdetermined system $\boldsymbol{F} = (F_0, \ldots, F_N)$ in $\mathbb{K}[\boldsymbol{u}][\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}]$ of multidegrees $\boldsymbol{d}_0, \ldots, \boldsymbol{d}_N$ (Def. 3.3.1), where $(F_1, \ldots, F_n)$ is a square star multilinear system such that, for every $j \in \{1, \ldots, B\}$, $\mathcal{E}_j \ge \beta_j$, and $\boldsymbol{d}_0$ is the multidegree of F_0. Following Eq.* (3.8) *we can rewrite the modules of the Weyman complex $K_v(\boldsymbol{m}) = \bigoplus_{p=0}^{N+1} K_{v,p} \otimes \mathbb{Z}[\boldsymbol{u}]$ in the more detailed form*

$$K_{v,p}(\boldsymbol{m}) = \bigoplus_{\substack{0\le s_0\le 1 \\ 0\le s_1\le \mathcal{E}_1 \\ \cdots \\ 0\le s_B\le \mathcal{E}_B \\ s_0+s_1+\cdots+s_B=p}} H_{\mathcal{P}}^{p-v}\Big(\boldsymbol{m} - (\sum_{j=1}^{B} s_j, \ldots, \sum_{j=1}^{B} s_j, s_1, \ldots, s_B) - s_0\, \boldsymbol{d}_0\Big) \otimes \bigoplus_{I\in\mathcal{I}_{s_0,s_1,\ldots,s_B}} \bigwedge_{k\in I} e_k.$$

Moreover, the following isomorphisms hold for the cohomologies:

$$H^{p-v}_{\mathcal{P}}\Big(\boldsymbol{m}-(\sum_{j=1}^{B}s_j,\dots,\sum_{j=1}^{B}s_j,s_1,\dots,s_B)-s_0\,\boldsymbol{d}_0\Big)\cong$$

$$\bigoplus_{\substack{r_{\boldsymbol{X}_1},\dots,r_{\boldsymbol{X}_1},r_{\boldsymbol{Y}_1},\dots,r_{\boldsymbol{Y}_B}\\ \sum_i r_{\boldsymbol{X}_i}+\sum_j r_{\boldsymbol{Y}_j}=p-v}}\Big(\bigotimes_{i=1}^{A}H^{r_{\boldsymbol{X}_i}}_{\mathbb{P}^{\alpha_i}}\Big(m_{\boldsymbol{X}_i}-\sum_{j=1}^{B}s_j-s_0\,d_{0,\boldsymbol{X}}\Big)\otimes\bigotimes_{j=1}^{B}H^{r_{\boldsymbol{Y}_j}}_{\mathbb{P}^{\beta_j}}(m_{\boldsymbol{Y}_j}-s_j-s_0\,d_{0,\boldsymbol{Y}_j})\Big).$$

Proof. We notice that if for $I,J\subset\{0,\dots,N\}$ it holds $\sum_{k\in I}\boldsymbol{d}_k=\sum_{k\in J}\boldsymbol{d}_k$, then

$$\Big(H^{p-v}_{\mathcal{P}}(\boldsymbol{m}-\sum_{k\in I}\boldsymbol{d}_k)\otimes\bigwedge_{k\in I}e_k\Big)\oplus\Big(H^{p-v}_{\mathcal{P}}(\boldsymbol{m}-\sum_{k\in J}\boldsymbol{d}_k)\otimes\bigwedge_{k\in J}e_k\Big)\cong$$
$$H^{p-v}_{\mathcal{P}}(\boldsymbol{m}-\sum_{k\in I}\boldsymbol{d}_k)\otimes\Big(\bigwedge_{k\in I}e_k\oplus\bigwedge_{k\in J}e_k\Big).$$

Now, if $I,J\subset\mathcal{I}_{s_0,s_1,\dots,s_B}$, then by definition $\sum\limits_{k\in I}\boldsymbol{d}_k=\sum\limits_{k\in J}\boldsymbol{d}_k=(\sum_{j=1}^{B}s_j,\dots,\sum_{j=1}^{B}s_j,s_1,\dots,s_B)+s_0\,\boldsymbol{d}_0$, and so

$$H^{p-v}_{\mathcal{P}}(\boldsymbol{m}-\sum_{k\in I}\boldsymbol{d}_k)=H^{p-v}_{\mathcal{P}}(\boldsymbol{m}-\sum_{k\in J}\boldsymbol{d}_k)=H^{p-v}_{\mathcal{P}}\Big(\boldsymbol{m}-(\sum_{j=1}^{B}s_j,\dots,\sum_{j=1}^{B}s_j,s_1,\dots,s_B)-s_0\,\boldsymbol{d}_0\Big).$$

By definition of $\mathcal{E}_1,\dots,\mathcal{E}_B$ (Def. 6.1.1) the set $\mathcal{I}_{s_0,s_1,\dots,s_B}$ is not empty if and only if $0\le s_0\le 1$ and for all $i\in\{1,\dots,B\}$ it holds $0\le s_i\le\mathcal{E}_i$. Hence,

$$\{I:I\subset\{0,\dots,N\},\#I=p\}=\bigcup_{\substack{0\le s_0\le 1\\ 0\le s_1\le\mathcal{E}_1\\ \dots\\ 0\le s_B\le\mathcal{E}_B\\ s_0+s_1+\dots+s_B=p}}\mathcal{I}_{s_0,s_1,\dots,s_B}.$$

Thus, we can write each direct summand of $K_v(\boldsymbol{m})=\bigoplus_{p=0}^{N+1}K_{v,p}\otimes\mathbb{K}[\boldsymbol{u}]$ as

$$K_{v,p}(\boldsymbol{m})=\bigoplus_{\substack{0\le s_0\le 1\\ 0\le s_1\le\mathcal{E}_1\\ \dots\\ 0\le s_B\le\mathcal{E}_B\\ s_0+s_1+\dots+s_B=p}}H^{p-v}_{\mathcal{P}}\Big(\boldsymbol{m}-(\sum_{j=1}^{B}s_j,\dots,\sum_{j=1}^{B}s_j,s_1,\dots,s_B)-s_0\,\boldsymbol{d}_0\Big)\otimes\bigoplus_{I\in\mathcal{I}_{s_0,s_1,\dots,s_B}}\bigwedge_{k\in I}e_k.\quad(6.4)$$

Finally, by means of Prop. 3.3.5, we have the isomorphism

$$H^{p-v}_{\mathcal{P}}\Big(\boldsymbol{m}-(\sum_{j=1}^{B}s_j,\dots,\sum_{j=1}^{B}s_j,s_1,\dots,s_B)-s_0\,\boldsymbol{d}_0\Big)\cong$$

$$\bigoplus_{\substack{r_{\boldsymbol{X}_1},\dots,r_{\boldsymbol{X}_1},r_{\boldsymbol{Y}_1},\dots,r_{\boldsymbol{Y}_B}\\ \sum_i r_{\boldsymbol{X}_i}+\sum_j r_{\boldsymbol{Y}_j}=p-v}}\Big(\bigotimes_{i=1}^{A}H^{r_{\boldsymbol{X}_i}}_{\mathbb{P}^{\alpha_i}}(m_{\boldsymbol{X}_i}-\sum_{j=1}^{B}s_j-s_0\,d_{0,\boldsymbol{X}})\otimes\bigotimes_{j=1}^{B}H^{r_{\boldsymbol{Y}_j}}_{\mathbb{P}^{\beta_j}}(m_{\boldsymbol{Y}_j}-s_j-s_0\,d_{0,\boldsymbol{Y}_j})\Big).$$

□

The following theorem identifies the degree vectors that reduce the Weyman complex to have just two elements and in this way it provides us a determinantal formula for square star multilinear systems.

Theorem 6.1.8. *Consider a generic overdetermined system $\boldsymbol{F} = (F_0, \dots, F_N)$ in $\mathbb{Z}[\boldsymbol{u}][\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}]$ of multidegrees $\boldsymbol{d}_0, \dots, \boldsymbol{d}_N$ (Def. 3.3.1), where $(F_1, \dots, F_n)$ is a square star multilinear system. For every $j \in \{1, \dots, B\}$ it holds $\mathcal{E}_j \geq \beta_j$ and the multidegree of F_0, $\boldsymbol{d}_0$, is a solution of the system in Eq. 6.1.*

For each determinantal data (P, D, c) *(Def. 6.1.6) and a permutation $\sigma : \{1, \dots, A\} \to \{1, \dots, A\}$, there is a degree vector $\boldsymbol{m} = (m_{\boldsymbol{X}_1}, \dots, m_{\boldsymbol{X}_A}, m_{\boldsymbol{Y}_1}, \dots, m_{\boldsymbol{Y}_B})$, such that*

$$\begin{cases} m_{\boldsymbol{X}_i} = \sum_{j \in D} \beta_j + \sum_{k=1}^{\sigma(i)-1} \alpha_{\sigma^{-1}(k)} + d_{0,\boldsymbol{X}_i} & \text{for } 1 \leq i \leq A \text{ and } \sigma(i) > c \\ m_{\boldsymbol{X}_i} = \sum_{j \in D} \beta_j + \sum_{k=1}^{\sigma(i)-1} \alpha_{\sigma^{-1}(k)} - 1 & \text{for } 1 \leq i \leq A \text{ and } \sigma(i) \leq c \\ m_{\boldsymbol{Y}_j} = \mathcal{E}_j - \beta_j + d_{0,\boldsymbol{Y}_j} & \text{for } j \in P \\ m_{\boldsymbol{Y}_j} = -1 & \text{for } j \in D \end{cases}$$

so that the Weyman complex of $K_\bullet(\boldsymbol{m})$ reduces to

$$K_\bullet(\boldsymbol{m}) : 0 \to K_{1,\omega+1}(\boldsymbol{m}) \otimes \mathbb{K}[\boldsymbol{u}] \xrightarrow{\delta_1(\boldsymbol{m})} K_{0,\omega}(\boldsymbol{m}) \otimes \mathbb{K}[\boldsymbol{u}] \to 0,$$

where $\omega = \sum_{i=1}^{c} \alpha_{\sigma^{-1}(i)} + \sum_{j \in D} \beta_j$.

Hence, the map $\delta_1(\boldsymbol{m})$ is a Koszul-type determinantal formula (Def. 3.3.14).

Proof. We rewrite Eq. (3.9) using Lem. 6.1.7 and consider

$$H^{p-v}_{\mathcal{P}}\Big(\boldsymbol{m} - (\sum_{j=1}^{B} s_j, \dots, \sum_{j=1}^{B} s_j, s_1, \dots, s_B) - s_0\, \boldsymbol{d}_0\Big) \cong$$

$$\bigoplus_{\substack{r_{\boldsymbol{X}_1}, \dots, r_{\boldsymbol{X}_1}, r_{\boldsymbol{Y}_1}, \dots, r_{\boldsymbol{Y}_B} \\ \sum_i r_{\boldsymbol{X}_i} + \sum_j r_{\boldsymbol{Y}_j} = p-v}} \left(\begin{array}{ll} \bigotimes_{j \in P} H^{r_{\boldsymbol{Y}_j}}_{\mathbb{P}^{\beta_j}}(\mathcal{E}_j - \beta_j + d_{0,\boldsymbol{Y}_j} - s_j - s_0\, d_{0,\boldsymbol{Y}_j}) \otimes & \text{[Case Y.1]} \\ \bigotimes_{j \in D} H^{r_{\boldsymbol{Y}_j}}_{\mathbb{P}^{\beta_j}}(-1 - s_j - s_0\, d_{0,\boldsymbol{Y}_j}) \otimes & \text{[Case Y.2]} \\ \bigotimes_{i=1}^{c} H^{r_{\boldsymbol{X}_i}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(i)}}}\Big(\sum_{j \in D} \beta_j + \sum_{k=1}^{i-1} \alpha_{\sigma^{-1}(k)} - 1 - \sum_{j=1}^{B} s_j - s_0\, d_{0,\boldsymbol{X}}\Big) \otimes & \text{[Case X.1]} \\ \bigotimes_{i=c+1}^{A} H^{r_{\boldsymbol{X}_i}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(i)}}}\Big(\sum_{j \in D} \beta_j + \sum_{k=1}^{i-1} \alpha_{\sigma^{-1}(k)} + d_{0,\boldsymbol{X}_i} - \sum_{j=1}^{B} s_j - s_0\, d_{0,\boldsymbol{X}_i}\Big) & \text{[Case X.2]} \end{array} \right) \tag{6.5}$$

We will study the values for $p, v, s_0, \dots, s_B, r_{\boldsymbol{X}_1}, \dots, r_{\boldsymbol{X}_A}, r_{\boldsymbol{Y}_1}, \dots, r_{\boldsymbol{Y}_B}$ such that $K_{v,p}(\boldsymbol{m})$ does not vanish. Clearly, if $0 \leq s_0 \leq 1$ and $(\forall i \in \{1, \dots, B\})\ 0 \leq s_i \leq \mathcal{E}_i$, then the module $\bigoplus_{I \in \mathcal{I}_{s_0, s_1, \dots, s_B}} \bigwedge_{k \in I} e_k$ is not zero. Hence, assuming $0 \leq s_0 \leq 1$ and $(\forall i \in \{1, \dots, B\})\ 0 \leq s_i \leq \mathcal{E}_i$, we study the vanishing of the modules in Eq. (6.5). We will study the cohomologies independently. By Prop. 3.3.6, the modules in the right hand side of Eq. (6.5) are not zero only when, for $1 \leq i \leq A$, $r_{\boldsymbol{X}_i} \in \{0, \alpha_i\}$ and, for $1 \leq j \leq B$, $r_{\boldsymbol{Y}_j} \in \{0, \beta_j\}$. In the following we prove that if Eq. (6.5) does not

vanish then,

$$\begin{cases} r_{\boldsymbol{Y}_j} = 0 & \text{for } j \in P & \text{[Case Y.1]} \\ r_{\boldsymbol{Y}_j} = \beta_j & \text{for } j \in D & \text{[Case Y.2]} \\ r_{\boldsymbol{X}_i} = \alpha_i & \text{for } 1 \leq i \leq A \text{ and } \sigma(i) \leq c & \text{[Case X.1]} \\ r_{\boldsymbol{X}_i} = 0 & \text{for } 1 \leq i \leq A \text{ and } \sigma(i) > c & \text{[Case X.2]} \end{cases} \tag{6.6}$$

Consider the modules related to the variables $\boldsymbol{Y}_j$, for $j \in \{1, \ldots, B\}$.

Case (Y.1) We consider the modules that involve the variables in the block $\boldsymbol{Y}_j$, for $j \in P$. As $s_j \leq \mathcal{E}_j$ and $0 \leq s_0, d_{0,\boldsymbol{Y}_j} \leq 1$, it holds $\mathcal{E}_j - \beta_j + d_{0,\boldsymbol{Y}_j} - s_j - s_0\, d_{0,\boldsymbol{Y}_j} > -\beta_j - 1$. Hence, by Cor. 3.3.7,

$$H^{r_{\boldsymbol{Y}_j}}_{\mathbb{P}^{\beta_j}}(\mathcal{E}_j - \beta_j + d_{0,\boldsymbol{Y}_j} - s_j - s_0\, d_{0,\boldsymbol{Y}_j}) \neq 0 \Longleftrightarrow \left(r_{\boldsymbol{Y}_j} = 0 \text{ and } \mathcal{E}_j - \beta_j + d_{0,\boldsymbol{Y}_j} \geq s_j + s_0\, d_{0,\boldsymbol{Y}_j}\right). \tag{6.7}$$

Case (Y.2) We consider the modules that involve the variables in the block $\boldsymbol{Y}_j$, for $j \in D$. As $s_j, s_0, d_{0,\boldsymbol{Y}_j} \geq 0$, then $-1 - s_j - s_0\, d_{0,\boldsymbol{Y}_j} < 0$. Hence, by Cor. 3.3.7,

$$H^{r_{\boldsymbol{Y}_j}}_{\mathbb{P}^{\beta_j}}(-1 - s_j - s_0\, d_{0,\boldsymbol{Y}_j}) \neq 0 \iff \left(r_{\boldsymbol{Y}_j} = \beta_j \quad \text{and} \quad s_j + s_0\, d_{0,\boldsymbol{Y}_j} \geq \beta_j\right). \tag{6.8}$$

Now we consider the cohomologies related to the blocks of variables $\boldsymbol{X}_i$, for $i \in \{1, \ldots, A\}$, assuming that the cohomologies related to the blocks of variables $\boldsymbol{Y}_j$ do not vanish.

Case (X.1) We consider the modules related to the blocks $\boldsymbol{X}_{\sigma^{-1}(1)} \ldots, \boldsymbol{X}_{\sigma^{-1}(c)}$. We only need to consider this case if $c > 0$, so we assume $c > 0$. We prove by induction that for each $1 \leq k \leq c$, if the cohomologies related to the variables in the blocks $\boldsymbol{Y}_j$, for $1 \leq j \leq B$, and the ones related to $\boldsymbol{X}_{\sigma^{-1}(1)} \ldots, \boldsymbol{X}_{\sigma^{-1}(k-1)}$, do not vanish, then

$$\begin{gathered} H^{r_{\boldsymbol{X}_{\sigma^{-1}(k)}}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(k)}}}\left(m_{\boldsymbol{X}_{\sigma^{-1}(k)}} - \sum_{j=1}^{B} s_j - s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}}\right) \neq 0 \iff \\ r_{\boldsymbol{X}_{\sigma^{-1}(k)}} = \alpha_{\sigma^{-1}(k)} \text{ and } \sum_{j=1}^{B} s_j + s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} \geq \sum_{j \in D} \beta_j + \sum_{i=1}^{k} \alpha_{\boldsymbol{X}_{\sigma^{-1}(i)}}. \end{gathered} \tag{6.9}$$

- Consider $k = 1$ and the cohomology related to the block $\boldsymbol{X}_{\sigma^{-1}(1)}$,

$$H^{r_{\boldsymbol{X}_{\sigma^{-1}(1)}}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(1)}}}\Big(\sum_{j \in D} \beta_j - 1 - \sum_{j=1}^{B} s_j - s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}}\Big).$$

As the cohomologies related to the blocks $\boldsymbol{Y}_j$, for each $j \in \{1, \ldots, B\}$ do not vanish, then

- For each $j \in D$, $\beta_j \leq s_j + s_0\, d_{0,\boldsymbol{Y}_j}$ (Eq. (6.8)).
- For each $j \in P$, $0 \leq s_j$.

Adding these inequalities we conclude that,

$$\sum_{j=1}^{B} s_j + s_0 \sum_{j\in D} d_{0,\boldsymbol{Y}_j} \geq \sum_{j\in D} \beta_j. \tag{6.10}$$

As we assumed that $c > 0$ and the triplet (P, D, c) is *determinantal data* (Def. 6.1.6), by definition either $d_{0,\boldsymbol{X}_{\sigma^{-1}(1)}} = 1$ or both $d_{0,\boldsymbol{X}_{\sigma^{-1}(1)}} = 0$ and $\sum_{j\in D} d_{0,\boldsymbol{Y}_j} = 0$. Also, it holds $0 \leq s_0, \sum_{j\in D} d_{0,\boldsymbol{Y}_j} \leq 1$. We conclude that,

$$\sum_{j\in D} \beta_j - 1 - \sum_{j=1}^{B} s_j - s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} \leq s_0 \sum_{j\in D} d_{0,\boldsymbol{Y}_j} - 1 - s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} < 0$$

Hence, by Cor. 3.3.7,

$$H^{r_{\boldsymbol{X}_{\sigma^{-1}(1)}}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(1)}}} \Big(\sum_{j\in D} \beta_j - 1 - \sum_{j=1}^{B} s_j - s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(1)}} \Big) \neq 0 \iff$$

$$r_{\boldsymbol{X}_{\sigma^{-1}(1)}} = \alpha_{\sigma^{-1}(1)} \quad \text{and} \quad \sum_{j=1}^{B} s_j + s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(1)}} \geq \sum_{j\in D} \beta_j + \alpha_{\boldsymbol{X}_{\sigma^{-1}(1)}}$$

- We proceed by induction, assuming that Eq. 6.9 holds for $k-1$, we prove the property for k. We consider the cohomology

$$H^{r_{\boldsymbol{X}_{\sigma^{-1}(k)}}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(k)}}} \Big(\sum_{j\in D} \beta_j - 1 + \sum_{i=1}^{k-1} \alpha_{\sigma^{-1}(i)} - \sum_{j=1}^{B} s_j - s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} \Big).$$

By definition (Eq. (6.1)), $d_{0,\boldsymbol{X}_{\sigma^{-1}(k-1)}} = d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}}$, and by inductive hypothesis, if the cohomologies related to the blocks $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_B, \boldsymbol{X}_{\sigma^{-1}(1)}, \ldots, \boldsymbol{X}_{\sigma^{-1}(k-1)}$ do not vanish, then

$$\sum_{j=1}^{B} s_j + s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} = \sum_{j=1}^{B} s_j + s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k-1)}} \geq \sum_{j\in D} \beta_j + \sum_{i=1}^{k-1} \alpha_{\boldsymbol{X}_{\sigma^{-1}(i)}}.$$

Hence, by Cor. 3.3.7,

$$H^{r_{\boldsymbol{X}_{\sigma^{-1}(k)}}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(k)}}} \Big(\sum_{j\in D} \beta_j - 1 + \sum_{i=1}^{k-1} \alpha_{\sigma^{-1}(i)} - \sum_{j=1}^{B} s_j - s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} \Big) \neq 0 \iff$$

$$r_{\boldsymbol{X}_{\sigma^{-1}(k)}} = \alpha_{\sigma^{-1}(k)} \quad \text{and} \quad \sum_{j=1}^{B} s_j + s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} \geq \sum_{j\in D} \beta_j + \sum_{i=1}^{k-1} \alpha_{\sigma^{-1}(i)} + \alpha_{\sigma^{-1}(k)}.$$

Case (X.2) We consider the module related to the blocks $\boldsymbol{X}_{\sigma^{-1}(c+1)} \ldots, \boldsymbol{X}_{\sigma^{-1}(A)}$. We only need to consider this case if $c < A$, so we assume $c < A$.

We prove by induction that for each $c < k \leq A$, if the cohomologies related to the variables in the blocks $\boldsymbol{Y}_j$, for $1 \leq j \leq B$, and the ones related to $\boldsymbol{X}_{\sigma^{-1}(k+1)}, \boldsymbol{X}_{\sigma^{-1}(k+2)}, \ldots, \boldsymbol{X}_{\sigma^{-1}(A)}$, do not vanish, then

$$\begin{aligned} &H^{r_{\boldsymbol{X}_{\sigma^{-1}(k)}}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(k)}}}\Big(\sum_{j\in D}\beta_j + \sum_{i=1}^{k-1}\alpha_{\sigma^{-1}(i)} + (1-s_0)\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} - \sum_{j=1}^{B} s_j\Big) \neq 0 \iff \\ &r_{\boldsymbol{X}_{\sigma^{-1}(k)}} = 0 \quad \text{and} \quad \sum_{j\in D}\beta_j + \sum_{i=1}^{k-1}\alpha_{\sigma^{-1}(i)} \geq \sum_{j=1}^{B} s_j + (s_0-1)\; d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}}. \end{aligned} \tag{6.11}$$

- Consider $k = A$ and the cohomology related to the block $\boldsymbol{X}_{\sigma^{-1}(A)}$,

$$H^{r_{\boldsymbol{X}_{\sigma^{-1}(A)}}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(A)}}}\Big(\sum_{j\in D}\beta_j + \sum_{i=1}^{A-1}\alpha_{\sigma^{-1}(i)} + (1-s_0)\, d_{0,\boldsymbol{X}_{\sigma^{-1}(A)}} - \sum_{j=1}^{B} s_j\Big).$$

  As the cohomologies related to the blocks $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_B$ do not vanish, we have

  - For $j \in P$, $\mathcal{E}_j - \beta_j + d_{0,\boldsymbol{Y}_j} \geq s_j + s_0\, d_{0,\boldsymbol{Y}_j}$ (Eq. (6.7)), and
  - For each $j \in D$, $\mathcal{E}_j \geq s_j$.

  By definition, it holds $N = \sum_{j=1}^{B}\mathcal{E}_j = \sum_{i=1}^{A}\alpha_i + \sum_{j\in P}\beta_j + \sum_{j\in D}\beta_j$. Hence, adding the inequalities we obtain

$$\sum_{i=1}^{A}\alpha_i + \sum_{j\in D}\beta_j = \sum_{j=1}^{B}\mathcal{E}_j - \sum_{j\in P}\beta_j \geq \sum_{j=1}^{B} s_j + (s_0-1)\sum_{j\in P} d_{0,\boldsymbol{Y}_j}. \tag{6.12}$$

  As we assumed that $c < A$ and the triplet (P, D, c) is *determinantal data* (Def. 6.1.6), either $d_{0,\boldsymbol{X}_{\sigma^{-1}(A)}} = 1$ or both $d_{0,\boldsymbol{X}_{\sigma^{-1}(A)}} = 0$ and $\sum_{j\in P} d_{0,\boldsymbol{Y}_j} = 0$. Also it holds $0 \leq s_0, \sum_{j\in P} d_{0,\boldsymbol{Y}_j} \leq 1$, we conclude that

$$\sum_{j\in D}\beta_j + \sum_{i=1}^{A-1}\alpha_{\sigma^{-1}(i)} + (1-s_0)\, d_{0,\boldsymbol{X}_{\sigma^{-1}(A)}} - \sum_{j=1}^{B} s_j \geq -\alpha_{\sigma^{-1}(A)}$$

  Hence, by Cor. 3.3.7,

$$\begin{aligned} &H^{r_{\boldsymbol{X}_{\sigma^{-1}(A)}}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(A)}}}\Big(\sum_{j\in D}\beta_j + \sum_{i=1}^{A-1}\alpha_{\sigma^{-1}(i)} + (1-s_0)\, d_{0,\boldsymbol{X}_{\sigma^{-1}(A)}} - \sum_{j=1}^{B} s_j\Big) \neq 0 \iff \\ &r_{\boldsymbol{X}_{\sigma^{-1}(A)}} = 0 \text{ and } \sum_{j\in D}\beta_j + \sum_{i=1}^{A-1}\alpha_{\sigma^{-1}(i)} \geq \sum_{j=1}^{B} s_j + (s_0-1)\; d_{0,\boldsymbol{X}_{\sigma^{-1}(A)}}. \end{aligned} \tag{6.13}$$

- We proceed by induction, assuming that Eq. (6.11) holds for $k+1$, we prove the property for $k > c$. We consider the cohomology

$$H^{r_{\boldsymbol{X}_{\sigma^{-1}(k)}}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(k)}}}\Big(\sum_{j\in D}\beta_j + \sum_{i=1}^{k-1}\alpha_{\sigma^{-1}(i)} + (1-s_0)\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} - \sum_{j=1}^{B} s_j\Big).$$

By definition (Eq. (6.1)), $d_{0,\boldsymbol{X}_{\sigma^{-1}(k+1)}} = d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}}$. So, if the cohomologies related to the blocks $\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_B,\boldsymbol{X}_{\sigma^{-1}(k+1)},\ldots,\boldsymbol{X}_{\sigma^{-1}(A)}$ do not vanish, by induction hypothesis,

$$\sum_{j\in D}\beta_j + \sum_{i=1}^{k}\alpha_{\sigma^{-1}(i)} \geq (s_0-1)\ d_{0,\boldsymbol{X}_{\sigma^{-1}(k+1)}} + \sum_{j=1}^{B} s_j = (s_0-1)\ d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} + \sum_{j=1}^{B} s_j$$

Equivalently,

$$\sum_{j\in D}\beta_j + \sum_{i=1}^{k-1}\alpha_{\sigma^{-1}(i)} + (1-s_0)\ d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} - \sum_{j=1}^{B} s_j \geq -\alpha_{\sigma^{-1}(k)}.$$

Hence, by Cor. 3.3.7,

$$H^{r_{\boldsymbol{X}_{\sigma^{-1}(k)}}}_{\mathbb{P}^{\alpha_{\sigma^{-1}(k)}}}\Big(\sum_{j\in D}\beta_j + \sum_{i=1}^{k-1}\alpha_{\sigma^{-1}(i)} + (1-s_0)\, d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}} - \sum_{j=1}^{B} s_j\Big) \neq 0 \iff$$

$$r_{\boldsymbol{X}_{\sigma^{-1}(k)}} = 0 \quad \text{and} \quad \sum_{j\in D}\beta_j + \sum_{i=1}^{k-1}\alpha_{\sigma^{-1}(i)} \geq \sum_{j=1}^{B} s_j + (s_0-1)\ d_{0,\boldsymbol{X}_{\sigma^{-1}(k)}}.$$

We proved that, if the cohomologies in Eq. (6.5) do not vanish then Eq. (6.6) holds. We study the possible values for v such that $K_{v,p}(\boldsymbol{m})$ does not vanish. From Eq. (6.4), it holds $p = \sum_{j=1}^{B} s_j + s_0$. By Prop. 3.3.5, $p - v = \sum_{i=1}^{A} r_{\boldsymbol{X}_i} + \sum_{j=1}^{B} r_{\boldsymbol{Y}_j}$. Hence, we deduce that, when $K_{v,p}(\boldsymbol{m})$ does not vanish, it holds,

$$v = \sum_{j=1}^{B} s_j + s_0 - \sum_{j\in D}\beta_j - \sum_{i=1}^{c}\alpha_{\sigma^{-1}(i)}.$$

We bound the values for v for which $K_{v,p}(\boldsymbol{m})$ does not vanish.

- First we lower-bound v. Assume that $c > 0$. By Eq. (6.9), if we consider $k = c$,

$$\sum_{j=1}^{B} s_j + s_0\, d_{0,\boldsymbol{X}_{\sigma^{-1}(c)}} \geq \sum_{j\in D}\beta_j + \sum_{i=1}^{c}\alpha_{\boldsymbol{X}_{\sigma^{-1}(i)}}.$$

Hence, as $0 \leq s_0, d_{0,\boldsymbol{X}_{\sigma^{-1}(c+1)}} \leq 1$, we conclude

$$v = s_0 + \sum_{j=1}^{B} s_j - \sum_{j\in D}\beta_j - \sum_{i=1}^{c}\alpha_{\sigma^{-1}(i)} \geq s_0\,(1 - d_{0,\boldsymbol{X}_{\sigma^{-1}(c)}}) \geq 0.$$

If $c = 0$, then $v = \sum_{j=1}^{B} s_j + s_0 - \sum_{j\in D} \beta_j$. As (P, D, c) is determinantal data, then $0 \leq \sum_{j\in D} d_{0,\boldsymbol{Y}_j} \leq 1$. Hence, by Eq. (6.10),

$$v \geq s_0 - s_0 \sum_{j\in D} d_{0,\boldsymbol{Y}_j} \geq 0$$

- Finally we upper-bound v. Assume that $c < A$. By Eq. (6.11), if we consider $k = c + 1$, then

$$\sum_{j\in D} \beta_j + \sum_{i=1}^{c} \alpha_{\sigma^{-1}(i)} \geq \sum_{j=1}^{B} s_j + (s_0 - 1)\ d_{0,\boldsymbol{X}_{\sigma^{-1}(c+1)}}.$$

Hence, as $0 \leq s_0, d_{0,\boldsymbol{X}_{\sigma^{-1}(c+1)}} \leq 1$,

$$v = s_0 + \sum_{j=1}^{B} s_j - \sum_{j\in D} \beta_j - \sum_{i=1}^{c} \alpha_{\sigma^{-1}(i)} \leq s_0 + (1 - s_0)\, d_{0,\boldsymbol{X}_{\sigma^{-1}(c+1)}} \leq 1.$$

If $c = A$, then $v = \sum_{j=1}^{B} s_j + s_0 - \sum_{j\in D} \beta_j - \sum_{i=1}^{A} \alpha_i$. As $0 \leq s_0, \sum_{j\in P} d_{0,\boldsymbol{Y}_j} \leq 1$, by Eq. (6.12),

$$v \leq s_0 - (s_0 - 1) \sum_{j\in P} d_{0,\boldsymbol{Y}_j} \leq 1.$$

We conclude that the possible values for $v, p, r_{\boldsymbol{X}_1}, \ldots, r_{\boldsymbol{X}_A}, r_{\boldsymbol{Y}_1}, \ldots, r_{\boldsymbol{Y}_B}$ such that Eq. (6.5) is not zero are $v \in \{0, 1\}$, the possible values for $r_{\boldsymbol{X}_1}, \ldots, r_{\boldsymbol{X}_A}, r_{\boldsymbol{Y}_1}, \ldots, r_{\boldsymbol{Y}_B}$ are the ones in Eq. (6.6) and $p = \sum_{i=1}^{c} \alpha_{\sigma^{-1}(i)} + \sum_{j\in D} \beta_j + v$. Let $\omega = \sum_{i=1}^{c} \alpha_{\sigma^{-1}(i)} + \sum_{j\in D} \beta_j$. Hence, our Weyman complex looks like Eq. (3.10), where

$$\delta_1(\boldsymbol{m}) : K_{1,\omega+1}(\boldsymbol{m}) \otimes \mathbb{K}[\boldsymbol{u}] \to K_{0,\omega}(\boldsymbol{m}) \otimes \mathbb{K}[\boldsymbol{u}]$$

is a Koszul-type determinantal formula. □

Remark 6.1.9. *Consider a degree vector $\boldsymbol{m}$ related to the* determinantal data (P, D, c) *and a permutation σ. Then, the triplet $(D, P, A - c)$ is also* determinantal data *and the map $\bar{\sigma}$ such that $\bar{\sigma}(i) = (A + 1 - \sigma(i))$ is a permutation of $\{1, \ldots, A\}$. Let $\bar{\boldsymbol{m}}$ be the degree vector associated to $(D, P, A - c)$ and $\bar{\sigma}$, then, by Prop. 3.3.8, $K_\bullet(\boldsymbol{m})$ is isomorphic to the dual complex of $K_\bullet(\bar{\boldsymbol{m}})$.*

Corollary 6.1.10 (Sylvester-type formulas)**.** *Consider $\boldsymbol{d}_0$ such that $d_{0,\boldsymbol{X}_1} = 1$ and $\sum_{j\in D} d_{0,\boldsymbol{Y}_j} = 0$ (Cases 1 and 3). Let $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ be any permutation. Note that the triplet $(\{1, \ldots, B\}, \emptyset, 0)$ is* determinantal data*. In this case, the overdetermined systems from Thm. 6.1.8 have a Sylvester-type formula coming from the degree vector $\boldsymbol{m}$ related to the triplet $(\{1, \ldots, B\}, \emptyset, 0)$ and the permutation σ.*

Size of the determinantal formulas

In this subsection we study the size of the determinantal formulas of Thm. 6.1.8 and we compare them with the number of solutions of the square system $(f_1, \dots, f_N)$.

The multihomogeneous Bézout bound (Prop. 2.10.9) implies the following lemma.

Lemma 6.1.11. *The expected number of solutions, Υ, of a square star multilinear system is*

$$\Upsilon := \frac{(\sum_{i=1}^{A} \alpha_i)!}{\prod_{i=1}^{A} \alpha_i!} \cdot \prod_{j=1}^{B} \binom{\mathcal{E}_j}{\beta_j}.$$

Lemma 6.1.12. *The sizes of the matrices corresponding to the determinantal formulas of Thm. 6.1.8, that is, the rank of the modules $K_0(\boldsymbol{m})$ and $K_1(\boldsymbol{m})$, are as follow:*

1. *If $f_0 \in \mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A]_{\mathbf{1}}$, then $Rank(K_0(\boldsymbol{m})) = Rank(K_1(\boldsymbol{m})) = \Upsilon \cdot \Big(1 + \sum_{i=1}^{A} \alpha_i\Big)$.*
2. *If $j \in [B]$ and $f_0 \in \mathbb{K}[\boldsymbol{Y}_j]_{\mathbf{1}}$, then $Rank(K_0(\boldsymbol{m})) = Rank(K_1(\boldsymbol{m})) = \Upsilon \cdot \frac{\mathcal{E}_j + \beta_j \, (\sum_{i=1}^{A} \alpha_i) + 1}{\mathcal{E}_j - \beta_j + 1}$.*
3. *If $j \in [B]$ and $f_0 \in \mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A, \boldsymbol{Y}_j]_{\mathbf{1}}$, then*

$$Rank(K_0(\boldsymbol{m})) = Rank(K_1(\boldsymbol{m})) = \Upsilon \cdot \frac{(1 + \sum_{i=1}^{A} \alpha_i)(\mathcal{E}_j + 1)}{\mathcal{E}_j - \beta_j + 1}.$$

4. *If $j_1, j_2 \in \{1, \dots, B\}$ and $j_1 \neq j_2$, $f_0 \in \mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A, \boldsymbol{Y}_{j_1}, \boldsymbol{Y}_{j_2}]_{\mathbf{1}}$, then*

$$Rank(K_0(\boldsymbol{m})) = Rank(K_1(\boldsymbol{m})) = \Upsilon \cdot \Big(1 + \sum_{i=1}^{A} \alpha_i\Big) \Big(1 + \frac{\beta_{j_1}}{\mathcal{E}_{j_1} - \beta_{j_1} + 1} + \frac{\beta_{j_2}}{\mathcal{E}_{j_2} - \beta_{j_2} + 1}\Big).$$

Proof. Following Cor. 3.3.16, the size of the Koszul determinantal matrix is the degree of the resultant. We use Prop. 3.3.2 to compute this degree in various cases. For each $j \in \{1, \dots, B\}$, let $I_j \in \{1, \dots, N\}$ be the index of a polynomial in $\boldsymbol{F}$ such that $F_j \in \mathbb{K}[\boldsymbol{X}_1, \dots, \boldsymbol{X}_A, \boldsymbol{Y}_{I_j}]$. Moreover, $\mathcal{E}_j$ is the number of polynomials with multidegree equal to $\boldsymbol{d}_{I_j}$ and so we can rewrite the degree of the resultant from Prop. 3.3.2 as

$$\deg(\mathtt{res}(\boldsymbol{d}_0, \dots, \boldsymbol{d}_n)) = \mathtt{MHB}(\boldsymbol{d}_1, \dots, \boldsymbol{d}_n) + \sum_{j=1}^{B} \mathcal{E}_j \, \mathtt{MHB}(\boldsymbol{d}_0, \dots, \boldsymbol{d}_{I_j - 1}, \boldsymbol{d}_{I_j + 1}, \dots, \boldsymbol{d}_n).$$

From Lem. 6.1.11, $\mathtt{MHB}(\boldsymbol{d}_1, \dots, \boldsymbol{d}_n) = \frac{(\sum_{i=1}^{A} \alpha_i)!}{\prod_{i=1}^{A} \alpha_i!} \cdot \prod_{j=1}^{B} \binom{\mathcal{E}_j}{\beta_j} = \Upsilon$. By Prop. 2.10.9, for every $1 \leq j \leq B$, $\mathtt{MHB}(\boldsymbol{d}_0, \dots, \boldsymbol{d}_{I_j - 1}, \boldsymbol{d}_{I_j + 1}, \dots, \boldsymbol{d}_n)$ is the coefficient of $(\prod_{i=1}^{A} Z_{\boldsymbol{X}_i}^{\alpha_i})(\prod_{t=1}^{B} Z_{\boldsymbol{Y}_t}^{\beta_t})$ in

$$\Big(\sum_{i=1}^{A} d_{0,\boldsymbol{X}_i} Z_{\boldsymbol{X}_i} + \sum_{t=1}^{B} d_{0,\boldsymbol{Y}_t} Z_{\boldsymbol{X}_t}\Big)\Big(\sum_{i=1}^{A} Z_{\boldsymbol{X}_i} + Z_{\boldsymbol{Y}_j}\Big)^{\mathcal{E}_j - 1} \prod_{k \in \{1, \dots, B\} \setminus \{j\}} \Big(\sum_{i=1}^{A} Z_{\boldsymbol{X}_i} + Z_{\boldsymbol{Y}_k}\Big)^{\mathcal{E}_k}.$$

Consider the last two factors of the previous equation, that is

$$\Big(\sum_{i=1}^{A} Z_{\boldsymbol{X}_i} + Z_{\boldsymbol{Y}_j}\Big)^{\mathcal{E}_j-1} \prod_{k\in\{1,\dots,B\}\setminus\{j\}} \Big(\sum_{i=1}^{A} Z_{\boldsymbol{X}_i} + Z_{\boldsymbol{Y}_k}\Big)^{\mathcal{E}_k}. \tag{6.14}$$

Then

$$\mathtt{MHB}(\boldsymbol{d}_0,\dots,\boldsymbol{d}_{I_j-1},\boldsymbol{d}_{I_j+1},\dots,\boldsymbol{d}_n) = \sum_{s=1}^{A} d_{0,\boldsymbol{X}_s}\theta^X_{j,s} + \sum_{l=1}^{B} d_{0,\boldsymbol{Y}_l}\theta^Y_{j,l},$$

where $\theta^X_{j,s}$ is the coefficient of $\frac{(\prod_{i=1}^{A} Z_{\boldsymbol{X}_i}^{\alpha_i})(\prod_{t=1}^{B} Z_{\boldsymbol{Y}_t}^{\beta_t})}{Z_{\boldsymbol{X}_s}}$ in Eq. (6.14), and $\theta^Y_{j,t}$ is the coefficient of $\frac{(\prod_{i=1}^{A} Z_{\boldsymbol{X}_i}^{\alpha_i})(\prod_{l=1}^{B} Z_{\boldsymbol{Y}_l}^{\beta_t})}{Z_{\boldsymbol{Y}_t}}$ in Eq. (6.14). After some computations, we have

$$\theta^X_{j,s} = \frac{((\sum_{i=1}^{A}\alpha_i)-1)!}{(\alpha_s-1)!\prod_{i\in\{1,\dots,A\}\setminus\{s\}}\alpha_i!}\binom{\mathcal{E}_j-1}{\beta_j}\prod_{k\in\{1,\dots,B\}\setminus\{j\}}\binom{\mathcal{E}_k}{\beta_k} = \Upsilon\cdot\frac{\alpha_s}{\sum_{i=1}^{A}\alpha_i}\frac{\mathcal{E}_j-\beta_j}{\mathcal{E}_j},$$

$$\theta^Y_{j,t} = \begin{cases} \frac{(\sum_{i=1}^{A}\alpha_i)!}{\prod_{i=1}^{A}\alpha_i!}\binom{\mathcal{E}_j-1}{\beta_j-1}\prod_{k\in\{1,\dots,B\}\setminus\{j\}}\binom{\mathcal{E}_k}{\beta_k} = \Upsilon\cdot\frac{\beta_j}{\mathcal{E}_j} & \text{if } t=j,\\ \frac{(\sum_{i=1}^{A}\alpha_i)!}{\prod_{i=1}^{A}\alpha_i!}\binom{\mathcal{E}_t}{\beta_t-1}\binom{\mathcal{E}_j-1}{\beta_j}\prod_{k\in\{1,\dots,B\}\setminus\{t,j\}}\binom{\mathcal{E}_k}{\beta_k} = \Upsilon\cdot\frac{\beta_t}{\mathcal{E}_t-\beta_t+1}\frac{\mathcal{E}_j-\beta_j}{\mathcal{E}_j} & \text{otherwise.}\end{cases}$$

Using the formulas for $\theta^X_{j,s}$ and $\theta^Y_{j,t}$ we get

$$\deg(\mathtt{res}(\boldsymbol{d}_0,\dots,\boldsymbol{d}_n)) = \mathtt{MHB}(\boldsymbol{d}_1,\dots,\boldsymbol{d}_n) + \sum_{j=1}^{B}\mathcal{E}_j\,\mathtt{MHB}(\boldsymbol{d}_0,\dots,\boldsymbol{d}_{I_j-1},\boldsymbol{d}_{I_j+1},\dots,\boldsymbol{d}_n) =$$

$$\Upsilon + \sum_{j=1}^{B}\mathcal{E}_j\,(\sum_{s=1}^{A} d_{0,\boldsymbol{X}_s}\theta^X_{j,s} + \sum_{l=1}^{B} d_{0,\boldsymbol{Y}_t}\theta^Y_{j,t}) = \Upsilon + \sum_{j=1}^{B}\mathcal{E}_j\,(\sum_{s=1}^{A} d_{0,\boldsymbol{X}_s}\theta^X_{j,s}) + \sum_{j=1}^{B}\mathcal{E}_j\,(\sum_{l=1}^{B} d_{0,\boldsymbol{Y}_t}\theta^Y_{j,t}).$$

Next, we simplify the last two summands of the previous equation. For the first one, as $\sum_{i=1}^{A}\alpha_i = \sum_{j=1}^{B}(\mathcal{E}_j-\beta_j)$ and for all s it holds $d_{0,\boldsymbol{X}_s} = d_{0,\boldsymbol{X}_1}$, we obtain

$$\sum_{j=1}^{B}\mathcal{E}_j\,\Big(\sum_{s=1}^{A} d_{0,\boldsymbol{X}_s}\theta^X_{j,s}\Big) = \sum_{j=1}^{B}\mathcal{E}_j\,\Big(\sum_{s=1}^{A} d_{0,\boldsymbol{X}_s}\Upsilon\frac{\alpha_s}{\sum_{i=1}^{A}\alpha_i}\frac{\mathcal{E}_j-\beta_j}{\mathcal{E}_j}\Big) = \Upsilon\, d_{0,\boldsymbol{X}_1}\sum_{i=1}^{A}\alpha_i.$$

For the second one, we perform the following direct calculations

$$\sum_{j=1}^{B}\mathcal{E}_j\,(\sum_{l=1}^{B}d_{0,\boldsymbol{Y}_t}\theta^{Y}_{j,t}) = \sum_{j=1}^{B}\mathcal{E}_j\,\Big(\sum_{t\in\{1,\ldots,B\}\setminus\{j\}}d_{0,\boldsymbol{Y}_t}\Upsilon\frac{\beta_t}{\mathcal{E}_t-\beta_t+1}\frac{\mathcal{E}_j-\beta_j}{\mathcal{E}_j}+d_{0,\boldsymbol{Y}_j}\Upsilon\frac{\beta_j}{\mathcal{E}_j}\Big) =$$

$$\Upsilon\sum_{j=1}^{B}\Big(\sum_{t=1}^{B}d_{0,\boldsymbol{Y}_t}\cdot\frac{\beta_t}{\mathcal{E}_t-\beta_t+1}(\mathcal{E}_j-\beta_j)-d_{0,\boldsymbol{Y}_j}\cdot\frac{\beta_j(\mathcal{E}_j-\beta_j)}{\mathcal{E}_j-\beta_j+1}+d_{0,\boldsymbol{Y}_j}\beta_j\Big) =$$

$$\Upsilon\sum_{j=1}^{B}(\mathcal{E}_j-\beta_j)\Big(\sum_{t=1}^{B}d_{0,\boldsymbol{Y}_t}\cdot\frac{\beta_t}{\mathcal{E}_t-\beta_t+1}\Big)+\Upsilon\sum_{j=1}^{B}d_{0,\boldsymbol{Y}_j}\cdot\frac{\beta_j}{\mathcal{E}_j-\beta_j+1} =$$

$$\Upsilon\,(1+\sum_{i=1}^{A}\alpha_i)\Big(\sum_{t=1}^{B}d_{0,\boldsymbol{Y}_t}\cdot\frac{\beta_t}{\mathcal{E}_t-\beta_t+1}\Big).$$

At last we have the formula

$$\deg(\mathtt{res}(\boldsymbol{d}_0,\ldots,\boldsymbol{d}_n)) = \Upsilon\Big(1+d_{0,\boldsymbol{X}_1}\sum_{i=1}^{A}\alpha_i+(1+\sum_{i=1}^{A}\alpha_i)\Big(\sum_{t=1}^{B}d_{0,\boldsymbol{Y}_t}\cdot\frac{\beta_t}{\mathcal{E}_t-\beta_t+1}\Big)\Big).$$

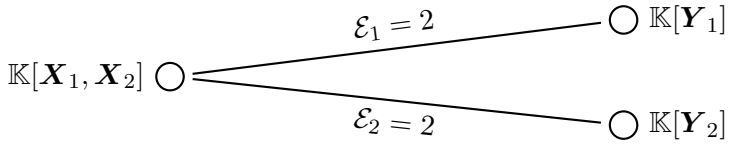
We complete the proof by instantiating the variables $d_{0,\boldsymbol{X}_1}, d_{0,\boldsymbol{Y}_1},\ldots,d_{0,\boldsymbol{Y}_B}$ in the previous formula according to each f_0. □

Example

We follow the notation from the beginning of Sec. 6.1.1. Consider four blocks of variables that we partition to two sets of cardinalities $A=2$ and $B=2$. The number of variables in the blocks of the first set are $\alpha=(1,1)$ and in the second $\beta=(1,1)$. That is, we consider

$$X_1 := \{X_{1,0}, X_{1,1}\}, X_2 := \{X_{2,0}, X_{2,1}\},\quad Y_1 := \{Y_{1,0}, Y_{1,1}\}, Y_2 := \{Y_{2,0}, Y_{2,1}\}.$$

Let $(f_1,\ldots,f_4)$ be a square star multilinear system corresponding to the following graph,



The expected number of solutions of the system is 8.

If we introduce a polynomial f_0 and we consider the multiprojective resultant of $(f_0,f_1,\ldots,f_N)$, then the degree of the resultant, depending on the choice of f_0, is as follows:

- If $f_0\in\mathbb{K}[\boldsymbol{X}_1,\boldsymbol{X}_2]_{\mathbf{1}}$, then the degree of the resultant is 24.
- If $f_0\in\mathbb{K}[\boldsymbol{Y}_j]_{\mathbf{1}}$, where $j\in\{1,2\}$, then the degree of the resultant is 20.
- If $f_0\in\mathbb{K}[\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{Y}_j]_{\mathbf{1}}$, where $j\in\{1,2\}$, then the degree of the resultant is 36.

- If $f_0 \in \mathbb{K}[\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}_1, \boldsymbol{Y}_2]_{\mathbf{1}}$, then the degree of the resultant is 48.

Let $f_0 \in \mathbb{K}[\boldsymbol{X}_1, \boldsymbol{X}_2]_{\mathbf{1}}$ and consider the overdetermined system $\boldsymbol{f}$,

$$\boldsymbol{f} := \left\{ \begin{array}{ll} f_0 := & (a_1\, x_{2;0} + a_2\, x_{2;1})\, x_{1;0} + (a_3\, x_{2;0} + a_4\, x_{2;1})\, x_{1;1} \\ f_1 := & ((b_1\, y_{1;0} + b_2\, y_{1;1})\, x_{2;0} + (b_3\, y_{1;0} + b_4\, y_{1;1})\, x_{2;1})\, x_{1;0} \\ & + ((b_5\, y_{1;0} + b_6\, y_{1;1})\, x_{2;0} + (b_7\, y_{1;0} + b_8\, y_{1;1})\, x_{2;1})\, x_{1;1} \\ f_2 := & ((c_1\, y_{1;0} + c_2\, y_{1;1})\, x_{2;0} + (c_3\, y_{1;0} + c_4\, y_{1;1})\, x_{2;1})\, x_{1;0} \\ & + ((c_5\, y_{1;0} + c_6\, y_{1;1})\, x_{2;0} + (c_7\, y_{1;0} + c_8\, y_{1;1})\, x_{2;1})\, x_{1;1} \\ f_3 := & ((d_1\, y_{2;0} + d_2\, y_{2;1})\, x_{2;0} + (d_3\, y_{2;0} + d_4\, y_{2;1})\, x_{2;1})\, x_{1;0} \\ & + ((d_5\, y_{2;0} + d_6\, y_{2;1})\, x_{2;0} + (d_7\, y_{2;0} + d_8\, y_{2;1})\, x_{2;1})\, x_{1;1} \\ f_4 := & ((e_1\, y_{2;0} + e_2\, y_{2;1})\, x_{2;0} + (e_3\, y_{2;0} + e_4\, y_{2;1})\, x_{2;1})\, x_{1;0} \\ & + ((e_5\, y_{2;0} + e_6\, y_{2;1})\, x_{2;0} + (e_7\, y_{2;0} + e_8\, y_{2;1})\, x_{2;1})\, x_{1;1} \end{array} \right. .$$

We consider the *determinantal data* $(\{1\}, \{2\}, 1)$ and the identity map $i \mapsto i$. Then, the degree vector of Thm. 6.1.8 is $\boldsymbol{m} = (0, 3, 1, -1)$ and the vector spaces of the Weyman complex $K(\boldsymbol{m}, \boldsymbol{f})$ become

$$\begin{aligned} K_1(\boldsymbol{m}, \boldsymbol{f}) = \quad & S^*_{\boldsymbol{X}_1}(-1) \otimes S_{\boldsymbol{X}_2}(0) \otimes S_{\boldsymbol{Y}_1}(0) \otimes S^*_{\boldsymbol{Y}_2}(0) \otimes \left\{ \begin{array}{l} (e_0 \wedge e_1 \wedge e_3) \oplus (e_0 \wedge e_1 \wedge e_4) \oplus \\ (e_0 \wedge e_2 \wedge e_3) \oplus (e_0 \wedge e_2 \wedge e_4) \end{array} \right\} \\ & \oplus\ S^*_{\boldsymbol{X}_1}(-1) \otimes S_{\boldsymbol{X}_2}(0) \otimes S_{\boldsymbol{Y}_1}(0) \otimes S^*_{\boldsymbol{Y}_2}(-1) \otimes \left\{ \ (e_1 \wedge e_3 \wedge e_4) \oplus (e_2 \wedge e_3 \wedge e_4) \ \right\} \\ & \oplus\ S^*_{\boldsymbol{X}_1}(-1) \otimes S_{\boldsymbol{X}_2}(0) \otimes S_{\boldsymbol{Y}_1}(1) \otimes S^*_{\boldsymbol{Y}_2}(-1) \otimes \left\{ \ (e_0 \wedge e_3 \wedge e_4) \ \right\}, \\ K_0(\boldsymbol{m}, \boldsymbol{f}) = \quad & S^*_{\boldsymbol{X}_1}(0) \otimes S_{\boldsymbol{X}_2}(1) \otimes S_{\boldsymbol{Y}_1}(0) \otimes S^*_{\boldsymbol{Y}_2}(0) \otimes \left\{ \begin{array}{l} (e_1 \wedge e_3) \oplus (e_1 \wedge e_4) \oplus \\ (e_2 \wedge e_3) \oplus (e_2 \wedge e_4) \end{array} \right\} \\ & \oplus\ S^*_{\boldsymbol{X}_1}(0) \otimes S_{\boldsymbol{X}_2}(1) \otimes S_{\boldsymbol{Y}_1}(1) \otimes S^*_{\boldsymbol{Y}_2}(0) \otimes \left\{ \ (e_0 \wedge e_3) \oplus (e_0 \wedge e_4) \ \right\} \\ & \oplus\ S^*_{\boldsymbol{X}_1}(0) \otimes S_{\boldsymbol{X}_2}(1) \otimes S_{\boldsymbol{Y}_1}(1) \otimes S^*_{\boldsymbol{Y}_2}(-1) \otimes \left\{ \ (e_3 \wedge e_4) \ \right\}. \end{aligned}$$

The Koszul determinantal matrix representing the map $\delta_1(\boldsymbol{m}, \boldsymbol{f})$ between the modules with respect to

the monomial basis is

$$\begin{bmatrix}
a_1\ a_2 & & & & & & & & -b_1\ -b_2\ -b_3\ -b_4 & & & & & & & & & & & & & & & \\
a_3\ a_4 & & & & & & & & -b_5\ -b_6\ -b_7\ -b_8 & & & & & & & & & & & & & & & \\
 & a_1\ a_2 & & & & & & & & & & & -b_1\ -b_2\ -b_3\ -b_4 & & & & & & & & & & & \\
 & a_3\ a_4 & & & & & & & & & & & -b_5\ -b_6\ -b_7\ -b_8 & & & & & & & & & & & \\
 & & a_1\ a_2 & & & & & & -c_1\ -c_2\ -c_3\ -c_4 & & & & & & & & & & & & & & & \\
 & & a_3\ a_4 & & & & & & -c_5\ -c_6\ -c_7\ -c_8 & & & & & & & & & & & & & & & \\
 & & & & a_1\ a_2 & & & & & & & & -c_1\ -c_2\ -c_3\ -c_4 & & & & & & & & & & & \\
 & & & & a_3\ a_4 & & & & & & & & -c_5\ -c_6\ -c_7\ -c_8 & & & & & & & & & & & \\
e_1\ e_3\ -d_1\ -d_3 & & & & & & & & & & & & & & & & b_1\ b_2\ b_3\ b_4 & & & & & & & \\
e_2\ e_4\ -d_2\ -d_4 & & & & & & & & & & & & & & & & & & & & b_1\ b_2\ b_3\ b_4 & & & \\
e_5\ e_7\ -d_5\ -d_7 & & & & & & & & & & & & & & & & b_5\ b_6\ b_7\ b_8 & & & & & & & \\
e_6\ e_8\ -d_6\ -d_8 & & & & & & & & & & & & & & & & & & & & b_5\ b_6\ b_7\ b_8 & & & \\
 & & & e_1\ e_3\ -d_1\ -d_3 & & & & & & & & & & & & & c_1\ c_2\ c_3\ c_4 & & & & & & & \\
 & & & e_2\ e_4\ -d_2\ -d_4 & & & & & & & & & & & & & & & & & & c_1\ c_2\ c_3\ c_4 & & & \\
 & & & e_5\ e_7\ -d_5\ -d_7 & & & & & & & & & & & & & c_5\ c_6\ c_7\ c_8 & & & & & & & \\
 & & & e_6\ e_8\ -d_6\ -d_8 & & & & & & & & & & & & & & & & & & c_5\ c_6\ c_7\ c_8 & & & \\
 & & & & & & & e_1 & & e_3 & & -d_1 & & -d_3 & & a_1 & & a_2 & & & & & & & \\
 & & & & & & & e_2 & & e_4 & & -d_2 & & -d_4 & & & & & & a_1 & & a_2 & & & \\
 & & & & & & & & e_1 & & e_3 & & -d_1 & & -d_3 & a_1 & & a_2 & & & & & & \\
 & & & & & & & & e_2 & & e_4 & & -d_2 & & -d_4 & & & & & & a_1 & & a_2 \\
 & & & & & & & e_5 & & e_7 & & -d_5 & & -d_7 & & a_3 & & a_4 & & & & & & \\
 & & & & & & & e_6 & & e_8 & & -d_6 & & -d_8 & & & & & & a_3 & & a_4 & & \\
 & & & & & & & & e_5 & & e_7 & & -d_5 & & -d_7 & a_3 & & a_4 & & & & & & \\
 & & & & & & & & e_6 & & e_8 & & -d_6 & & -d_8 & & & & & & a_3 & & a_4
\end{bmatrix}.$$

6.1.3 Determinantal formulas for bipartite bilinear systems

We consider two different kinds of overdetermined multihomogeneous systems, related to *bipartite bilinear systems* (Def. 6.1.3) and we construct determinantal formulas for each of them. These formulas are Koszul- and Sylvester-type determinantal formulas (Def. 3.3.14). We study overdetermined polynomial systems $(f_0, f_1, \ldots, f_N)$ in $\mathbb{K}[\bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}]$ where $(f_1, \ldots, f_N)$ is a square bipartite bilinear system and f_0 is a multilinear polynomial.

We consider different types of polynomials f_0. The obvious choice for f_0 is to have the same structure as one of the polynomials $f_1, \ldots, f_N$; still we also choose f_0 to have a different support. This leads to resultants of smaller degrees and so to matrices of smaller size. The following f_0 lead to determinantal formulas:

1. $f_0 \in \mathbb{K}[\boldsymbol{X}_i]_{\mathbf{1}}$, for any $i \in \{1, \ldots, A\}$.
2. $f_0 \in \mathbb{K}[\boldsymbol{Y}_j]_{\mathbf{1}}$, for any $j \in \{1, \ldots, B\}$.
3. $f_0 \in \mathbb{K}[\boldsymbol{X}_i, \boldsymbol{Y}_j]_{\mathbf{1}}$, for any $i \in \{1, \ldots, A\}$ and $j \in \{1, \ldots, B\}$.

We can interpret the various multidegrees of f_0, $\boldsymbol{d}_0 = (d_{0,X_1}, \ldots, d_{0,X_A}, d_{0,Y_1}, \ldots, d_{0,Y_B})$, that lead to determinantal formulas as solutions of the following system of inequalities:

$$\begin{cases}
(\forall 1 \le i \le A) & 0 \le d_{0,\boldsymbol{X}_i} \le 1 \\
(\forall 1 \le j \le B) & 0 \le d_{0,\boldsymbol{Y}_j} \le 1 \\
 & \sum_{i=1}^{A} d_{0,\boldsymbol{X}_i} \le 1 \\
 & \sum_{j=1}^{B} d_{0,\boldsymbol{Y}_j} \le 1.
\end{cases} \tag{6.15}$$

Consider the set $\{0,\dots,N\}$ that corresponds to generic polynomials $\boldsymbol{F}=(F_0,\dots,F_N)$ (Def. 3.3.1). For each tuple $s_0, s_{1,1},\dots,s_{A,B}\in\mathbb{N}$, let $\mathcal{I}_{s_0,s_{1,1},\dots,s_{A,B}}$ be the set of all the subsets of $\{0,\dots,N\}$, such that

- For $1\leq i\leq A$ and $1\leq j\leq B$, the index $s_{i,j}$ indicates that we consider $s_{i,j}$ polynomials from $\boldsymbol{F}$ that belong to $\mathbb{Z}[\boldsymbol{u}][\boldsymbol{X}_i,\boldsymbol{Y}_j]_{\mathbf{1}}$.
- In addition, if $s_0=1$, then 0 belongs to all the sets in $\mathcal{I}_{s_0,s_{1,1},\dots,s_{A,B}}$

That is,

$$\mathcal{I}_{s_0,s_{1,1},\dots,s_{A,B}}:=\Big\{I: I\subset\{0,\dots,n\},(0\in I\Leftrightarrow s_0=1)\text{ and }\\ (\forall 1\leq i\leq A)(\forall 1\leq j\leq B)\ s_{i,j}=\#\{k\in I: f_k\in\mathbb{K}[\boldsymbol{X}_i,\boldsymbol{Y}_j]\}\Big\}. \tag{6.16}$$

Notice that if $I,J\in\mathcal{I}_{s_0,s_{1,1},\dots,s_{A,B}}$, then I and J have the same cardinality and $\sum_{k\in I}\boldsymbol{d}_k=\sum_{k\in J}\boldsymbol{d}_k$, as they correspond to subsets of polynomials of $\boldsymbol{F}$ with the same supports.

The following lemma exploits the sets $\mathcal{I}_{s_0,s_{1,1},\dots,s_{A,B}}$ to rewrite the cohomologies of Eq. (3.8).

Lemma 6.1.13. *Consider a generic overdetermined system $\boldsymbol{F}=(F_0,\dots,F_N)$ in $\mathbb{K}[\boldsymbol{u}][\bar{\boldsymbol{X}},\bar{\boldsymbol{Y}}]$ of multidegrees $\boldsymbol{d}_0,\dots,\boldsymbol{d}_N$ (Def. 3.3.1), where $(F_1,\dots,F_n)$ is a square bipartite bilinear system such that, for each $i\in\{1,\dots,A\}$, $\sum_{j=1}^{B}\mathcal{E}_{i,j}\geq\alpha_i$ and for each $j\in\{1,\dots,B\}$, $\sum_{i=1}^{A}\mathcal{E}_{i,j}\geq\beta_j$, and $\boldsymbol{d}_0$ is the multidegree of F_0. Following Eq. (3.8) we can rewrite the modules of the Weyman complex $K_v(\boldsymbol{m})=\bigoplus_{p=0}^{N+1}K_{v,p}\otimes\mathbb{Z}[\boldsymbol{u}]$ in the more detailed form*

$$K_{v,p}(\boldsymbol{m})=\bigoplus_{\substack{0\leq s_0\leq 1\\(\forall 1\leq i\leq A)(\forall 1\leq j\leq B)\ 0\leq s_{i,j}\leq\mathcal{E}_{i,j}\\ s_0+\sum_{i=1}^{A}\sum_{j=1}^{B}s_{i,j}=p}}\Big(H_{\mathcal{P}}^{p-v}\Big(\boldsymbol{m}-(\sum_{j=1}^{B}s_{1,j},\dots,\sum_{j=1}^{B}s_{A,j},\sum_{i=1}^{A}s_{i,1},\dots,\sum_{i=1}^{A}s_{i,B})-s_0\,\boldsymbol{d}_0\Big)\otimes\bigoplus_{I\in\mathcal{I}_{s_0,s_{1,1},\dots,s_{A,B}}}\bigwedge_{k\in I}e_k\Big). \tag{6.17}$$

Moreover, the following isomorphisms hold for the cohomologies:

$$H_{\mathcal{P}}^{p-v}\Big(\boldsymbol{m}-(\sum_{j=1}^{B}s_{1,j},\dots,\sum_{j=1}^{B}s_{A,j},\sum_{i=1}^{A}s_{i,1},\dots,\sum_{i=1}^{A}s_{i,B})-s_0\,\boldsymbol{d}_0\Big)\cong\\ \bigoplus_{r_{\boldsymbol{X}_1}+\cdots+r_{\boldsymbol{X}_A}+r_{\boldsymbol{Y}_1}+\cdots+r_{\boldsymbol{Y}_B}=p-v}\left(\begin{array}{l}\bigotimes_{i=1}^{A}H_{\mathbb{P}^{\alpha_i}}^{r_{\boldsymbol{X}_i}}\Big(m_{\boldsymbol{X}_i}-\sum_{j=1}^{B}s_{i,j}-s_0)\,d_{0,X_i}\Big)\\ \otimes\bigotimes_{j=1}^{B}H_{\mathbb{P}^{\beta_j}}^{r_{\boldsymbol{Y}_j}}\Big(m_{\boldsymbol{Y}_j}-\sum_{i=1}^{A}s_{i,j}-d_{0,Y_j}s_0\Big)\end{array}\right). \tag{6.18}$$

The proof is similar, mutatis mutandis, to the one of Lem. 6.1.7.

Theorem 6.1.14. *Consider a generic overdetermined system $\boldsymbol{F}=(F_0,\dots,F_N)$ in $\mathbb{Z}[\boldsymbol{u}][\bar{\boldsymbol{X}},\bar{\boldsymbol{Y}}]$ of multidegrees $\boldsymbol{d}_0,\dots,\boldsymbol{d}_N$ (Def. 3.3.1), where $(F_1,\dots,F_n)$ is a square bipartite bilinear systems such that,*

for each $i \in \{1,\dots,A\}$, $\sum_{j=1}^{B} \mathcal{E}_{i,j} \geq \alpha_i$ *and for each* $j \in \{1,\dots,B\}$, $\sum_{i=1}^{A} \mathcal{E}_{i,j} \geq \beta_j$, *and* F_0 *is multilinear polynomial whose degree* $\boldsymbol{d}_0$ *is a solution of the system in Eq. 6.15.*

There are is a degree vector $\boldsymbol{m} = (m_{\boldsymbol{X}_1},\dots,m_{\boldsymbol{X}_A},m_{\boldsymbol{Y}_1},\dots,m_{\boldsymbol{Y}_B})$,

$$\begin{cases} m_{\boldsymbol{X}_i} = \sum_{j=1}^{B} \mathcal{E}_{i,j} - \alpha_i + d_{0,\boldsymbol{X}_i} & \text{for } 1 \leq i \leq A \\ m_{\boldsymbol{Y}_j} = -1 & \text{for } 1 \leq j \leq B \end{cases}$$

such that the Weyman complex of $K_\bullet(\boldsymbol{m})$ *reduces to*

$$K_\bullet(\boldsymbol{m}) : 0 \to K_{1,\omega+1}(\boldsymbol{m}) \xrightarrow{\delta_1(\boldsymbol{m})} K_{0,\omega}(\boldsymbol{m}) \to 0$$

where $\omega = \sum_{j=1}^{B} \beta_j$.

Hence, the map $\delta_1(\boldsymbol{m})$ *is a Koszul-type determinantal formula (Def. 3.3.14).*

Remark 6.1.15. *Let* $\boldsymbol{m}$ *be the degree vector in Thm. 6.1.14. Consider the degree vector* $\bar{\boldsymbol{m}} = (\bar{m}_{\boldsymbol{X}_1},\dots,\bar{m}_{\boldsymbol{X}_A},\bar{m}_{\boldsymbol{Y}_1},\dots,\bar{m}_{\boldsymbol{Y}_B})$, *where*

$$\begin{cases} \bar{m}_{\boldsymbol{X}_i} = -1 & \text{for } 1 \leq i \leq A \\ \bar{m}_{\boldsymbol{Y}_j} = \sum_{i=1}^{A} \mathcal{E}_{i,j} - \beta_j + d_{0,\boldsymbol{Y}_j} & \text{for } 1 \leq j \leq B \end{cases}$$

Note that $\boldsymbol{m} + \bar{\boldsymbol{m}} = \sum_{k=0}^{n} \boldsymbol{d}_k - (n_{x_1}+1,\dots,n_{x_A}+1,\beta_1+1,\dots,\beta_B+1)$. *Hence, by Prop. 3.3.8, the Weyman complex* $K_\bullet(\boldsymbol{m})$ *is the dual complex of* $K_\bullet(\bar{\boldsymbol{m}})$. *Therefore, the Weyman complex of* $K_\bullet(\bar{\boldsymbol{m}})$ *reduces to*

$$K_\bullet(\bar{\boldsymbol{m}}) : 0 \to K_{1,\sum_{i=1}^{A}\alpha_i+1}(\bar{\boldsymbol{m}}) \xrightarrow{\delta_1(\bar{\boldsymbol{m}})} K_{0,\sum_{i=1}^{A}\alpha_i}(\bar{\boldsymbol{m}}) \to 0.$$

Proof. This proof follows the same idea as the proof of Thm. 6.1.8. We rewrite Eq. (3.8) as in Lem. 6.1.13 and consider

$$H^{p-v}_{\mathcal{P}}\left(\boldsymbol{m} - \left(\sum_{j=1}^{B} s_{1,j},\dots,\sum_{j=1}^{B} s_{A,j},\sum_{i=1}^{A} s_{i,1},\dots,\sum_{i=1}^{A} s_{i,B}\right) - s_0\,\boldsymbol{d}_0\right) \cong$$
$$\bigoplus_{r_{\boldsymbol{X}_1}+\cdots+r_{\boldsymbol{X}_A}+r_{\boldsymbol{Y}_1}+\cdots+r_{\boldsymbol{Y}_B}=p-v} \begin{pmatrix} \bigotimes_{i=1}^{A} H^{r_{\boldsymbol{X}_i}}_{\mathbb{P}^{\alpha_i}}\left(\sum_{j=1}^{B}(\mathcal{E}_{i,j}-s_{i,j}) - \alpha_i + (1-s_0)\,d_{0,X_i}\right) & \text{[Case X]} \\ \otimes \bigotimes_{j=1}^{B} H^{r_{\boldsymbol{Y}_j}}_{\mathbb{P}^{\beta_j}}\left(-1 - \sum_{i=1}^{A} s_{i,j} - d_{0,Y_j} s_0\right) & \text{[Case Y]} \end{pmatrix} \tag{6.19}$$

We will study the values for $p, v, s_0, s_{1,1},\dots,s_{A,B}, r_{\boldsymbol{X}_1},\dots,r_{\boldsymbol{X}_A}, r_{\boldsymbol{Y}_1},\dots,r_{\boldsymbol{Y}_B}$ such that $K_{v,p}(\boldsymbol{m})$ does not vanish. Clearly, if $0 \leq s_0 \leq 1$ and $(\forall i \in \{1,\dots,A\})\,(\forall j \in \{1,\dots,B\})\ 0 \leq s_{i,j} \leq \mathcal{E}_{i.j}$, then the module $\bigoplus_{I \in \mathcal{I}_{s_0,s_{1,1},\dots,s_{A,B}}} \bigwedge_{k\in I} e_k$ is not zero. Hence, assuming $0 \leq s_0 \leq 1$ and $(\forall i \in \{1,\dots,A\})\,(\forall j \in \{1,\dots,B\})\ 0 \leq s_{i,j} \leq \mathcal{E}_{i.j}$ $(\forall i \in \{1,\dots,B\})\ 0 \leq s_i \leq \mathcal{E}_i$, we study the vanishing of the modules in Eq. (6.19). We will study the cohomologies independently. By Prop. 3.3.6,

the modules in the right hand side of Eq. (6.19) are not zero only when, for $1 \leq i \leq A$, $r_{\boldsymbol{X}_i} \in \{0, \alpha_i\}$ and, for $1 \leq j \leq B$, $r_{\boldsymbol{Y}_j} \in \{0, \beta_j\}$. In the following we prove that if Eq. (6.19) does not vanish then,

$$\begin{cases} r_{\boldsymbol{X}_i} = 0 & \text{for } 1 \leq i \leq A \quad \text{[Case X]} \\ r_{\boldsymbol{Y}_j} = \beta_j & \text{for } 1 \leq j \leq B \quad \text{[Case Y]} \end{cases} \tag{6.20}$$

Case (X) We consider the modules that involve the variables in the block $\boldsymbol{X}_i$, for $1 \leq i \leq A$. As $(\forall j)\, s_{i,j} \leq \mathcal{E}_{i,j}$, $0 \leq s_0 \leq 1$ and $0 \leq d_{0,\boldsymbol{X}_i} \leq 1$, we have $\sum_{j=1}^{B}(\mathcal{E}_{i,j} - s_{i,j}) - \alpha_i + (1 - s_0)\, d_{0,\boldsymbol{X}_i} > -1 - \alpha_i$. Hence, by Cor. 3.3.7,

$$\begin{aligned} &H^{r_{\boldsymbol{X}_i}}_{\mathbb{P}^{\alpha_i}}(\sum_{j=1}^{B}(\mathcal{E}_{i,j} - s_{i,j}) - \alpha_i + (1 - s_0)\, d_{0,\boldsymbol{X}_i}) \neq 0 \iff \\ r_{\boldsymbol{X}_i} = 0 \quad &\text{and} \quad \sum_{j=1}^{B}(\mathcal{E}_{i,j} - s_{i,j}) - \alpha_i + (1 - s_0)\, d_{0,\boldsymbol{X}_i} \geq 0. \end{aligned} \tag{6.21}$$

Case (Y) We consider the modules that involve the variables in the block $\boldsymbol{Y}_j$, for $1 \leq j \leq B$. As $(\forall j \in \{1, \ldots, B\})\, s_{i,j} \geq 0$ and $s_0, d_{0,\boldsymbol{Y}_j} \geq 0$, then $-1 - \sum_{i=1}^{A} s_{i,j} - s_0\, d_{0,\boldsymbol{Y}_j} < 0$, and so by Cor. 3.3.7,

$$\begin{aligned} &H^{r_{\boldsymbol{Y}_j}}_{\mathbb{P}^{\beta_j}}(-1 - \sum_{i=1}^{A} s_{i,j} - s_0\, d_{0,\boldsymbol{Y}_j}) \neq 0 \iff \\ r_{\boldsymbol{Y}_j} = \beta_j \quad &\text{and} \quad \sum_{i=1}^{A} s_{i,j} + s_0\, d_{0,\boldsymbol{Y}_j} - \beta_j \geq 0. \end{aligned} \tag{6.22}$$

We proved that, if the cohomologies in Eq. (6.19) do not vanish then Eq. (6.20) holds. We study the possible values for v such that $K_{v,p}(\boldsymbol{m})$ does not vanish. From Eq. (6.17), it holds $p = \sum_{i=1}^{A}\sum_{j=1}^{B} s_{i,j} + s_0$. By Prop. 3.3.5, $p - v = \sum_{i=1}^{A} r_{\boldsymbol{X}_i} + \sum_{j=1}^{B} r_{\boldsymbol{Y}_j}$. Hence, we deduce that, when $K_{v,p}(\boldsymbol{m})$ does not vanish, it holds,

$$v = \sum_{i=1}^{A}\sum_{j=1}^{B} s_{i,j} + s_0 - \sum_{j=1}^{B} \beta_j = \sum_{j=1}^{B}\left(\sum_{i=1}^{A} s_{i,j} - \beta_j\right) + s_0.$$

We bound the values for v for which $K_{v,p}(\boldsymbol{m})$ does not vanish.

- First, we lower-bound v. Assume that the cohomologies involving $\boldsymbol{Y}_j$ are not zero. Hence, if we sum over $j \in \{1, \ldots, B\}$ the inequalities of Eq. (6.22), we conclude that

$$0 \leq \sum_{j=1}^{B}(\sum_{i=1}^{A} s_{i,j} - \beta_j) + s_0 \sum_{j=1}^{B} d_{0,\boldsymbol{Y}_j} = v + s_0 \left(\sum_{j=1}^{B} d_{0,\boldsymbol{Y}_j} - 1\right).$$

  By definition, Eq. (6.15), $0 \leq \sum_{j=1}^{B} d_{0,\boldsymbol{Y}_j} \leq 1$, and $0 \leq s_0 \leq 1$, hence $0 \leq v$.

- Finally, we upper-bound v. Assume that the cohomologies involving $\boldsymbol{X}_j$ are not zero. Hence, if we sum over $i \in \{1, \dots, A\}$ the inequalities of Eq. (6.21), we conclude that

$$\begin{aligned} 0 &\leq \sum_{i=1}^{A} \left(\sum_{j=1}^{B} (\mathcal{E}_{i,j} - s_{i,j}) - \alpha_i + (1 - s_0)\, d_{0,\boldsymbol{X}_i} \right) \\ &= \sum_{i=1}^{A} \sum_{j=1}^{B} \mathcal{E}_{i,j} - \sum_{i=1}^{A} \alpha_i - \sum_{i=1}^{A} \sum_{j=1}^{B} s_{i,j} + (1 - s_0) \left(\sum_{i=1}^{A} d_{0,\boldsymbol{X}_i} \right). \end{aligned}$$

  Recall that $N = \sum_{i=1}^{A} \sum_{j=1}^{B} \mathcal{E}_{i,j} = \sum_{i=1}^{A} \alpha_i + \sum_{i=j}^{B} \beta_j$ and $v = \sum_{i=1}^{A} \sum_{j=1}^{B} s_{i,j} + s_0 - \sum_{i=j}^{B} \beta_j$. Also, as $\boldsymbol{d}_0$ is a solution of Eq. (6.15), it holds $0 \leq \sum_{j=1}^{B} d_{0,\boldsymbol{Y}_j} \leq 1$, and $0 \leq s_0 \leq 1$. Hence

$$v = \sum_{i=1}^{A} \sum_{j=1}^{B} s_{i,j} + s_0 - \sum_{i=j}^{B} \beta_j \leq s_0 + (1 - s_0) \left(\sum_{i=1}^{A} d_{0,\boldsymbol{X}_i} \right) \leq 1.$$

We conclude that the possible values for v, p, $r_{\boldsymbol{X}_1}, \dots, r_{\boldsymbol{X}_A}$, $r_{\boldsymbol{Y}_1}, \dots, r_{\boldsymbol{Y}_B}$ such that Eq. (6.19) is not zero are $v \in \{0, 1\}$, the possible values for $r_{\boldsymbol{X}_1}, \dots, r_{\boldsymbol{X}_A}$, $r_{\boldsymbol{Y}_1}, \dots, r_{\boldsymbol{Y}_B}$ are the ones in Eq. (6.20) and $p = \sum_{j=1}^{B} \beta_j + v$. Let $\omega = \sum_{j=1}^{B} \beta_j$. Hence, our Weyman complex looks like Eq. (3.10), where

$$\delta_1 : K_{1,\omega+1}(\boldsymbol{m}) \to K_{0,\omega}(\boldsymbol{m})$$

is a Koszul-type determinantal formula. □

Size of determinantal formulas

In this subsection we study the size of the determinantal formulas that we deduced in Thm. 6.1.14 and we compare them with the number of solutions of the square bipartite bilinear system $(f_1, \dots, f_N)$.

Lemma 6.1.16. *The expected number of solutions for a bipartite bilinear system is*

$$\Upsilon := \sum_{(s_{1,1},\dots,s_{A,B}) \in \pi} \prod_{i=1}^{A} \prod_{j=1}^{B} \binom{\mathcal{E}_{i,j}}{s_{i,j}},$$

where π is the set of solutions to the following system of inequalities,

$$\pi := \left\{ (s_{1,1}, \dots, s_{A,B}) \in \mathbb{N}^{A \cdot B} : \left\{ \begin{array}{cc} \forall (i,j) \in [A] \times [B] & 0 \leq s_{i,j} \leq \mathcal{E}_{i,j}, \\ \forall i \in [A] & \sum_{j=1}^{B} s_{i,j} = \alpha_i, \text{ and} \\ \forall j \in [B] & \sum_{i=1}^{A} s_{i,j} = \sum_i \mathcal{E}_{i,j} - \beta_j \end{array} \right\} \right\}. \quad (6.23)$$

Proof. By Prop. 2.10.9, multihomogeneous Bézout bound is equal to the coefficient of

$$\prod_{i=1}^{A} Z_{\boldsymbol{X}_i}^{\alpha_i} \prod_{j=1}^{B} Z_{\boldsymbol{Y}_j}^{\beta_j} \quad (6.24)$$

in the polynomial

$$P := \prod_{i=1}^{A}\prod_{j=1}^{B}(Z_{\boldsymbol{X}_i} + Z_{\boldsymbol{Y}_j})^{\mathcal{E}_{i,j}}. \tag{6.25}$$

We consider a new variables $s_{i,j}$ and the Newton identities

$$(Z_{\boldsymbol{X}_i} + Z_{\boldsymbol{Y}_j})^{\mathcal{E}_{i,j}} = \sum_{s_{i,j}=0}^{\mathcal{E}_{i,j}} \binom{\mathcal{E}_{i,j}}{s_{i,j}} Z_{\boldsymbol{X}_i}^{s_{i,j}}\, Z_{\boldsymbol{Y}_j}^{\mathcal{E}_{i,j}-s_{i,j}}.$$

If we expand Eq. (6.25) taking into account these identities we obtain,

$$\prod_{i=1}^{A}\prod_{j=1}^{B}(Z_{\boldsymbol{X}_i} + Z_{\boldsymbol{Y}_j})^{\mathcal{E}_{i,j}} = \prod_{i=1}^{A}\prod_{j=1}^{B}\sum_{s_{i,j}=0}^{\mathcal{E}_{i,j}} \binom{\mathcal{E}_{i,j}}{s_{i,j}} Z_{\boldsymbol{X}_i}^{s_{i,j}}\, Z_{\boldsymbol{Y}_j}^{\mathcal{E}_{i,j}-s_{i,j}} =$$

$$\sum_{s_{1,1}=0}^{\mathcal{E}_{1,1}} \cdots \sum_{s_{A,B}=0}^{\mathcal{E}_{A,B}} \left(\prod_{i=1}^{A}\prod_{j=1}^{B}\binom{\mathcal{E}_{i,j}}{s_{i,j}}\right) Z_{\boldsymbol{X}_1}^{\sum_{j=1}^{B} s_{1,j}} \cdots Z_{\boldsymbol{X}_A}^{\sum_{j=1}^{B} s_{A,j}}\, Z_{\boldsymbol{Y}_1}^{\sum_{i=1}^{A}(\mathcal{E}_{i,1}-s_{i,1})} \cdots Z_{\boldsymbol{Y}_B}^{\sum_{i=1}^{A}(\mathcal{E}_{i,B}-s_{i,B})}.$$

Hence, the coefficient of the monomial in Eq. (6.24) is the sum of $\left(\prod_{i=1}^{A}\prod_{j=1}^{B}\binom{\mathcal{E}_{i,j}}{s_{i,j}}\right)$ over the $(s_{1,1}, \ldots, s_{A,B})$ such that

$$(\forall i \in \{1, \ldots, A\})\ Z_{\boldsymbol{X}_i}^{\sum_{j=1}^{B} s_{i,j}} = Z_{\boldsymbol{X}_i}^{\alpha_i} \quad \text{and} \quad (\forall j \in \{1, \ldots, B\})\ Z_{\boldsymbol{Y}_j}^{\sum_{i=1}^{A}(\mathcal{E}_{i,j}-s_{i,j})} = Z_{\boldsymbol{Y}_j}^{\beta_j}.$$

□

The set π (Eq. 6.23) is the set of integer points in a (restricted) transportation polytope. Understanding the combinatorial structure of a transportation polytope is a hard and important problem by itself; we refer the reader to [DLK14, Sec. 2.3] for a review on the combinatorics of its number of integer points. Because of this intrinsic hardness, we do not expect to derive sharply bounds for Υ, let alone closed formulas for its value. Hence, we only study the relation between Υ and the degree of the resultant in special cases. We will consider $\boldsymbol{d}_0$ to be the multidegree of a bilinear polynomial $f_0 \in \mathbb{K}[\boldsymbol{X}_p, \boldsymbol{Y}_q]_{\mathbf{1}}$, for some $1 \leq p \leq A$ and $1 \leq q \leq B$. The degree of the resultant in the other cases is bounded by the degree of the resultant in this case.

Lemma 6.1.17. *With the same notation as Thm. 6.1.14, assume that $\boldsymbol{d}_0$ corresponds to the multidegree of a bilinear polynomial $f_0 \in \mathbb{K}[\boldsymbol{X}_p, \boldsymbol{Y}_q]_{\mathbf{1}}$ and*

$$\begin{cases} \textit{For any } i_1, i_2 \in [A], & \alpha_{i_1} = \alpha_{i_2} \\ \textit{For any } j_1, j_2 \in [B], & \beta_{j_1} = \beta_{j_2} \\ \textit{For any } (i,j) \in ([A] \times [B]) \setminus \{(p,q)\}, & \mathcal{E}_{i,j} = \mathcal{E}_{p,q} + 1 \end{cases}.$$

Then, by symmetry, the size of the matrix related to the determinantal formulas that we obtained in Thm. 6.1.14 is exactly $(N+1)$ times Υ (Lem. 6.1.16), that is,

$$Rank(K_0(\boldsymbol{m})) = Rank(K_1(\boldsymbol{m})) = \Upsilon \cdot (N+1).$$

Lemma 6.1.18. *With the same notation as Thm. 6.1.14, assume that $\boldsymbol{d}_0$ corresponds to the multidegree of a bilinear polynomial $f_0 \in \mathbb{K}[\boldsymbol{X}_p, \boldsymbol{Y}_q]_{\mathbf{1}}$ and it holds*

$$\begin{cases} & \mathcal{E}_{p,q} \geq \beta_q \\ & \mathcal{E}_{p,q} \geq \alpha_p \\ (\forall q \neq r) & \mathcal{E}_{p,r} > \alpha_p \end{cases}.$$

Then the size of the matrix related to the determinantal formulas that we obtained in Thm. 6.1.14 is bounded by

$$Rank(K_0(\boldsymbol{m})) = Rank(K_1(\boldsymbol{m})) \leq \Upsilon \cdot \frac{(\mathcal{E}_{p,q}+1)\,(N+1)}{\mathcal{E}_{p,q} - \max(\alpha_p, \beta_q) + 1} \cdot \max\Big(\frac{\alpha_p}{2}, 1\Big)$$

where Υ is the multihomogeneous Bézout bound of the square subsystem, see Lem. 6.1.16.

Proof. As the matrices represent determinantal formulas of degree one (Def. 3.2.14), we bound their size, equal to the degree of the resultant, using the multihomogeneous Bézout bound. By Prop. 3.3.2, it holds,

$$\text{degree}(\texttt{res}) = \sum_{k=0}^{N} \texttt{MHB}(\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \ldots, \boldsymbol{d}_N).$$

Assume that $f_0 \in \mathbb{K}[\boldsymbol{X}_p, \boldsymbol{Y}_q]_{\mathbf{1}}$ has degree $\boldsymbol{d}_0$. By Prop. 2.10.9, for each $f_k \in \mathbb{K}[\boldsymbol{X}_t, \boldsymbol{Y}_r]_{\mathbf{1}}$ of degree $\boldsymbol{d}_k$, the multihomogeneous Bézout bound $\texttt{MHB}(\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \ldots, \boldsymbol{d}_N)$ is equal to the coefficient of the monomial $\prod_{i=1}^{A} Z_{\boldsymbol{X}_i}^{\alpha_i} \prod_{j=1}^{B} Z_{\boldsymbol{Y}_j}^{\beta_j}$ in

$$(Z_{\boldsymbol{X}_p} + Z_{\boldsymbol{Y}_q}) \frac{\prod_{(i,j)\in[A]\times[B]} (Z_{\boldsymbol{X}_i} + Z_{\boldsymbol{Y}_j})^{\mathcal{E}_{i,j}}}{(Z_{\boldsymbol{X}_t} + Z_{\boldsymbol{Y}_r})}.$$

We proceed as in Lem. 6.1.18 and write this coefficient as a weighted sum of the integer points of a polytope η. We bound the value of this coefficient by mapping the polytope η into the polytope π from Eq. (6.23).

First we assume that $f_k \in \mathbb{K}[\boldsymbol{X}_t, \boldsymbol{Y}_r]_{\mathbf{1}}$ and $q \neq r$. Then

$$\texttt{MHB}(\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \ldots, \boldsymbol{d}_N) = \sum_{(e_{1,1},\ldots,e_{A,B})\in\eta} \binom{\mathcal{E}_{t,r}-1}{e_{t,r}} \binom{\mathcal{E}_{p,q}+1}{e_{p,q}} \prod_{\substack{1\leq i\leq A \\ 1\leq j\leq B \\ (i,j)\notin\{(p,q),(t,r)\}}} \binom{\mathcal{E}_{i,j}}{e_{i,j}}. \tag{6.26}$$

where η is the set of solutions to the following system of inequalities,

$$\eta := \left\{ (e_{1,1}, \ldots, e_{A,B}) \in \mathbb{N}^{A\cdot B} : \left\{ \begin{array}{rr} (\forall (i,j) \in ([A]\times[B]) \setminus \{(p,q),(t,r)\}) & 0 \leq e_{i,j} \leq \mathcal{E}_{i,j}, \\ & 0 \leq e_{p,q} \leq \mathcal{E}_{p,q}+1, \\ & 0 \leq e_{t,r} \leq \mathcal{E}_{t,r}-1, \\ (\forall i \in [A]) & \sum_{j=1}^{B} e_{i,j} = \alpha_i, \\ (\forall j \in [B]\setminus\{q,r\}) & \sum_{i=1}^{A} e_{i,j} = \sum_i \mathcal{E}_{i,j} - \beta_j, \\ & \sum_{i=1}^{A} e_{i,q} = \sum_i \mathcal{E}_{i,q} + 1 - \beta_q, \\ & \sum_{i=1}^{A} e_{i,r} = \sum_i \mathcal{E}_{i,r} - 1 - \beta_r \end{array} \right\} \right\}$$

We consider the subset $\bar{\eta} \subset \eta$ such that

$$\bar{\eta} := \{(e_{1,1}, \dots, e_{A,B}) \in \eta : e_{p,q} \geq 1 \text{ and } e_{p,r} \leq \mathcal{E}_{p,r} - 1\}.$$

Under our assumptions, it holds $\eta = \bar{\eta}$ because:

- For every $(e_{1,1}, \dots, e_{A,B}) \in \eta$ it holds $e_{i,q} \leq \mathcal{E}_{i,q}$ for all $i \in [A] \setminus \{q\}$ and $e_{p,q} + \sum_{i \in [A] \setminus \{q\}} e_{i,q} = \mathcal{E}_{p,q} + 1 - \beta_q + \sum_{i \in [A] \setminus \{q\}} \mathcal{E}_{i,q}$, then $e_{p,q} \geq \mathcal{E}_{p,q} + 1 - \beta_q$. As we assumed that $\mathcal{E}_{p,q} \geq \beta$, then $e_{p,q} \geq 1$.
- For every $(e_{1,1}, \dots, e_{A,B}) \in \eta$ it holds, $0 \leq e_{i,j}$ for all $(i,j) \in [A] \times [B]$ and $\sum_{j=1}^{B} e_{i,j} = \alpha_i$. Hence, $e_{p,r} \leq \alpha_i$. As we assumed that $\mathcal{E}_{p,r} > \alpha_p$ for all $r \in [B] \setminus \{q\}$, then $\mathcal{E}_{p,r} > e_{p,r}$.

We consider the injective map $\phi : \bar{\eta} \to \pi$ that, for each element $(e_{1,1}, \dots, e_{A,B}) \in \bar{\eta}$, it subtracts one from $e_{p,q}$ and adds one to $e_{p,r}$, that is

$$\phi((e_{1,1}, \dots, e_{A,B})) = (e_{1,1}, \dots, e_{p,q} - 1, \dots, a_{p,r} + 1, \dots, e_{A,B}).$$

Hence, when $p \neq t$, we can rewrite Eq. (6.26) as

$$\sum_{(s_{1,1},\dots,s_{A,B}) \in \phi(\bar{\eta})} \binom{\mathcal{E}_{t,r} - 1}{s_{t,r}} \binom{\mathcal{E}_{p,r}}{s_{p,r} - 1} \binom{\mathcal{E}_{p,q} + 1}{s_{p,q} + 1} \prod_{\substack{1 \leq i \leq A \\ 1 \leq j \leq B \\ (i,j) \notin \{(p,q),(t,r),(p,r)\}}} \binom{\mathcal{E}_{i,j}}{s_{i,j}}$$

Following the same argument as above, for any $(s_{1,1}, \dots, s_{A,B}) \in \pi$ it holds $s_{p,q} \geq \mathcal{E}_{p,q} - \beta_q$ and $s_{p,r} \leq \alpha_p < \mathcal{E}_{p,r}$. Hence, we get the following bounds

$$\begin{cases} \binom{\mathcal{E}_{t,r} - 1}{s_{t,r}} = \dfrac{\mathcal{E}_{t,r} - s_{t,r}}{\mathcal{E}_{t,r}} \binom{\mathcal{E}_{t,r}}{s_{t,r}} \leq \binom{\mathcal{E}_{t,r}}{s_{t,r}} \\ \binom{\mathcal{E}_{p,r}}{s_{p,r} - 1} = \dfrac{s_{p,r}}{\mathcal{E}_{p,r} - s_{p,r} + 1} \binom{\mathcal{E}_{p,r}}{s_{p,r}} \leq \dfrac{\alpha_p}{2} \binom{\mathcal{E}_{p,r}}{s_{p,r}} \\ \binom{\mathcal{E}_{p,q} + 1}{s_{p,q} + 1} = \dfrac{\mathcal{E}_{p,q} + 1}{s_{p,q} + 1} \binom{\mathcal{E}_{p,q}}{s_{p,q}} \leq \dfrac{\mathcal{E}_{p,q} + 1}{\mathcal{E}_{p,q} - \beta_q + 1} \binom{\mathcal{E}_{p,q}}{s_{p,q}} \end{cases}.$$

We conclude that, when $q \neq r$ and $p \neq t$, we can upper bound $\texttt{MHB}(\boldsymbol{d}_0, \dots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \dots, \boldsymbol{d}_N)$ by

$$\frac{\alpha_p}{2} \frac{\mathcal{E}_{p,q} + 1}{\mathcal{E}_{p,q} - \beta_q + 1} \sum_{(s_{1,1},\dots,s_{A,B}) \in \phi(\bar{\eta})} \prod_{\substack{1 \leq i \leq A \\ 1 \leq j \leq B}} \binom{\mathcal{E}_{i,j}}{s_{i,j}} \leq \frac{\alpha_p}{2} \frac{\mathcal{E}_{p,q} + 1}{\mathcal{E}_{p,q} - \beta_q + 1} \Upsilon$$

Following the same procedure, if $f_k \in \mathbb{K}[\boldsymbol{X}_p, \boldsymbol{Y}_r]_{\mathbf{1}}$ and $q \neq r$, we can deduce a similar bound,

$$\texttt{MHB}(\boldsymbol{d}_0, \dots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \dots, \boldsymbol{d}_N) \leq \frac{\mathcal{E}_{p,q} + 1}{\mathcal{E}_{p,q} - \beta_q + 1} \Upsilon.$$

If $f_k \in \mathbb{K}[\boldsymbol{X}_p, \boldsymbol{Y}_q]_{\mathbf{1}}$, by definition of Υ, it holds

$$\texttt{MHB}(\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \ldots, \boldsymbol{d}_N) = \Upsilon.$$

Finally, we consider $\texttt{MHB}(\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \ldots, \boldsymbol{d}_N)$ when $f_k \in \mathbb{K}[\boldsymbol{X}_t, \boldsymbol{Y}_q]_{\mathbf{1}}$ and $t \neq p$. We proceed as before, and consider a the set of points,

$$\bar{\pi} := \left\{ (s_{1,1}, \ldots, s_{A,B}) \in \mathbb{N}^{A \cdot B} : \left\{ \begin{array}{rr} (\forall (i,j) \in ([A] \times [B]) \setminus \{(p,q),(t,q)\}) & 0 \leq e_{i,j} \leq \mathcal{E}_{i,j}, \\ & 0 \leq e_{p,q} \leq \mathcal{E}_{p,q} + 1, \\ & 0 \leq e_{t,q} \leq \mathcal{E}_{t,q} - 1, \\ (\forall i \in [A]) & \sum_{j=1}^{B} s_{i,j} = \alpha_i, \text{ and} \\ (\forall j \in [B]) & \sum_{i=1}^{A} s_{i,j} = \sum_i \mathcal{E}_{i,j} - \beta_j \end{array} \right\} \right\}.$$

Under our assumptions, it holds $\bar{\pi} = \{(s_{1,1}, \ldots, s_{A,B}) \in \pi : s_{t,q} \leq \mathcal{E}_{t,q} - 1\}$ because if $(s_{1,1}, \ldots, s_{A,B}) \in \bar{\pi}$, then $s_{p,q} \leq \alpha_p$. We assumed that $\alpha_p \leq \mathcal{E}_{p,q}$, and so $s_{p,q} \leq \mathcal{E}_{p,q}$. Therefore, we can write the multihomogeneous Bézout bound as

$$\texttt{MHB}(\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \ldots, \boldsymbol{d}_N) = \sum_{\substack{(s_{1,1}, \ldots, s_{A,B}) \in \pi \\ s_{t,q} \leq \mathcal{E}_{t,q} - 1}} \binom{\mathcal{E}_{t,q} - 1}{s_{t,q}} \binom{\mathcal{E}_{p,q} + 1}{s_{p,q}} \prod_{\substack{1 \leq i \leq A \\ 1 \leq j \leq B \\ (i,j) \notin \{(p,q),(t,q)\}}} \binom{\mathcal{E}_{i,j}}{s_{i,j}}.$$

Moreover, we have the following inequalities,

$$\left\{ \begin{array}{l} \binom{\mathcal{E}_{t,q} - 1}{s_{t,q}} = \dfrac{\mathcal{E}_{t,q} - s_{t,q}}{\mathcal{E}_{t,q}} \binom{\mathcal{E}_{t,q}}{s_{t,q}} \leq \binom{\mathcal{E}_{t,q}}{s_{t,q}} \\ \\ \binom{\mathcal{E}_{p,q} + 1}{s_{p,q}} = \dfrac{\mathcal{E}_{p,q} + 1}{\mathcal{E}_{p,q} + 1 - s_{p,q}} \binom{\mathcal{E}_{p,q}}{s_{p,q}} \leq \dfrac{\mathcal{E}_{p,q} + 1}{\mathcal{E}_{p,q} - \alpha_p + 1} \binom{\mathcal{E}_{p,q}}{s_{p,q}} \end{array} \right.$$

Hence, we can bound the multihomogeneous Bézout bound,

$$\texttt{MHB}(\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \ldots, \boldsymbol{d}_N) \leq \frac{\mathcal{E}_{p,q} + 1}{\mathcal{E}_{p,q} - \alpha_p + 1} \sum_{\substack{(s_{1,1}, \ldots, s_{A,B}) \in \pi \\ s_{t,q} \leq \mathcal{E}_{t,q} - 1}} \prod_{\substack{1 \leq i \leq A \\ 1 \leq j \leq B}} \binom{\mathcal{E}_{i,j}}{s_{i,j}} \leq \frac{\mathcal{E}_{p,q} + 1}{\mathcal{E}_{p,q} - \alpha_p + 1} \Upsilon$$

To conclude, we use the previous bounds to bound the degree of the multiprojective resultant, $\deg(\texttt{res}) = \sum_{k=0}^{N} \texttt{MHB}(\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}, \boldsymbol{d}_{k+1}, \ldots, \boldsymbol{d}_N)$. It holds,

$$\deg(\texttt{res}) \leq \left(\begin{array}{l} \displaystyle\sum_{\substack{1 \leq t \leq A \\ 1 \leq r \leq B \\ t \neq p \wedge r \neq q}} \mathcal{E}_{t,r} \, \frac{\alpha_p}{2} \, \frac{\mathcal{E}_{p,r} + 1}{\mathcal{E}_{p,q} - \beta_q + 1} \, \Upsilon \\ + \displaystyle\sum_{\substack{1 \leq r \leq B \\ r \neq q}} \mathcal{E}_{p,r} \, \frac{\mathcal{E}_{p,q} + 1}{\mathcal{E}_{p,q} - \beta_p + 1} \, \Upsilon \\ + \displaystyle\sum_{\substack{1 \leq t \leq A \\ t \neq p}} \mathcal{E}_{t,q} \, \frac{\mathcal{E}_{p,q} + 1}{\mathcal{E}_{p,q} - \alpha_p + 1} \, \Upsilon \\ + \left(\mathcal{E}_{p,q} + 1\right) \Upsilon \end{array} \right) \leq \Upsilon \cdot \frac{\left(\mathcal{E}_{p,q} + 1\right)(N+1)}{\mathcal{E}_{p,q} - \max(\alpha_p, \beta_q) + 1} \cdot \max\Big(\frac{\alpha_p}{2}, 1\Big)$$

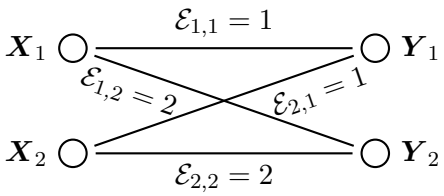
□

Example

Consider four blocks of variables such that $A = 2$, $B = 2$, $\alpha = (1, 2)$, $\beta = (1, 2)$, and

$$\begin{cases} X_1 := & \{X_{1,0}, X_{1,1}\} \\ X_2 := & \{X_{2,0}, X_{2,1}, X_{2,2}\} \\ Y_1 := & \{Y_{1,0}, Y_{1,1}\} \\ Y_2 := & \{Y_{2,0}, Y_{2,1}, Y_{2,2}\}. \end{cases}$$

Let $(f_1, \ldots, f_6)$ be the square bipartite bilinear system represented by the following graph:



The expected number of solutions of the system is 5. If we introduce a polynomial $f_0 \in \mathbb{K}[\boldsymbol{X}_1, \boldsymbol{Y}_1]_{\mathbf{1}}$, the system $(f_0, f_1, \ldots, f_6)$ satisfies the hypothesis of Lem. 6.1.18, and so the degree of the resultant is upper bounded by 50. By Prop. 2.10.9, the degree of the resultant is 24.

We consider the overdetermined system $\boldsymbol{f}$ where,

$$\boldsymbol{f} := \begin{cases} f_0 := & (a_1\, y_{1;0} + a_2\, y_{1;1})\, x_{1;0} + (a_3\, y_{1;0} + a_4\, y_{1;1})\, x_{1;1} \\ f_1 := & (b_1\, y_{1;0} + b_2\, y_{1;1})\, x_{1;0} + (b_3\, y_{1;0} + b_4\, y_{1;1})\, x_{1;1} \\ f_2 := & (c_1\, y_{1;0} + c_2\, y_{1;1})\, x_{2;0} + (c_5\, y_{1;0} + c_6\, y_{1;1})\, x_{2;1} + (c_3\, y_{1;0} + c_4\, y_{1;1})\, x_{2;2} \\ f_3 := & (d_1\, y_{2;0} + d_2\, y_{2;2} + d_3\, y_{2;1})\, x_{1;0} + (d_4\, y_{2;0} + d_5\, y_{2;2} + d_6\, y_{2;1})\, x_{1;1} \\ f_4 := & (e_1\, y_{2;0} + e_2\, y_{2;2} + e_3\, y_{2;1})\, x_{1;0} + (e_4\, y_{2;0} + e_5\, y_{2;2} + e_6\, y_{2;1})\, x_{1;1} \\ f_5 := & (g_1\, y_{2;0} + g_2\, y_{2;2} + g_3\, y_{2;1})\, x_{2;0} + (g_7\, y_{2;0} + g_8\, y_{2;2} + g_9\, y_{2;1})\, x_{2;1} \\ & \qquad + (g_4\, y_{2;0} + g_5\, y_{2;2} + g_6\, y_{2;1})\, x_{2;2} \\ f_6 := & (h_1\, y_{2;0} + h_2\, y_{2;2} + h_3\, y_{2;1})\, x_{2;0} + (h_7\, y_{2;0} + h_8\, y_{2;2} + h_9\, y_{2;1})\, x_{2;1} \\ & \qquad + (h_4\, y_{2;0} + h_5\, y_{2;2} + h_6\, y_{2;1})\, x_{2;2} \end{cases} \tag{6.27}$$

Following Sec. 6.1.3, we consider the degree vector $\boldsymbol{m} = (3, 1, -1, -1)$. The vector spaces of the Weyman complex $K(\boldsymbol{m}, \boldsymbol{f})$ looks like,

$$\begin{aligned} K_1(\boldsymbol{m}, \boldsymbol{f}) = \quad & S_{\boldsymbol{X}_1}(0) \otimes S_{\boldsymbol{X}_2}(0) \otimes S^*_{\boldsymbol{Y}_1}(0) \otimes S^*_{\boldsymbol{Y}_2}(-1) \otimes \left\{ \begin{array}{l} (e_0 \wedge e_3 \wedge e_4 \wedge e_5) \oplus (e_0 \wedge e_3 \wedge e_4 \wedge e_6) \oplus \\ (e_2 \wedge e_3 \wedge e_4 \wedge e_5) \oplus (e_2 \wedge e_3 \wedge e_4 \wedge e_6) \end{array} \right\} \\ & \oplus S_{\boldsymbol{X}_1}(0) \otimes S_{\boldsymbol{X}_2}(0) \otimes S^*_{\boldsymbol{Y}_1}(-1) \otimes S^*_{\boldsymbol{Y}_2}(0) \otimes \left\{ \begin{array}{l} (e_0 \wedge e_2 \wedge e_3 \wedge e_6) \oplus (e_0 \wedge e_2 \wedge e_4 \wedge e_5) \oplus \\ (e_0 \wedge e_2 \wedge e_4 \wedge e_6) \oplus (e_0 \wedge e_2 \wedge e_3 \wedge e_4) \oplus \\ (e_2 \wedge e_2 \wedge e_3 \wedge e_4) \oplus (e_0 \wedge e_2 \wedge e_3 \wedge e_5) \end{array} \right\} \\ K_0(\boldsymbol{m}, \boldsymbol{f}) = \quad & S_{\boldsymbol{X}_1}(0) \otimes S_{\boldsymbol{X}_2}(1) \otimes S^*_{\boldsymbol{Y}_1}(0) \otimes S^*_{\boldsymbol{Y}_2}(0) \otimes \left\{ \begin{array}{l} (e_0 \wedge e_3 \wedge e_4) \oplus (e_2 \wedge e_3 \wedge e_4) \end{array} \right\} \\ & \oplus S_{\boldsymbol{X}_1}(1) \otimes S_{\boldsymbol{X}_2}(0) \otimes S^*_{\boldsymbol{Y}_1}(0) \otimes S^*_{\boldsymbol{Y}_2}(0) \otimes \left\{ \begin{array}{l} (e_0 \wedge e_3 \wedge e_5) \oplus (e_0 \wedge e_3 \wedge e_6) \oplus \\ (e_0 \wedge e_4 \wedge e_5) \oplus (e_0 \wedge e_4 \wedge e_6) \oplus \\ (e_1 \wedge e_3 \wedge e_5) \oplus (e_1 \wedge e_3 \wedge e_6) \oplus \\ (e_1 \wedge e_4 \wedge e_5) \oplus (e_1 \wedge e_4 \wedge e_6) \oplus \\ (e_2 \wedge e_3 \wedge e_4) \end{array} \right\} \end{aligned}$$

The Koszul determinantal matrix representing the map $\delta_1(\boldsymbol{m}, \boldsymbol{f})$ between the modules with respect to the monomial basis is,

$$
\begin{bmatrix}
 & & & -b_1 & -b_3 & & & & & & & & & a_1 & a_3 & & & & & & & & \\
 & & & -b_2 & -b_4 & & & & & & & & & a_2 & a_4 & & & & & & & & \\
 & & & & & -b_1 & -b_3 & & & & & & & & & a_1 & a_3 & & & & & & \\
 & & & & & -b_2 & -b_4 & & & & & & & & & a_2 & a_4 & & & & & & \\
 & & & & & & & -b_1 & -b_3 & & & & & & & & & a_1 & a_3 & & & \\
 & & & & & & & -b_2 & -b_4 & & & & & & & & & a_2 & a_4 & & & \\
 & & & & & & & & & -b_1 & -b_3 & & & & & & & & & a_1 & a_3 & \\
 & & & & & & & & & -b_2 & -b_4 & & & & & & & & & a_2 & a_4 & \\
-c_1 & -c_5 & -c_3 & & & & & & & & & & & & & & & & & & & a_1 & a_3 \\
-c_2 & -c_6 & -c_4 & & & & & & & & & & & & & & & & & & & a_2 & a_4 \\
-g_1 & -g_7 & -g_4 & e_1 & e_4 & & & -d_1 & -d_4 & & & & & & & & & & & & & & \\
-g_3 & -g_9 & -g_6 & e_3 & e_6 & & & -d_3 & -d_6 & & & & & & & & & & & & & & \\
-g_2 & -g_8 & -g_5 & e_2 & e_5 & & & -d_2 & -d_5 & & & & & & & & & & & & & & \\
-h_1 & -h_7 & -h_4 & & & e_1 & e_4 & & & -d_1 & -d_4 & & & & & & & & & & & & \\
-h_3 & -h_9 & -h_6 & & & e_3 & e_6 & & & -d_3 & -d_6 & & & & & & & & & & & & \\
-h_2 & -h_8 & -h_5 & & & e_2 & e_5 & & & -d_2 & -d_5 & & & & & & & & & & & & \\
 & & & & & & & & & & & -c_1 & -c_5 & -c_3 & & & & & & & & b_1 & b_3 \\
 & & & & & & & & & & & -c_2 & -c_6 & -c_4 & & & & & & & & b_2 & b_4 \\
 & & & & & & & & & & & -g_1 & -g_7 & -g_4 & e_1 & e_4 & & -d_1 & -d_4 & & & & \\
 & & & & & & & & & & & -g_3 & -g_9 & -g_6 & e_3 & e_6 & & -d_3 & -d_6 & & & & \\
 & & & & & & & & & & & -g_2 & -g_8 & -g_5 & e_2 & e_5 & & -d_2 & -d_5 & & & & \\
 & & & & & & & & & & & -h_1 & -h_7 & -h_4 & & & e_1 & e_4 & & -d_1 & -d_4 & & \\
 & & & & & & & & & & & -h_3 & -h_9 & -h_6 & & & e_3 & e_6 & & -d_3 & -d_6 & & \\
 & & & & & & & & & & & -h_2 & -h_8 & -h_5 & & & e_2 & e_5 & & -d_2 & -d_5 & &
\end{bmatrix}
\tag{6.28}
$$

6.2 Generalized eigenvalue criterion

As we discussed in Sec. 5.2.1, a strategy to solve a zero-dimensional system $f_1, \dots, f_N \in \mathbb{K}[\boldsymbol{x}]$ is to consider the quotient ring $\mathbb{K}[\boldsymbol{x}]/\langle f_1, \dots, f_N\rangle$, which is a finite a dimensional vector space over $\mathbb{K}$: We fix a monomial basis, choose $f_0 \in \mathbb{K}[\boldsymbol{x}]$, and compute the matrix that represents the multiplication by f_0 in this quotient ring. Its eigenvalues are the evaluations of f_0 at the solutions, see Prop. 5.2.3. For a suitable basis, from the eigenvectors we can recover the coordinates of all the solutions, see Prop. 5.2.4. To compute these matrices we can use Sylvester-type formulas, see Sec. 5.3.2. In this section, we extend these techniques to a general family of matrices, that include Koszul resultant matrices (Prop. 3.3.13).

As we did in Sec. 3.3.1, in this section we fix multidegrees $\boldsymbol{d}_0, \dots, \boldsymbol{d}_N$. We will follow the notation from Sec. 3.3, but we warn the reader that all the results of this section are independent from the multi-projective resultant and we can use them in more general settings, for example, with the sparse resultant. The only result restricted to multihomogeneous systems is Cor. 6.2.8, because we need to perform a generic linear change of coordinates, which we cannot necessarily do in the sparse setting.

Definition 6.2.1 (Property Π_θ)**.** *Given an exponent of a monomial of degree $\boldsymbol{d}_0$, that is $\theta \in \mathcal{A}(\boldsymbol{d}_0)$, and a matrix $M := \left[\begin{smallmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2}\end{smallmatrix}\right] \in \mathbb{Z}[\boldsymbol{u}]^{\mathcal{K}\times\mathcal{K}}$, we say that M has the property $\Pi_\theta(\boldsymbol{d}_0, \dots, \boldsymbol{d}_N)$, or simply Π_θ, when:*

- *the resultant, as a polynomial $\mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$, divides $\det(M) \in \mathbb{Z}[\boldsymbol{u}]$,*
- *the submatrix $M_{2,2}$ is square and its diagonal entries equal to $u_{0,\boldsymbol{\theta}}$, and*
- *the coefficient $u_{0,\boldsymbol{\theta}}$ does not appear anywhere in M expect from the diagonal of $M_{2,2}$.*

Koszul resultant matrices (Prop. 3.3.13) and so, Sylvester matrices satisfy this property.

Lemma 6.2.2. *Assume that we have a Koszul-type determinantal formula for the resultant $\mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$ related to a generic multihomogeneous system of multidegrees $\boldsymbol{d}_0, \dots, \boldsymbol{d}_N$, respectively, see Def. 3.3.1. Hence, the Koszul resultant matrix M related to this formulas satisfies property Π_θ, for any $\theta \in \mathcal{A}(\boldsymbol{d}_0)$.*

Proof. We only need to check that the Koszul resultant matrices satisfy the second and third condition of property Π_θ. By construction, the entries of the Koszul resultant matrix are the variables of $\boldsymbol{u}$ up to sign. Note that if $u_{i,\sigma} \in \boldsymbol{u}$ appears in an entry, then it does not appear in the other entries in the same row or column. Hence, we can rearrange the elements of the matrix in such a way that the coefficient $u_{0,\theta}$ only appears in the diagonal of $M_{2,2}$. □

Given a multihomogeneous system $\boldsymbol{f_0} := (f_0, \dots, f_N)$ in $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$ of multidegrees $\boldsymbol{d}_0, \dots, \boldsymbol{d}_N$, respectively, let $M(\boldsymbol{f_0})$ be the specialization of the variables $\boldsymbol{u}$ of M at the coefficients of the polynomial system $\boldsymbol{f_0}$, see Def. 3.2.2. If $M_{1,1}(\boldsymbol{f_0})$ is invertible, then the Schur complement of $M_{2,2}(\boldsymbol{f_0})$ (Def. 5.3.5) is

$$M_{2,2}(\boldsymbol{f_0}) - M_{2,1}(\boldsymbol{f_0}) \cdot (M_{1,1}(\boldsymbol{f_0}))^{-1} \cdot M_{1,2}(\boldsymbol{f_0}).$$

To simplify notation, we write this complement as $(M_{2,2} - M_{2,1} \cdot M_{1,1}^{-1} \cdot M_{1,2})(\boldsymbol{f_0})$.

Theorem 6.2.3. *Consider $\theta \in \mathcal{A}(\boldsymbol{d}_0)$ and a matrix $M \in \mathbb{K}[\boldsymbol{u}]^{\mathcal{K}\times\mathcal{K}}$ such that Π_θ holds (Def. 6.2.1). Consider a system $\boldsymbol{f_0} := (f_0, \dots, f_N)$ in $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$ such that the specialization $M_{1,1}(\boldsymbol{f_0})$ is an invertible matrix. Then, for each solution $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(\langle f_1, \dots, f_N\rangle)$ of the subsystem $(f_1, \dots, f_N)$ such that $\boldsymbol{x}^{\boldsymbol{\theta}}(\boldsymbol{p}) \neq 0$, $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$ is an eigenvalue of the Schur complement of $M_{2,2}(\boldsymbol{f_0})$.*

Proof. The idea of the proof is as follows: For each solution $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(\langle f_1,\ldots,f_N\rangle)$ of the subsystem $(f_1,\ldots,f_N)$, we consider a system $\boldsymbol{g_0}$, slightly different from $\boldsymbol{f_0}$, with $\boldsymbol{p}$ as a solution. We study the matrices $M(\boldsymbol{f_0})$ and $M(\boldsymbol{g_0})$ and from the kernel of $M(\boldsymbol{g_0})$ we construct an eigenvector for the Schur complement of $M_{2,2}(\boldsymbol{f_0})$ corresponding to an eigenvalue equal to $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$.

Let $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(\langle f_1,\ldots,f_N\rangle)$ such that $\boldsymbol{x}^{\boldsymbol{\theta}}(\boldsymbol{p}) \neq 0$. Consider the polynomial $g_0 := f_0 - \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}) \cdot \boldsymbol{x}^{\boldsymbol{\theta}}$ and a new system $\boldsymbol{g_0} := (g_0, f_1, \ldots, f_N)$. The coefficients of the polynomials g_0 and f_0 are the same, with exception of the coefficient of the monomial $\boldsymbol{x}^{\boldsymbol{\theta}}$, so the specializations $u_{i,\boldsymbol{\sigma}}(\boldsymbol{f_0})$ and $u_{i,\boldsymbol{\sigma}}(\boldsymbol{g_0})$ (Def. 3.2.2) differ if and only if $i = 0$ and $\boldsymbol{\sigma} = \boldsymbol{\theta}$. Hence, as M satisfies Π_θ, then $u_{0,\boldsymbol{\theta}}$ does not appear in $M_{1,1}$, $M_{2,1}$, and $M_{1,2}$, and $M_{1,1}(\boldsymbol{g_0}) = M_{1,1}(\boldsymbol{f_0})$, $M_{1,2}(\boldsymbol{g_0}) = M_{1,2}(\boldsymbol{f_0})$, and $M_{2,1}(\boldsymbol{g_0}) = M_{2,1}(\boldsymbol{f_0})$.

The specialization of $u_{0,\boldsymbol{\theta}}$ is a ring homomorphism, so $u_{0,\boldsymbol{\theta}}(\boldsymbol{g_0}) = u_{0,\boldsymbol{\theta}}(\boldsymbol{f_0}) - \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$. As Π_θ holds, $u_{0,\boldsymbol{\theta}}$ only appears in the diagonal of $M_{2,2}$. Hence, $M_{2,2}(\boldsymbol{g_0}) = M_{2,2}(\boldsymbol{f_0}) - \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}) \cdot I$, where I is the identity matrix. Therefore,

$$M(\boldsymbol{g_0}) = \begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{bmatrix}(\boldsymbol{f_0}) - \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}) \cdot \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}.$$

By construction, it holds $g_0(\boldsymbol{p}) = 0$ and $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(\langle f_1,\ldots,f_N\rangle)$. Thus, $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(\langle \boldsymbol{g_0}\rangle)$, and so $\mathtt{res}(\boldsymbol{g_0})$ vanishes, see Eq. (3.7). By property Π_θ, $\det(M)$ is a multiple of $\mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$, hence $M(\boldsymbol{g_0})$ is a singular square matrix. Let $v \in \mathtt{Ker}(M(\boldsymbol{g_0}))$, then

$$M(\boldsymbol{g_0}) \cdot v = 0 \iff \begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{bmatrix}(\boldsymbol{f_0}) \cdot v = \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}) \cdot \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \cdot v \ .$$

Multiplying this equality by the non-singular matrix related to the Schur complement of $M_{2,2}(\boldsymbol{f_0})$, $\begin{bmatrix} I & 0 \\ -M_{2,1}\cdot M_{1,1}^{-1} & I \end{bmatrix}(\boldsymbol{f_0})$, we obtain

$$\begin{bmatrix} M_{1,1} & M_{1,2} \\ 0 & (M_{2,2} - M_{2,1}\cdot M_{1,1}^{-1}\cdot M_{1,2}) \end{bmatrix}(\boldsymbol{f_0}) \cdot v = \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}) \cdot \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \cdot v \ .$$

Consider the lower part of the matrices in the previous identity,

$$\left[\ 0 \ \middle|\ M_{2,2} - M_{2,1}\cdot M_{1,1}^{-1}\cdot M_{1,2}\ \right](\boldsymbol{f_0}) \cdot v = \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}) \cdot \left[\ 0 \ \middle|\ I\ \right] \cdot v$$

and let $\bar{v} := \left[\ 0 \ \middle|\ I\ \right] \cdot v$ be a truncation of the vector v. Then,

$$(M_{2,2} - M_{2,1}\cdot M_{1,1}^{-1}\cdot M_{1,2})(\boldsymbol{f_0}) \cdot \bar{v} = \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}) \cdot \bar{v} \ .$$

This equality proves that $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$ is an eigenvalue of the Schur complement of $M_{2,2}(\boldsymbol{f_0})$ with eigenvector $\bar{v}$. □

Definition 6.2.4. *Given a square multihomogeneous system $\boldsymbol{f}$, a multihomogeneous polynomial $f_0 \in \mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]_{\boldsymbol{d}_0}$ of degree $\boldsymbol{d}_0$ and an exponent $\theta \in \mathcal{A}(\boldsymbol{d}_0)$, we say that the rational function $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}$* separates the zeros *of the system $\boldsymbol{f}$, if for all $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(\langle \boldsymbol{f}\rangle)$, $\boldsymbol{x}^{\boldsymbol{\theta}}(\boldsymbol{p}) \neq 0$ and for all $\boldsymbol{p}, \boldsymbol{p}' \in \mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(\langle \boldsymbol{f}\rangle)$, $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}) = \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}') \iff \boldsymbol{p} = \boldsymbol{p}'$.*

Corollary 6.2.5. *Under the assumptions of Thm. 6.2.3, if*

- *the row dimension of $M_{2,2}$ is the multihomogeneous Bézout bound of a multihomogeneous systems of multidegrees $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N$, $\mathtt{MHB}(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N)$ (Prop. 2.10.9),*
- *$\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}$ separates the zeros of $(f_1, \ldots, f_N)$, and*
- *there are $\mathtt{MHB}(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N)$ different solutions $(f_1, \ldots, f_N)$ (over $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$),*

Then, the Schur complement of $M_{2,2}(\boldsymbol{f_0})$ is diagonalizable with eigenvalues $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$, for $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}}(\langle f_1, \ldots, f_N \rangle)$.

Proof. As a consequence of Thm. 6.2.3, for each $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}}(\boldsymbol{f})$ we have an eigenvalue $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$ for the Schur complement of $M_{2,2}(\boldsymbol{f_0})$. As $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}$ separates these zeros, all the eigenvalues are different. Hence, we have as many different eigenvalues as the dimension of the matrix, so the matrix is diagonalizable. □

Note that, as the multihomogeneous Bézout bound bounds the number of isolated solutions counting multiplicities, we cannot use Thm. 6.2.5 when we have a square system $(f_1, \ldots, f_N)$ such that its solutions over $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$ have multiplicities.

Whenever we have a determinantal formula, all the eigenvalues in Thm. 6.2.3 are related to a solution of the system.

Lemma 6.2.6. *Under the assumptions of Thm. 6.2.3, assume that $\mathtt{res}(\boldsymbol{f_0}) \neq 0$ and $\det(M) = t \cdot \mathtt{res}$, where t is a non-zero constant in $\mathbb{K}$, that is M is a determinantal formula, see Def. 3.2.14. If λ is an eigenvalue of the Schur complement of $M_{2,2}(\boldsymbol{f_0})$, then there is $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}}(\boldsymbol{f})$ such that $\lambda = \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$.*

Proof. Consider the system $\boldsymbol{g_0} := ((f_0 - \lambda \cdot \boldsymbol{x}^{\boldsymbol{\theta}}), f_1, \ldots, f_N)$. As the matrix of the Schur complement in the proof of Thm. 6.2.3 is invertible, we extend $\bar{v}$ to $v = \begin{bmatrix} M_{1,1}^{-1} \cdot M_{2,1} \\ I \end{bmatrix}(\boldsymbol{f_0})\, \bar{v}$, and reverse the argument in this proof to show that $M(\boldsymbol{g_0})$ is singular. As the determinant of M is a non-zero constant multiple of the resultant, we deduce that $\mathtt{res}(\boldsymbol{g_0})$ is zero. Hence, we can consider $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}}(\boldsymbol{g_0})$ and conclude that $\boldsymbol{p} \subset \mathbb{V}_{\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}}(\langle f_1, \ldots, f_N \rangle)$ and $(f_0 - \lambda \cdot \boldsymbol{x}^{\boldsymbol{\theta}})(\boldsymbol{p}) = 0$, equivalently, $f_0(\boldsymbol{p}) = \lambda \cdot \boldsymbol{x}^{\boldsymbol{\theta}}(\boldsymbol{p})$. As we assumed that $\mathtt{res}(\boldsymbol{f_0}) \neq 0$, then $f_0(\boldsymbol{p}) \neq 0$ and so $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}) = \lambda$. □

Moreover, when we have a determinantal formula, the singularity of the matrix $M_{1,1}(\boldsymbol{f_0})$ does not depend on f_0.

Proposition 6.2.7. *Under the assumptions of Thm. 6.2.3, assume $\det(M) = t \cdot \mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$, where t is a non-zero constant in $\mathbb{K}$, that is, $\det(M)$ is a determinantal formula for the resultant, and that the (row) dimension of $M_{2,2}$ is $\mathtt{MHB}(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N)$. Then, for any system $\boldsymbol{f_0} := (f_0, \ldots, f_N)$,*

$$\mathbb{V}_{\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}}(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1, \ldots, f_N) = \emptyset \iff M_{1,1}(\boldsymbol{f_0}) \text{ is non-singular.}$$

Proof. Consider the determinant of M. As we assumed that it is a multiple of the resultant and the resultant is a multihomogeneous polynomial of degree $\mathtt{MHB}(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N)$ with respect to $\boldsymbol{u_0}$ (Prop. 3.3.2), we can write $\det(M) = P(\boldsymbol{u}) \cdot u_{0,\boldsymbol{\theta}}^{\mathtt{MHB}(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N)} + Q(\boldsymbol{u})$, where $P(\boldsymbol{u}) \in \mathbb{K}[\boldsymbol{u}]$ does not involve any variable in the block $\boldsymbol{u}_0$ and $Q(\boldsymbol{u}) \in \mathbb{K}[\boldsymbol{u}]$ is a polynomial such that none of its monomials are multiple of

$u_{0,\boldsymbol{\theta}}^{\mathtt{MHB}(\boldsymbol{d}_1,\ldots,\boldsymbol{d}_N)}$. As M satisfies Π_θ, $u_{0,\boldsymbol{\theta}}$ only appears in the diagonal of $M_{2,2}$. Consider the expansion by minors of $\det(M)$. If the (row) dimension of $M_{2,2}$ is $\mathtt{MHB}(\boldsymbol{d}_1,\ldots,\boldsymbol{d}_N)$, then $P(\boldsymbol{u}) = \pm\det(M_{1,1})$. The polynomial $P(\boldsymbol{u})$ is a constant multiple of the cofactor of $u_{0,\boldsymbol{\theta}}^{\mathtt{MHB}(\boldsymbol{d}_1,\ldots,\boldsymbol{d}_N)}$ in the resultant $\mathtt{res} \in \mathbb{Z}[\boldsymbol{u}]$.

By construction, $Q(\boldsymbol{u})$ is a homogeneous polynomial with respect to the variables $\boldsymbol{u}_0$ of degree $\mathtt{MHB}(\boldsymbol{d}_1,\ldots,\boldsymbol{d}_N)$. As $u_{0,\boldsymbol{\theta}}^{\mathtt{MHB}(\boldsymbol{d}_1,\ldots,\boldsymbol{d}_N)}$ does not divide any monomial in $Q(\boldsymbol{u})$, each monomial involves a variables of $\boldsymbol{u}_0$ different to $u_{0,\theta}$. Hence, for any system $\boldsymbol{f_0}$, we have $Q(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1,\ldots,f_N) = 0$. By construction, the polynomial $P(\boldsymbol{u})$ does not involve any of the variables of $\boldsymbol{u}_0$. Therefore, $\det(M_{1,1})(\boldsymbol{f_0}) = \det(M_{1,1})(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1,\ldots,f_N)$. Hence, for any system $\boldsymbol{f_0}$, $t\cdot\mathtt{res}(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1,\ldots,f_N) = \det(M)(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1\ldots f_N) = \pm\det(M_{1,1})(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1\ldots f_N) = \pm\det(M_{1,1})(\boldsymbol{f_0})$. The determinant of M is a non-zero constant multiple of the resultant, and so $\det(M_{1,1})(\boldsymbol{f_0}) \neq 0$ if and only if the system $(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1,\ldots,f_N)$ has no solutions over $\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}$, i.e., $\mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1,\ldots,f_N) = \emptyset$. □

If the square system $(f_1,\ldots,f_N)$ has no solutions at infinity in $\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}$, that is all the coordinates of the solutions are not zero, then the evaluation of the solutions of $\boldsymbol{f}$ at any monomial in $\mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]_{\boldsymbol{d}_0}$ is not zero. Hence, for any $\boldsymbol{x}^{\boldsymbol{\theta}} \in \mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]_{\boldsymbol{d}_0}$, it holds $\mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1,\ldots,f_N) = \emptyset$. By Prop. 6.2.7, $M_{1,1}(f_0, f_1,\ldots,f_N)$ is invertible. To avoid solutions at infinity, in the zero-dimensional multihomogeneous case, we perform a generic linear change of coordinates that preserves the multihomogeneous structure. We state the following corollary without proof.

Corollary 6.2.8. *Consider a square multihomogeneous system $(f_1,\ldots,f_N)$ in $\mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]$ of degrees $\boldsymbol{d}_1,\ldots,\boldsymbol{d}_N$, respectively, whose solution set $\mathbb{V}_{\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}}(\langle f_1,\ldots,f_N\rangle)$ is finite. Assume we can construct a determinantal formula for the resultant of a generic multihomogeneous system of multidegrees $\boldsymbol{d}_0,\boldsymbol{d}_1,\ldots,\boldsymbol{d}_N$ as the resultant of a matrix M satisfying Π_θ. Then, for any $f_0 \in \mathbb{K}[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_q]_{\boldsymbol{d}_0}$ and a generic linear change of coordinates A, preserving the multihomogeneous structure, the matrix $M_{1,1}(f_0, f_1\circ A,\ldots,f_N\circ A)$ is invertible.*

As a particular case of these results, we can use Thm. 6.2.3 to solve any multihomogeneous systems for which we know a Koszul-type determinantal formula, see Sec. 6.1. We present an example in Sec. 6.3.1.

6.3 Eigenvector criterion for 2-bilinear systems

In this section we only consider Koszul-type determinantal formulas for 2-bilinear systems (see below). We study the eigenvectors of the matrices from Sec. 6.2. Our proof is ad-hoc for the 2-bilinear case, so we introduce some specific notation.

A *2-bilinear system* is a particular case of a star multilinear system where $A = 1$ and $B = 2$, see Def. 6.1.1. These systems involve three blocks of variables which we identify by $\boldsymbol{x}$ (instead of $\boldsymbol{X_1}$), $\boldsymbol{y}$ (instead of $\boldsymbol{Y_1}$) and $\boldsymbol{z}$ (instead of $\boldsymbol{Y_2}$). We set $\alpha := \#\boldsymbol{x} - 1$, $\beta_y := \#\boldsymbol{y} - 1$, $\beta_z := \#\boldsymbol{z} - 1$, $N := \alpha + \beta_y + \beta_z$ and for all $d \in \mathbb{N}$, $S_{\boldsymbol{x}}(d) = \mathbb{K}[\boldsymbol{x}]_d$, $S^*_{\boldsymbol{x}}(-d) = \mathbb{K}[\boldsymbol{x}]_d$, $S_{\boldsymbol{y}}(d) = \mathbb{K}[\boldsymbol{y}]_d$, $S^*_{\boldsymbol{y}}(-d) = \mathbb{K}[\boldsymbol{y}]_d$, $S_{\boldsymbol{z}}(d) = \mathbb{K}[\boldsymbol{z}]_d$, and $S^*_{\boldsymbol{z}}(-d) = \mathbb{K}[\boldsymbol{z}]_d$.

Given a square 2-bilinear system $(f_1, \dots, f_N)$ we denote by $\mathcal{E}_y$ the number of equations of the 2-bilinear system belonging to $S_{\boldsymbol{x}}(1) \otimes S_{\boldsymbol{y}}(1)$ and by $\mathcal{E}_z$ the number of equations belonging to $S_{\boldsymbol{x}}(1) \otimes S_{\boldsymbol{z}}(1)$. Without loss of generality, we assume that $f_1, \dots, f_{\mathcal{E}_y} \in S_{\boldsymbol{x}}(1) \otimes S_{\boldsymbol{y}}(1)$ and $f_{\mathcal{E}_y+1}, \dots, f_N \in S_{\boldsymbol{x}}(1) \otimes S_N(1)$. We introduce a trilinear polynomial $f_0 \in S_{\boldsymbol{x}}(1) \otimes S_{\boldsymbol{y}}(1) \otimes S_{\boldsymbol{z}}(1)$ and consider a Koszul-type determinantal formula for the resultant of the overdetermined system $\boldsymbol{f_0} := (f_0, f_1, \dots, f_N)$. In this case, the polynomial f_0 corresponds to $\boldsymbol{d}_0 := (1, 1, 1)$ and so, to the case (4) in the list appearing in the page 115.

According to Thm. 6.1.8, we can construct a Koszul-type determinantal formula for the resultant of $\boldsymbol{f_0}$ by considering the degree vector $\boldsymbol{m} = (\mathcal{E}_y - 1, -1, \mathcal{E}_z - \beta_z + 1)$. In this case, the Weyman complex reduces to

$$K_\bullet(\boldsymbol{f_0}, \boldsymbol{m}) : 0 \to K_{1,\alpha+\beta_y+1}(\boldsymbol{m}) \xrightarrow{\delta_1} K_{0,\alpha+\beta_y}(\boldsymbol{m}) \to 0\,, \tag{6.29}$$

where (see Eq. 6.3),

$$\begin{aligned} K_{1,\alpha+\beta_y+1} &\cong L_{1,1} \oplus L_{1,2} \qquad (6.30)\\ &= \Big(S^*_{\boldsymbol{x}}(-1) \otimes S^*_{\boldsymbol{y}}(\beta_y - \mathcal{E}_y) \otimes S_{\boldsymbol{z}}(0) \otimes \bigoplus_{I \in \mathcal{I}_{0,\mathcal{E}_y,\mathcal{E}_z-\beta_z+1}} \bigwedge_{k\in I} e_k.\Big) \oplus \\ &\quad \Big(S^*_{\boldsymbol{x}}(1) \otimes S^*_{\boldsymbol{y}}(\beta_y - \mathcal{E}_y - 1) \otimes S_{\boldsymbol{z}}(0) \otimes \bigoplus_{I \in \mathcal{I}_{1,\mathcal{E}_y,\mathcal{E}_z-\beta_z}} \bigwedge_{k\in I} e_k.\Big). \\ K_{0,\alpha+\beta_y} &\cong L_{0,1} \oplus L_{0,2} \oplus L_{0,3} \oplus L_{0,4} \qquad (6.31)\\ &= \Big(S^*_{\boldsymbol{x}}(0) \otimes S^*_{\boldsymbol{y}}(\beta_y + 1 - \mathcal{E}_y) \otimes S_{\boldsymbol{z}}(0) \otimes \bigoplus_{I \in \mathcal{I}_{0,\mathcal{E}_y-1,\mathcal{E}_z-\beta_z+1}} \bigwedge_{k\in I} e_k.\Big) \oplus \\ &\quad \Big(S^*_{\boldsymbol{x}}(0) \otimes S^*_{\boldsymbol{y}}(\beta_y - \mathcal{E}_y) \otimes S_{\boldsymbol{z}}(1) \otimes \bigoplus_{I \in \mathcal{I}_{0,\mathcal{E}_y,\mathcal{E}_z-\beta_z}} \bigwedge_{k\in I} e_k.\Big) \oplus \\ &\quad \Big(S^*_{\boldsymbol{x}}(0) \otimes S^*_{\boldsymbol{y}}(\beta_y - \mathcal{E}_y) \otimes S_{\boldsymbol{z}}(0) \otimes \bigoplus_{I \in \mathcal{I}_{1,\mathcal{E}_y-1,\mathcal{E}_z-\beta_z}} \bigwedge_{k\in I} e_k.\Big) \oplus \\ &\quad \Big(S^*_{\boldsymbol{x}}(0) \otimes S^*_{\boldsymbol{y}}(\beta_y - \mathcal{E}_y - 1) \otimes S_{\boldsymbol{z}}(1) \otimes \bigoplus_{I \in \mathcal{I}_{1,\mathcal{E}_y,\mathcal{E}_z-\beta_z-1}} \bigwedge_{k\in I} e_k.\Big). \end{aligned}$$

Let M be the Koszul resultant matrix associated to the map δ_1 in Eq. (6.29), see Prop. 3.3.13. We split M as in Def. 6.2.1 and study the right eigenvectors of the Schur complement of $M_{2,2}$ to recover the

coordinates of all the solutions of the 2-bilinear system $(f_1, \dots, f_N)$. We reduce this analysis to study a map in a strand of the Koszul complex (Sec. 2.7) of a linear system with common solutions.

We will assume that the number of different solutions is $\#\mathbb{V}_{\mathbb{P}^{\alpha}\times\mathbb{P}^{\beta_y}\times\mathbb{P}^{\beta_z}}(\langle f_1,\dots,f_N\rangle) = \binom{\mathcal{E}_y}{\beta_y}\binom{\mathcal{E}_z}{\beta_z}$, that is, the multihomogeneous Bézout bound (Lem. 6.1.11).

Let $\boldsymbol{p} = (\boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{p}_z) \in \mathbb{P}^{\alpha}\times\mathbb{P}^{\beta_y}\times\mathbb{P}^{\beta_z}$, and without loss of generality assume that $(x_0\, y_0\, z_0)(\boldsymbol{p}) \neq 0$. First, we study the kernel of $M(\boldsymbol{f_0})$, when the overdetermined system $\boldsymbol{f_0}$ *has* common solutions. We relate this kernel to the eigenvectors, as we did in the proof of Thm. 6.2.3. For each variable $t \in \{x, y, z\}$, consider the dual form (see notation in page 74)

$$\mathbb{1}^t_{\boldsymbol{p}}(d_t) := \sum_{\theta_t \in \mathcal{A}(d_t)} \frac{\boldsymbol{t}^{\theta_t}}{t_0^{d_t}}(\boldsymbol{p}_t)\ \boldsymbol{\partial t}^{\theta} \in S^*_{\boldsymbol{t}}(-d_t)$$

for $d_t \geq 0$. If $d_t < 0$, then we take $\mathbb{1}^t_{\boldsymbol{p}}(d_t) := 0$.

Observation 6.3.1. *For each variable $t \in \{x, y, z\}$, given a polynomial $g_t \in S_t(\bar{d}_t)$, such that $\bar{d}_t \leq d_t$, then the map $\mu_{(t)}$ from Eq.* (3.11) *acts over g_t and $\mathbb{1}^t_{\boldsymbol{p}}(d_t)$ as the evaluation of $\frac{g_t}{t_0^{\bar{d}_t}}$ at $\mathbf{p}$, that is*

$$\mu_{(t)}(g_t, \mathbb{1}^t_{\boldsymbol{p}}(d_t)) = \frac{g_t}{t_0^{\bar{d}_t}}(\boldsymbol{p}_t) \cdot \mathbb{1}^t_{\boldsymbol{p}}(d_t - \bar{d}_t).$$

To simplify notation, we consider $S(d_x, d_y, d_z) = S_{\boldsymbol{x}}(d_x) \otimes S_{\boldsymbol{y}}(d_y) \otimes S_{\boldsymbol{z}}(d_z)$, and given multihomogeneous polynomials $f \in S(d_x, d_y, d_z)$ and a point $(\boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{p}_z) \in \mathbb{P}^{\alpha} \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}$, we denote by $f(\boldsymbol{p}_x, \boldsymbol{p}_y) \in S_{\boldsymbol{z}}(d_z)$ the partial evaluation of the rational function $\frac{f}{x_0^{d_x} y_0^{d_y}}$ at $\boldsymbol{x} = \boldsymbol{p}_x$ and $\boldsymbol{y} = \boldsymbol{p}_y$. This specialization is well-defined because the numerator and denominator share the same degrees with respect to $\boldsymbol{x}$ and $\boldsymbol{y}$.

Lemma 6.3.2. *Consider $\mathbf{d} = (d_x, d_y, d_z)$, $\bar{\mathbf{d}} = (\bar{d}_x, \bar{d}_y, \bar{d}_z)$. Let $f \in S(\bar{\mathbf{d}})$ and $g_z \in S_{\boldsymbol{z}}(d_z)$. If $d_x \geq \bar{d}_x$ and $d_y \geq \bar{d}_y$, then the map μ from Eq.* (3.12) *acts over $\mathbb{1}^x_{\boldsymbol{p}}(d_x) \otimes \mathbb{1}^y_{\boldsymbol{p}}(d_y) \otimes g_z$ and f as the multiplication of g_z and $f(\boldsymbol{p}_x, \boldsymbol{p}_y)$, that is*

$$\mu(\mathbb{1}^x_{\boldsymbol{p}}(d_x) \otimes \mathbb{1}^y_{\boldsymbol{p}}(d_y) \otimes g_z, f) = \mathbb{1}^x_{\boldsymbol{p}}(d_x - \bar{d}_x) \otimes \mathbb{1}^y_{\boldsymbol{p}}(d_y - \bar{d}_y) \otimes \big(g_z \cdot f(\boldsymbol{p}_x, \boldsymbol{p}_y)\big).$$

Proof. Consider $f = \sum_{\sigma} c_{\sigma} \boldsymbol{x}^{\sigma_x} \boldsymbol{y}^{\sigma_y} \boldsymbol{z}^{\sigma_z}$. As μ is a bilinear map and the tensor product is multilinear, it is enough to prove this lemma only for the monomials $\boldsymbol{x}^{\sigma_x} \boldsymbol{y}^{\sigma_y} \boldsymbol{z}^{\sigma_z} \in S(\bar{\boldsymbol{d}})$. For that reason, we study the monomial case,

$$\begin{aligned}
&\mu(\mathbb{1}^x_{\alpha}(d_x) \otimes \mathbb{1}^y_{\alpha}(d_y) \otimes g_z, \boldsymbol{x}^{\sigma_x} \otimes \boldsymbol{y}^{\sigma_y} \otimes \boldsymbol{z}^{\sigma_z}) = \\
&\qquad \mu_{(x)}\Big(\boldsymbol{x}^{\sigma_x}, \mathbb{1}^x_{\alpha}(d_x)\Big) \otimes \mu_{(y)}\Big(\boldsymbol{y}^{\sigma_y}, \mathbb{1}^y_{\alpha}(d_y)\Big) \otimes \mu_{(z)}\Big(g_z, \boldsymbol{z}^{\sigma_z}\Big) = \\
&\Big(\frac{\boldsymbol{x}^{\sigma_x}}{x_0^{\bar{d}_x}}(\alpha)\, \mathbb{1}^x_{\alpha}(d_x - \bar{d}_x)\Big) \otimes \Big(\frac{\boldsymbol{y}^{\sigma_y}}{y_0^{\bar{d}_y}}(\alpha)\, \mathbb{1}^y_{\alpha}(d_y - \bar{d}_y)\Big) \otimes \Big(g_z \cdot \boldsymbol{z}^{\sigma_z}\Big) = \\
&\qquad \Big(\mathbb{1}^x_{\alpha}(d_x - \bar{d}_x)\Big) \otimes \Big(\mathbb{1}^y_{\alpha}(d_y - \bar{d}_y)\Big) \otimes \Big(g_z \cdot \frac{\boldsymbol{x}^{\sigma_x}}{x_0^{\bar{d}_x}}(\alpha) \frac{\boldsymbol{y}^{\sigma_y}}{y_0^{\bar{d}_y}}(\alpha) \cdot \boldsymbol{z}^{\sigma_z}\Big) \qquad \square
\end{aligned}$$

Let $\omega^{(1)} := \mathcal{I}_{0,\mathcal{E}_y,\mathcal{E}_z-\beta_z+1}$ and $\omega^{(2)} := \mathcal{I}_{1,\mathcal{E}_y,\mathcal{E}_z-\beta_z}$, see Eq. (6.3). With the notation of Eq. (6.30), for each $\boldsymbol{p} = (\boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{p}_z) \in \mathbb{P}^\alpha \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}$ such that $(x_0\, y_0\, z_0)(\boldsymbol{p}) \neq 0$, consider the map

$$\rho_{\boldsymbol{p}} : \mathbb{K}^{\#\omega^{(1)}} \times \mathbb{K}^{\#\omega^{(2)}} \to L_{1,1} \oplus L_{1,2},$$

such that

$$\rho_{\boldsymbol{p}}(\boldsymbol{\lambda^{(1)}}, \boldsymbol{\lambda^{(2)}}) := \sum_{I\in\omega^{(1)}} \lambda_I^{(1)} \cdot \Big(\mathbb{1}^x_{\boldsymbol{p}}(1) \otimes \mathbb{1}^y_{\boldsymbol{p}}(\mathcal{E}_y - \beta_y) \otimes 1 \otimes \bigwedge_{k\in I} e_k\Big) + \sum_{J\in\omega^{(2)}} \lambda_J^{(2)} \cdot \Big(\mathbb{1}^x_{\boldsymbol{p}}(1) \otimes \mathbb{1}^y_{\boldsymbol{p}}(\mathcal{E}_y - \beta_y + 1) \otimes 1 \otimes \bigwedge_{k\in J} e_k\Big).$$

Note that $\#\omega^{(1)} + \#\omega^{(2)} = \binom{\mathcal{E}_z+1}{\mathcal{E}_z-\beta_z+1}$. Hence, we write $\rho_{\boldsymbol{p}} : \mathbb{K}^{\binom{\mathcal{E}_z+1}{\mathcal{E}_z-\beta_z+1}} \to K_{1,\alpha+\beta_y+1}$.

Lemma 6.3.3. *Let $\delta_1(\boldsymbol{f_0}, \boldsymbol{m})$ be the map from Eq.* (6.29) *and consider $\boldsymbol{p} = (\boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{p}_z) \in \mathbb{V}_{\mathbb{P}^\alpha\times\mathbb{P}^{\beta_y}\times\mathbb{P}^{\beta_z}}(\boldsymbol{f_0})$ such that $(x_0\, y_0\, z_0)(\boldsymbol{p}) \neq 0$. The linear map*

$$\delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho_{\boldsymbol{p}} : \mathbb{K}^{\binom{\mathcal{E}_z+1}{\mathcal{E}_z-\beta_z+1}} \to K_{0,\alpha+\beta_y}$$

is equivalent to the $(\mathcal{E}_z - \beta_z + 1)$-th map of the Koszul complex of the system in Eq. (6.32)*, consisting of $\mathcal{E}_z + 1$ linear polynomials in $\boldsymbol{z}$,*

$$\boldsymbol{f_z} := \Big(f_0(\boldsymbol{p}_x, \boldsymbol{p}_y), f_{\mathcal{E}_y+1}(\boldsymbol{p}_x, \boldsymbol{p}_y), \ldots, f_n(\boldsymbol{p}_x, \boldsymbol{p}_y)\Big), \tag{6.32}$$

restricted to its 0-graded part, i.e. the strand of the Koszul complex such that its $(\mathcal{E}_z - \beta_z + 1)$-th module is isomorphic to $\mathbb{K}^{\binom{\mathcal{E}_z+1}{\mathcal{E}_z-\beta_z+1}}$.

Proof. With the notation of Eq. (6.30), we split the map ρ as $\rho(\boldsymbol{\lambda^{(1)}}, \boldsymbol{\lambda^{(2)}}) := \rho^{(1)}_{\boldsymbol{p}}(\boldsymbol{\lambda^{(1)}}) + \rho^{(2)}_{\boldsymbol{p}}(\boldsymbol{\lambda^{(2)}})$, where $\rho^{(1)}_{\boldsymbol{p}} : \mathbb{K}^{\#\omega^{(1)}} \to L_{1,1}$ is such that,

$$\rho^{(1)}_{\boldsymbol{p}}(\boldsymbol{\lambda^{(1)}}) := \sum_{I\in\omega^{(1)}} \Big(\mathbb{1}^x_{\boldsymbol{p}}(1) \otimes \mathbb{1}^y_{\boldsymbol{p}}(\mathcal{E}_y - \beta_y) \otimes \lambda_I^{(1)} \otimes \bigwedge_{k\in I} e_k\Big),$$

and $\rho^{(2)}_{\boldsymbol{p}} : \mathbb{K}^{\#\omega^{(2)}} \to L_{1,2}$ is such that

$$\rho^{(2)}_{\boldsymbol{p}}(\boldsymbol{\lambda^{(2)}}) := \sum_{J\in\omega^{(2)}} \Big(\mathbb{1}^x_{\boldsymbol{p}}(1) \otimes \mathbb{1}^y_{\boldsymbol{p}}(\mathcal{E}_y - \beta_y + 1) \otimes \lambda_J^{(2)} \otimes \bigwedge_{k\in J} e_k\Big).$$

Both maps are injective. As $\delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho_{\boldsymbol{p}} = \delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho^{(1)}_{\boldsymbol{p}} + \delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho^{(2)}_{\boldsymbol{p}}$, we study $\delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho^{(1)}_{\boldsymbol{p}}$ and $\delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho^{(2)}_{\boldsymbol{p}}$ separately. Following the definition of δ_1, see Prop. 3.3.13, it

holds,

$$\delta_1(\boldsymbol{f_0},\boldsymbol{m})\circ\rho_{\boldsymbol{p}}^{(1)} = \sum_{I\in\omega^{(1)}}\lambda_I^{(1)}\delta_1(\boldsymbol{f_0},\boldsymbol{m})\Big(\mathbb{1}_{\boldsymbol{p}}^x(1)\otimes\mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y-\beta_y)\otimes 1\otimes\bigwedge_{k\in I}e_k\Big)$$

$$= \sum_{I\in\omega^{(1)}}\lambda_I^{(1)}\left(\begin{array}{l}\displaystyle\sum_{\alpha+\beta_y+1}^{\mathcal{E}_y}(-1)^{i-1}\mu(\mathbb{1}_{\boldsymbol{p}}^x(1)\otimes\mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y-\beta_y)\otimes 1,f_{I_i})\otimes\bigwedge_{k\in I\setminus\{I_i\}}e_k+\\ \displaystyle\sum_{i=\mathcal{E}_y+1}(-1)^{i-1}\mu(\mathbb{1}_{\boldsymbol{p}}^x(1)\otimes\mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y-\beta_y)\otimes 1,f_{I_i})\otimes\bigwedge_{k\in I\setminus\{I_i\}}e_k\end{array}\right).$$

By Lem. 6.3.2, $\delta_1(\boldsymbol{f_0},\boldsymbol{m})\circ\rho_{\boldsymbol{p}}^{(1)}$ is equivalent to,

$$\sum_{I\in\omega^{(1)}}\lambda_I^{(1)}\left(\begin{array}{l}\displaystyle\sum_{i=1}^{\mathcal{E}_y}(-1)^{i-1}\mathbb{1}_{\boldsymbol{p}}^x(0)\otimes\mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y-\beta_y-1)\otimes f_{I_i}(\boldsymbol{p}_x,\boldsymbol{p}_y)\otimes\bigwedge_{k\in I\setminus\{I_i\}}e_k+\\ \displaystyle\sum_{\alpha+\beta_y+1}^{\alpha+\beta_y+1}(-1)^{i-1}\mathbb{1}_{\boldsymbol{p}}^x(0)\otimes\mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y-\beta_y)\otimes f_{I_i}(\boldsymbol{p}_x,\boldsymbol{p}_y)\otimes\bigwedge_{k\in I\setminus\{I_i\}}e_k\end{array}\right).$$

As $\boldsymbol{p}=(\boldsymbol{p}_x,\boldsymbol{p}_y,\boldsymbol{p}_z)\in\mathbb{V}_{\mathbb{P}^{\alpha}\times\mathbb{P}^{\beta_y}\times\mathbb{P}^{\beta_z}}(\boldsymbol{f_0})$ and, for $1\le i\le\mathcal{E}_y$, $f_{I_i}\in S(1,1,0)$, it holds

$$f_{I_i}(\boldsymbol{p}_x,\boldsymbol{p}_y)=f_{I_i}(\boldsymbol{p})=0.$$

We conclude that the image of $\delta_1(\boldsymbol{f_0},\boldsymbol{m})\circ\rho_{\boldsymbol{p}}^{(1)}$ belongs to $L_{0,2}$, and it holds,

$$\delta_1(\boldsymbol{f_0},\boldsymbol{m})\circ\rho_{\boldsymbol{p}}^{(1)}=\mathbb{1}_{\boldsymbol{p}}^x(0)\otimes\mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y-\beta_y)\otimes\big(\sum_{I\in\omega^{(1)}}\sum_{i=\mathcal{E}_y+1}^{\alpha+\beta_y+1}(-1)^{i-1}\lambda_I^{(1)}f_{I_i}(\boldsymbol{p}_x,\boldsymbol{p}_y)\otimes\bigwedge_{k\in I\setminus\{I_i\}}e_k\big).$$

Now consider $\delta_1(\boldsymbol{f_0},\boldsymbol{m})\circ\rho_{\boldsymbol{p}}^{(2)}$. Following a similar procedure, we deduce

$$\begin{aligned}\delta_1(\boldsymbol{f_0},\boldsymbol{m})\circ\rho_{\boldsymbol{p}}^{(2)} =\\ &\mathbb{1}_{\boldsymbol{p}}^x(0)\otimes\mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y-\beta_y)\otimes\sum_{I\in\omega^{(2)}}\Big(\lambda_I^{(2)}f_0(\boldsymbol{p}_x,\boldsymbol{p}_y)\otimes\bigwedge_{k\in I\setminus\{0\}}e_k\Big)+\\ \mathbb{1}_{\boldsymbol{p}}^x(0)\otimes\mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y-\beta_y-1)\otimes&\sum_{I\in\omega^{(2)}}\sum_{i=2}^{\mathcal{E}_y+1}\Big((-1)^{i-1}\lambda_I^{(2)}f_{I_i}(\boldsymbol{p}_x,\boldsymbol{p}_y)\otimes\bigwedge_{k\in I\setminus\{I_i\}}e_k\Big)+\\ &\mathbb{1}_{\boldsymbol{p}}^x(0)\otimes\mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y-\beta_y+1)\otimes\sum_{I\in\omega^{(2)}}\sum_{i=\mathcal{E}_y+1}^{\alpha+\beta_y+1}\Big((-1)^{i-1}\lambda_I^{(2)}f_{I_i}(\boldsymbol{p}_x,\boldsymbol{p}_y)\otimes\bigwedge_{k\in I\setminus\{I_i\}}e_k\Big).\end{aligned}$$

As $\boldsymbol{p}=(\boldsymbol{p}_x,\boldsymbol{p}_y,\boldsymbol{p}_z)\in\mathbb{V}_{\mathbb{P}^{\alpha}\times\mathbb{P}^{\beta_y}\times\mathbb{P}^{\beta_z}}(\boldsymbol{f_0})$ and, for $1\le i\le\mathcal{E}_y+1$, $f_{I_i}\in S(1,1,0)$, it holds

$$f_{I_i}(\boldsymbol{p}_x,\boldsymbol{p}_y)=f_{I_i}(\boldsymbol{p})=0.$$

Hence, the image of $\delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho_{\boldsymbol{p}}^{(2)}$ belongs to $L_{0,2} \oplus L_{0,4}$, and it holds

$$\begin{aligned}\delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho_{\boldsymbol{p}}^{(2)} = \\ \mathbb{1}_{\boldsymbol{p}}^x(0) \otimes \mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y - \beta_y) \otimes \sum_{I \in \omega^{(2)}} \Big(\lambda_I^{(2)} f_0(\boldsymbol{p}_x, \boldsymbol{p}_y) \otimes \bigwedge_{k \in I \setminus \{0\}} e_k\Big) + \\ \mathbb{1}_{\boldsymbol{p}}^x(0) \otimes \mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y - \beta_y + 1) \otimes \sum_{I \in \omega^{(2)}} \sum_{i=\mathcal{E}_y+1}^{\alpha+\beta_y+1} \Big((-1)^{i-1} \lambda_I^{(2)} f_{I_i}(\boldsymbol{p}_x, \boldsymbol{p}_y) \otimes \bigwedge_{k \in I \setminus \{I_i\}} e_k\Big).\end{aligned}$$

Using the identity $\alpha + \beta_y + \beta_z = \mathcal{E}_y + \mathcal{E}_z$, we can rewrite $\delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho_{\boldsymbol{p}} : \mathbb{K}^{\binom{\mathcal{E}_z+1}{\mathcal{E}_z-\beta_z+1}} \to L_{0,2} \oplus L_{0,4}$ as

$$(\delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho_{\boldsymbol{p}})(\boldsymbol{\lambda}) = \mathbb{1}_{\boldsymbol{p}}^x(0) \otimes \mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y - \beta_y) \otimes P_1(\boldsymbol{\lambda}) + \mathbb{1}_{\boldsymbol{p}}^x(0) \otimes \mathbb{1}_{\boldsymbol{p}}^y(\mathcal{E}_y - \beta_y + 1) \otimes (-1)^{\mathcal{E}_y} P_2(\boldsymbol{\lambda})$$

where

$$P_1(\boldsymbol{\lambda}) := \sum_{I \in \omega^{(2)}} \lambda_I f_0(\boldsymbol{p}_x, \boldsymbol{p}_y) \otimes \bigwedge_{k \in I \setminus \{0\}} e_k + \sum_{J \in \omega^{(1)}} \sum_{j=1}^{\mathcal{E}_z - \beta_z + 1} (-1)^{j-1} \lambda_J f_{J_j}(\boldsymbol{p}_x, \boldsymbol{p}_y) \otimes \bigwedge_{k \in J \setminus \{J_j\}} e_k,$$

$$P_2(\boldsymbol{\lambda}) := \sum_{I \subset \omega^{(2)}} \sum_{j=2}^{\mathcal{E}_z - \beta_z} (-1)^{r+j-1} \lambda_I f_{I_j}(\boldsymbol{p}_x, \boldsymbol{p}_y) \otimes \bigwedge_{k \in I \setminus \{I_j\}} e_k.$$

Note that the intersection between the image of P_1 and $-P_2$ is trivial. Hence, $P_1 + P_2$ vanishes if and only if P_1 and P_2 vanish. The map $\delta_1 \circ \rho_{\boldsymbol{p}}$ is equivalent to the map $\boldsymbol{\lambda} \mapsto P_1(\boldsymbol{\lambda}) + P_2(\boldsymbol{\lambda})$. Note that, for all $I \in \omega^{(1)} \cup \omega^{(2)}$, $\{1, \ldots, \mathcal{E}_y\} \subset I$. Therefore, if we expand this map we conclude that it is equivalent to the 0-graded part of the $(\mathcal{E}_z - \beta_z + 1)$-th map of the Koszul complex of the linear system $\boldsymbol{f_z}$, that is,

$$P_1(\boldsymbol{\lambda}) + P_2(\boldsymbol{\lambda}) = \sum_{\substack{J \subset \{0, \mathcal{E}_y+1, \ldots, n\} \\ \#J = \mathcal{E}_z - \beta_z + 1}} \sum_{j=1}^{\mathcal{E}_z - \beta_z + 1} (-1)^{j-1} \lambda_J f_{J_j}(\boldsymbol{p}_x, \boldsymbol{p}_y) \otimes \bigwedge_{k \in \{1 \ldots \mathcal{E}_y\} \cup J \setminus \{J_j\}} e_k \,.$$

□

If $\boldsymbol{f_0}$ has a solution $(\boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{p}_z) \in \mathbb{V}_{\mathbb{P}^{\alpha} \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}}(\boldsymbol{f_0})$, then, $\boldsymbol{p}_z$ is a solution of the linear system $\boldsymbol{f_z}$, that is $\boldsymbol{p}_z \in \mathbb{V}_{\mathbb{P}^{\beta_z}}(\boldsymbol{f_z})$. As $\boldsymbol{f_z}$ is an overdetermined system, the Koszul complex of $\boldsymbol{f_z}$ is not exact [Lan02, Thm. XXI.4.6].

Lemma 6.3.4. *Let $\boldsymbol{f_0}$ be an overdetermined 2-bilinear system. If $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^{\alpha} \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}}(\boldsymbol{f_0})$, then there is a non-zero $\widehat{\boldsymbol{\lambda}}_{\boldsymbol{p}} \in \mathbb{K}^{\binom{s+1}{s-n_z+1}}$ such that $\delta_1(\boldsymbol{f_0}, \boldsymbol{m}) \circ \rho_{\boldsymbol{p}}(\widehat{\boldsymbol{\lambda}}_{\boldsymbol{p}}) = 0$.*

Proof. Following Lem. 6.3.3, if we compose $\delta_1(\boldsymbol{f}, \boldsymbol{m})$ and $\rho_{\boldsymbol{p}}$, then we obtain a map which is similar to the 0-graded part of the $(\mathcal{E}_z - \beta_z + 1)$-th map of the Koszul complex of the $\mathcal{E}_z + 1$ linear polynomials in $\boldsymbol{z}$, $\boldsymbol{f_z}$ (Eq. 6.32). As the linear system $\boldsymbol{f_z}$ has a solution $\boldsymbol{p}_z$, at most β_z of its polynomials are linearly independent. Hence, the Koszul complex of $\boldsymbol{f_z}$ is isomorphic to a Koszul complex $\mathcal{K}(\widetilde{f}_1, \ldots, \widetilde{f}_{\beta_z}, 0, \ldots, 0)$ of a system of $\mathcal{E}_z + 1$ linear polynomials, where $(\mathcal{E}_z - \beta_z + 1)$ of them are equal to zero [Lan02, Lem. XXI.4.2]. The $(\mathcal{E}_z - \beta_z + 1)$-th map of $\mathcal{K}(\widetilde{f}_1, \ldots, \widetilde{f}_{\beta_z}, 0, \ldots, 0)$ maps $e_{\beta_z+1} \wedge \ldots \wedge e_{\mathcal{E}_z - \beta_z + 1}$ to zero. Hence, its 0-graded part has a non-trivial kernel, and so there is a non-zero $\widehat{\boldsymbol{\lambda}}_{\boldsymbol{p}} \in \mathbb{K}^{\binom{\mathcal{E}_z}{\varepsilon_z - \beta_z + 1}}$ such that $\delta_1(\boldsymbol{f}, \boldsymbol{m}) \circ \rho_{\boldsymbol{p}}(\widehat{\boldsymbol{\lambda}}_{\boldsymbol{p}}) = 0$. □

Theorem 6.3.5. *Let $\boldsymbol{f} := (f_1, \ldots, f_N)$ be a square 2-bilinear system such that it has $\binom{\mathcal{E}_y}{\beta_y} \cdot \binom{\mathcal{E}_z}{\beta_z}$ different solutions over $\mathbb{P}^\alpha \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}$. Consider $\boldsymbol{\theta} \in \mathcal{A}(1,1,1)$ such that the multiprojective resultant of the system $(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1, \ldots, f_n)$ does not vanish, that is*

$$\texttt{res}(\boldsymbol{x}^{\boldsymbol{\theta}}, f_1, \ldots, f_n) \neq 0.$$

Let $f_0 \in S(1,1,1)$ be a trilinear polynomial such that $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}$ separates the elements in $\mathbb{V}_{\mathbb{P}^\alpha \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}}(\boldsymbol{f})$. Let M be the Koszul resultant matrix related to Koszul-type resultant formula detailed at the beginning of the section (page 143). Then, the Schur complement of $M_{2,2}(\boldsymbol{f_0})$ is diagonalizable, each eigenvalue equals $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$, for $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^\alpha \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}}(f_1, \ldots, f_N)$, and we can extend the eigenvector $\bar{v}_{\boldsymbol{p}}$ related to $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$ to $v_{\boldsymbol{p}} := \left[\begin{smallmatrix} M_{1,1}^{-1} \cdot M_{2,1} \\ I \end{smallmatrix}\right](\boldsymbol{f_0}) \cdot \bar{v}_{\boldsymbol{p}}$ such that $v_{\boldsymbol{p}}$ is of the form $\rho_{\boldsymbol{p}}(\widehat{\boldsymbol{\lambda}}_{\boldsymbol{p}})$, for some $\widehat{\boldsymbol{\lambda}}_{\boldsymbol{p}} \in \mathbb{K}^{\binom{\mathcal{E}_z+1}{\varepsilon_z - \beta_z + 1}}$.

Proof. By Cor. 6.2.5, the Schur complex of $M_{2,2}(\boldsymbol{f_0})$ is diagonalizable and every eigenvalue is different. For each $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^\alpha \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}}(\boldsymbol{f})$, consider the eigenvalue $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$, its related eigenvector $\bar{v}_{\boldsymbol{p}}$, and the system $\boldsymbol{g_p} := (f_0 - \frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})\, \boldsymbol{x}^{\boldsymbol{\theta}}, f_1, \ldots, f_N)$. As $\boldsymbol{p} \in \mathbb{V}_{\mathbb{P}^\alpha \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}}(\boldsymbol{g_p})$, by Lem. 6.3.4, there is a $\boldsymbol{\lambda_p} \in \mathbb{K}$ such that $\delta_1(\boldsymbol{g_p}, \boldsymbol{m}) \circ \rho(\boldsymbol{\lambda_p}) = 0$. Hence, there is a vector $w_{\boldsymbol{p}}$ in the kernel of $M(\boldsymbol{g_p})$, representing the element $\rho(\boldsymbol{\lambda_p})$. Following the proof of Thm. 6.2.3, each element in the kernel of the Schur complement of $M_{2,2}(\boldsymbol{g_p})$ is related to an eigenvector of the Schur complement of $M_{2,2}(\boldsymbol{f_0})$ with corresponding eigenvalue $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p})$. As for each eigenvalue we have only one eigenvector, then the dimension of this kernel is one. Hence, the truncation of $w_{\boldsymbol{p}}$, $\bar{w}_{\boldsymbol{p}} := \left[0 \,|\, I\right] \cdot w_{\boldsymbol{p}}$, is a multiple of $\bar{v}_{\boldsymbol{p}}$, where 0 is the zero matrix of appropriate dimension and I the identity.

As $M_{1,1}(\boldsymbol{g_p})$ is invertible and $M(\boldsymbol{g_p}) \cdot w_{\boldsymbol{p}} = 0$, it holds that $\left[\begin{smallmatrix} M_{1,1}^{-1} \cdot M_{2,1} \\ I \end{smallmatrix}\right](\boldsymbol{g_p}) \bar{w}_{\boldsymbol{p}} = w_{\boldsymbol{p}}$. As $\left[\begin{smallmatrix} M_{1,1}^{-1} \cdot M_{2,1} \\ I \end{smallmatrix}\right](\boldsymbol{g_p})$ does not involve $u_{0,\theta}$, then $\left[\begin{smallmatrix} M_{1,1}^{-1} \cdot M_{2,1} \\ I \end{smallmatrix}\right](\boldsymbol{g_p}) = \left[\begin{smallmatrix} M_{1,1}^{-1} \cdot M_{2,1} \\ I \end{smallmatrix}\right](\boldsymbol{f_0})$. Therefore, we conclude that, as $\bar{v}_{\boldsymbol{p}}$ is a multiple of $\bar{w}_{\boldsymbol{p}}$, then $v_{\boldsymbol{p}} = \left[\begin{smallmatrix} M_{1,1}^{-1} \cdot M_{2,1} \\ I \end{smallmatrix}\right](\boldsymbol{f_0}) \cdot \bar{v}_{\boldsymbol{p}}$ is a multiple of $w_{\boldsymbol{p}}$. □

Hence, we can recover the coordinates $(\boldsymbol{p}_x, \boldsymbol{p}_y)$ of $(\boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{p}_z) \in \mathbb{V}_{\mathbb{P}^\alpha \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}}(f_1, \ldots, f_N)$ by inverting a monomial map as we do in Prop. 5.2.4. Algorithm 6 summarizes our strategy to solve 2-bilinear systems.

6.3.1 Example: Solving 2-bilinear systems

We consider a *2-bilinear system* such that $\alpha = \beta_1 = \beta_2 = 1$, $\mathcal{E}_1 = 2$ and $\mathcal{E}_2 = 1$. The following (Eq. 6.33) is a square 2-bilinear system and has two solutions over $\mathbb{P}^\alpha \times \mathbb{P}^{\beta_y} \times \mathbb{P}^{\beta_z}$, namely $\boldsymbol{p}_1 := (1 :$

Algorithm 6 `Solve2Bilinear`

Input: $(\bar{f}_1, \dots, \bar{f}_N)$ is a square 2-bilinear system such that $\mathbb{V}_{\mathbb{P}^{\alpha}\times\mathbb{P}^{\beta_y}\times\mathbb{P}^{\beta_z}}(\bar{f}_1, \dots, \bar{f}_N)$ is finite and has no multiplicities.

1: $A \leftarrow$ Random linear change of coordinates preserving the structure.
2: $(f_1, \dots, f_N) \leftarrow (\bar{f}_1 \circ A, \dots, \bar{f}_N \circ A)$. (Cor. 6.2.8)
3: $f_0 \leftarrow$ Random trilinear polynomial in $S(1,1,1)$.
4: $\left[\begin{smallmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{smallmatrix}\right] \leftarrow$ Matrix corresponding to $\delta_1(\boldsymbol{m}, (f_0, \dots, f_N))$, split with respect to the monomial $\boldsymbol{x}^{\boldsymbol{\theta}}$. (Def. 6.2.1)
5: $\left\{\left(\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}), \bar{v}_{\boldsymbol{p}}\right)\right\}_{\boldsymbol{p}} \leftarrow$ Set of pairs Eigenvalue-Eigenvector of the Schur complement of $M_{2,2}$. (Thm. 6.2.3)
6: **for all** $\left(\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}), \bar{v}_{\boldsymbol{p}}\right) \in \left\{\left(\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}), \bar{v}_{\boldsymbol{p}}\right)\right\}_{\boldsymbol{p}}$ **do**
7: Extract the coordinates $\boldsymbol{p}_x, \boldsymbol{p}_y$ from $\rho_{\boldsymbol{p}}(\widehat{\boldsymbol{\lambda}}_{\boldsymbol{p}})$ by recovering it from $\left[\begin{smallmatrix} M_{1,1}^{-1}\cdot M_{2,1} \\ I \end{smallmatrix}\right] \cdot \bar{v}$. (Thm. 6.3.5)
8: Let $\boldsymbol{p}_z \in \mathbb{P}^{n_z}$ be the unique solution to the linear system, over $\mathbb{K}[\boldsymbol{z}]$, given by

$$\{f_1(\boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{z}) = 0, \dots, f_N(\boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{z}) = 0\}.$$

9: Recover the solution of the system $(\bar{f}_1, \dots, \bar{f}_N)$, as $A\big((\boldsymbol{p}_x, \boldsymbol{p}_y, \boldsymbol{p}_z)\big)$.
10: **end for**

$1\ ;\ 1{:}1\ ;\ 1{:}1)$ and $\boldsymbol{p}_2 := (1{:}3\ ;\ 1{:}2\ ;\ 1{:}3)$.

$$\begin{cases} f_1 := 7x_0y_0 - 8x_0y_1 - x_1y_0 + 2x_1y_1 \\ f_2 := -5x_0y_0 + 7x_0y_1 - x_1y_0 - x_1y_1 \\ f_3 := -6x_0z_0 + 9x_0z_1 - x_1z_0 - 2x_1z_1 \end{cases}. \tag{6.33}$$

We introduce a trilinear polynomial f_0 and consider the overdetermined 2-bilinear system $\boldsymbol{f_0} := (f_0, f_1, f_2, f_3)$, where

$$\begin{aligned} f_0 := {} & 3\,x_0y_0z_0 - x_0y_0z_1 - 4\,x_0y_1z_0 + 2\,x_0y_1z_1 \\ & + x_1y_0z_0 + 2\,x_1y_0z_1 + 2\,x_1y_1z_0 - 2\,x_1y_1z_1. \end{aligned}$$

In this case, the polynomial f_0 corresponds to case (4) in the list appearing in the page 115. According to Thm. 6.1.8, we can construct a Koszul-type determinantal formula for the resultant of $\boldsymbol{f_0}$ by considering the degree vector $\boldsymbol{m} = (0, -1, 1)$. To construct the associated Koszul resultant matrix, we consider the following monomial basis,

Basis of $K_{1,3}$ (Columns)	
(A)	$\partial x_0 \partial y_1^2 \boldsymbol{e}_{\{0,1,2\}}$
(B)	$\partial x_1 \partial y_0^2 \boldsymbol{e}_{\{0,1,2\}}$
(C)	$\partial x_1 \partial y_1^2 \boldsymbol{e}_{\{0,1,2\}}$
(D)	$\partial x_0 \partial y_0 \boldsymbol{e}_{\{1,2,3\}}$
(E)	$\partial x_0 \partial y_1 \boldsymbol{e}_{\{1,2,3\}}$
(F)	$\partial x_1 \partial y_0 \boldsymbol{e}_{\{1,2,3\}}$
(G)	$\partial x_1 \partial y_1 \boldsymbol{e}_{\{1,2,3\}}$
(H)	$\partial x_0 \partial y_0 \partial y_1 \boldsymbol{e}_{\{0,1,2\}}$
(I)	$\partial x_0 \partial y_0^2 \boldsymbol{e}_{\{0,1,2\}}$
(J)	$\partial x_1 \partial y_0 \partial y_1 \boldsymbol{e}_{\{0,1,2\}}$

Basis of $K_{0,2}$ (Rows)	
(I)	$\boldsymbol{e}_{\{1,3\}}$
(II)	$\boldsymbol{e}_{\{2,3\}}$
(III)	$\partial y_0 \boldsymbol{e}_{\{0,1\}}$
(IV)	$\partial y_1 \boldsymbol{e}_{\{0,1\}}$
(V)	$\partial y_0 \boldsymbol{e}_{\{0,2\}}$
(VI)	$\partial y_1 \boldsymbol{e}_{\{0,2\}}$
(VII)	$\partial y_0 z_1 \boldsymbol{e}_{\{1,2\}}$
(VIII)	$\partial y_1 z_1 \boldsymbol{e}_{\{1,2\}}$
(IX)	$\partial y_0 z_0 \boldsymbol{e}_{\{1,2\}}$
(X)	$\partial y_1 z_0 \boldsymbol{e}_{\{1,2\}}$

The following matrix represents the map $\delta_1(\boldsymbol{m}, \boldsymbol{f_0})$ of the Weyman complex $K_\bullet(\boldsymbol{m}; \boldsymbol{f})$ (Def. 3.3.3), with respect to the basis above.

	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)
(I)	0	0	0	**5**	**−7**	**1**	**1**	0	0	0
(II)	0	0	0	**7**	**−8**	**−1**	**2**	0	0	0
(III)	0	**−1**	0	0	0	0	0	**−1**	**−5**	**7**
(IV)	**7**	0	**−1**	0	0	0	0	**−1**	0	**−5**
(V)	0	**1**	0	0	0	0	0	**−2**	**−7**	**8**
(VI)	**8**	0	**−2**	0	0	0	0	**1**	0	**−7**
(VII)	0	**2**	0	**9**	0	**−2**	0	**−2**	**−1**	**2**
$(VIII)$	**2**	0	**−2**	0	**9**	0	**−2**	**2**	0	**−1**
(IX)	0	**1**	0	**−6**	0	**−1**	0	**2**	**3**	**−4**
(X)	**−4**	0	**2**	0	**−6**	0	**−1**	**1**	0	**3**

The splitting of the matrix corresponds to its partition as $\left[\begin{smallmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{smallmatrix}\right]$ with respect to the exponent of the monomial $x_0\, y_0\, z_0$ (Def. 6.2.1).

The Schur complement of this matrix is $\left[\begin{smallmatrix} 5 & -2 \\ 4 & -1 \end{smallmatrix}\right]$. Its characteristic polynomial is $T^2 - 4\,T + 3$, and so its eigenvalues are $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}_1) = 3$ and $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}_2) = 1$, in accordance to Thm. 6.2.3 and Lem. 6.2.6.

The eigenvector of $\frac{f_0}{\boldsymbol{x}^{\boldsymbol{\theta}}}(\boldsymbol{p}_2) = 1$ is $\bar{v}_{\boldsymbol{p}_2} := (1, 2)^\top$. By extending $\bar{v}_{\boldsymbol{p}_2}$, we get

$$v_{\boldsymbol{p}_2} := \left[\begin{smallmatrix} M_{1,1}^{-1} \cdot M_{2,1} \\ I \end{smallmatrix}\right](\boldsymbol{f_0}) \cdot \left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right) = (4, 3, 12, 1, 2, 3, 6, 6, 1, 2)^\top$$

which represents $\rho_{\boldsymbol{p}_2}(1, 1) =$

$$\left(\boldsymbol{\partial x}^{(1,0)} + 3\,\boldsymbol{\partial x}^{(0,1)}\right) \otimes \left(\boldsymbol{\partial y}^{(2,0)} + 2\,\boldsymbol{\partial y}^{(1,1)} + 4\,\boldsymbol{\partial y}^{(0,2)}\right) \otimes 1 \otimes \boldsymbol{e}_{\{0,1,2\}} + \left(\boldsymbol{\partial x}^{(1,0)} + 3\,\boldsymbol{\partial x}^{(0,1)}\right) \otimes \left(\boldsymbol{\partial y}^{(1,0)} + 2\,\boldsymbol{\partial y}^{(0,1)}\right) \otimes 1 \otimes \boldsymbol{e}_{\{1,2,3\}}$$

Hence, $\mathbb{1}^x_{\boldsymbol{p}_2}(1) = \left(1\,\boldsymbol{\partial x}^{(1,0)} + 3\,\boldsymbol{\partial x}^{(0,1)}\right)$, and so $\boldsymbol{p}_{2,x} = (1 : 3) \in \mathbb{P}^1$. Also, $\mathbb{1}^y_{\boldsymbol{p}_2}(1) = \left(1\,\boldsymbol{\partial y}^{(1,0)} + 2\,\boldsymbol{\partial y}^{(0,1)}\right)$, and then $\boldsymbol{p}_{2,y} = (1 : 2) \in \mathbb{P}^1$. We note that

$$\mathbb{1}^y_{\boldsymbol{p}_2}(2) = \left(1 \cdot 1 \cdot \boldsymbol{\partial y}^{(2,0)} + 1 \cdot 2 \cdot \boldsymbol{\partial y}^{(1,1)} + 2 \cdot 2 \cdot \boldsymbol{\partial y}^{(0,2)}\right).$$

We can recover $\boldsymbol{p}_{2,z}$ as the solution of $\boldsymbol{f}(\boldsymbol{p}_{2,x}, \boldsymbol{p}_{2,y}, \boldsymbol{z}) = 0$,

$$\begin{cases} f_1(\boldsymbol{p}_{2,x}, \boldsymbol{p}_{2,y}, \boldsymbol{z}) = f_2(\boldsymbol{p}_{2,x}, \boldsymbol{p}_{2,y}, \boldsymbol{z}) = 0 \\ f_3(\boldsymbol{p}_{2,x}, \boldsymbol{p}_{2,y}, \boldsymbol{z}) = -9\,z_0 + 3\,z_1 \end{cases}$$

Thus, $\boldsymbol{p}_{2,z} = (1 : 3) \in \mathbb{P}^1$ and so $\boldsymbol{p}_2 = (1\!:\!3\ ;\ 1\!:\!2\ ;\ 1\!:\!3) \in \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$.

6.4 Applications: Multiparameter Eigenvalue Problem

A motivating application for the systems and the determinantal formulas that we study in this chapter comes from the *multiparameter eigenvalue problem* (MEP). We can model MEP using *star multilinear systems*. The resultant matrices that we construct together with the eigenvalue and eigenvector criterion for polynomial systems, see Sec. 5.2.1, lead to a novel approach for solving MEP. We illustrate it through a detailed example in Sec. 6.4.1. Because of the importance of MEP in various areas of computational and applied mathematics we give a detailed presentation of the problem, we mention some of the related works, and we show how to model it using polynomial systems.

MEP is a generalization of the classical eigenvalue problem. It arises in mathematical physics as a way of solving ordinary and partial differential equations when we can use separation of variables (Fourier method) to solve boundary eigenvalue problems. Its applications include Spectral theory and Sturm-Liouville theory, among others [Atk72, Vol88, AM10, HKP04, GHPR12]. MEP allows to solve many different eigenvalue problems, for example the polynomial and the quadratic two-parameter eigenvalue problems [GLR05, MP10].

The precise definition of the problem is as follows. Assume $\alpha \in \mathbb{N}$, $\beta_1, \ldots, \beta_\alpha \in \mathbb{N}$, and square matrices $\{M^{(i,j)}\}_{0\leq j\leq \alpha} \in \mathbb{K}^{(\beta_i+1)\times(\beta_i+1)}$, for $0 \leq i \leq \alpha$. MEP consists in finding $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_\alpha) \in \mathbb{P}^\alpha(\mathbb{K})$ and $\boldsymbol{v}_1 \in \mathbb{P}^{\beta_1}, \ldots, \boldsymbol{v}_\alpha \in \mathbb{P}^{\beta_\alpha}$ such that

$$\left\{ \quad \left(\sum\nolimits_{j=0}^{\alpha} \lambda_j\, M^{(1,j)}\right) \boldsymbol{v}_1 = 0, \cdots, \left(\sum\nolimits_{j=0}^{\alpha} \lambda_j\, M^{(\alpha,j)}\right) \boldsymbol{v}_\alpha = 0 \quad \right\}. \tag{6.34}$$

We refer to $\boldsymbol{\lambda}$ as the *MEP-eigenvalue*, $\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_\alpha$ as the *MEP-eigenvector*, and to $(\boldsymbol{\lambda}, \boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_\alpha)$ as an *MEP-eigenpair*.

We model MEP as a mixed square bilinear system. For this, we introduce the variables $(x_0, \ldots, x_\alpha)$ to represent the MEP-eigenvalues and, for each $1 \leq i \leq \alpha$, the vectors $(y_{i,0}, \ldots, y_{i,\beta_i})$ to represent the MEP-eigenvectors. In this way, we obtain a bilinear polynomial system $\boldsymbol{f} = (f_{1,0}, \ldots, f_{\alpha,\beta_\alpha})$, where for each $1 \leq t \leq \alpha$,

$$\left\{ \begin{pmatrix} f_{t,0} \\ \vdots \\ f_{t,\beta_t} \end{pmatrix} := \left(\sum_{j=0}^{\alpha} x_j\, M^{(t,j)}\right) \cdot \begin{pmatrix} y_{t,0} \\ y_{t,1} \\ \vdots \\ y_{t,\beta_t} \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{\beta_t}\sum_{j=0}^{\alpha} M_{i,0}^{(t,j)}\, x_j\, y_{t,i} \\ \vdots \\ \sum_{i=0}^{\beta_t}\sum_{j=0}^{\alpha} M_{i,\beta_t}^{(t,j)}\, x_j\, y_{t,i} \end{pmatrix}, \right. \tag{6.35}$$

and, for each $1 \leq t \leq \alpha$, $f_{t,0}, \ldots, f_{t,\beta_t} \in \mathbb{K}[x_0, \ldots, x_\alpha]_1 \otimes \mathbb{K}[y_{t,0}, \ldots, y_{t,\beta_t}]_1$; that is $f_{t,1}, \ldots, f_{t,\beta_t}$ are bilinear polynomials in the blocks of variables $\{x_0, \ldots, x_\alpha\}$ and $\{y_{t,0}, \ldots, y_{t,\beta_t}\}$.

There is an one-to-one correspondence between the MEP-eigenpairs and the solutions of $\boldsymbol{f}$. That is,

$$(\boldsymbol{\lambda}, \boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_\alpha) \text{ is an MEP-eigenpair } \{M^{(i,j)}\} \Longleftrightarrow (\boldsymbol{\lambda} \otimes \boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_\alpha) \in \mathbb{V}_{\mathbb{P}^\alpha \times \mathbb{P}^{\beta_1} \times \cdots \times \mathbb{P}^{\beta_\alpha}}(\boldsymbol{f}).$$

The standard method to solve MEP is Atkinson's *Delta method* [Atk72, Ch. 6, 8]. For each $0 \leq k \leq \alpha$, it considers the overdetermined system $\boldsymbol{f}_{(k)}$ resulting from $\boldsymbol{f}$ (Eq. 6.35) by setting $x_k = 0$. Then, it constructs a matrix which is nonsingular if and only if $\boldsymbol{f}_{(k)}$ has no solutions [Atk72, Eq. 6.4.4]. Subsequently, it applies linear algebra operations to these matrices to solve the MEP $\boldsymbol{f}$ [Atk72, Thm. 6.8.1]

. It turns out that these matrices are determinantal formulas for the resultants of the corresponding overdetermined systems $\boldsymbol{f}_{(k)}$. The degree of the determinantal formulas is α [Atk72, Thm. 8.2.1], see Def. 3.2.14. One limitation of this method is that the Delta method can solve only *nonsingular MEPs*, that is, when there is a linear form in $\mathbb{K}[x_0, \ldots, x_\alpha]$ such that it does not vanish at any MEP-eigenvalue [Atk72, Thm. 8.7.1].

There are also recent algorithms that exploit homotopy continuation methods as the *diagonal coefficient homotopy method* [DYY16] and the *fiber product homotopy method* [RLY18]. These methods seems to be slower than the Delta method but, as the construction of the matrices of the Delta method is computationally expensive, they can tackle MEP of bigger size. Also, in some cases, these algorithms can solve *singular MEPs*. We can also use general purpose polynomial system solving algorithms that exploit sparsity to tackle MEP. We refer reader to [FSEDS11], see also [Spa12], for an algorithm to solve unmixed multilinear systems using Gröbner bases, and to [EMT16] using resultants. We also mention our Gröbner-basis-based algorithm to solve square mixed multihomogeneous systems, see Sec. 8.2.

The system in Eq. (6.35) is a particular case of a *star multilinear system* (Def. 6.1.1), where $A = 1$, $B = \alpha$ and, for each $j \in [B]$, $\mathcal{E}_j = \beta_j + 1$. Its expected number of solutions is $\prod_{j=1}^{\alpha}(\beta_j + 1)$ (Lem. 6.1.11). We introduce an extra linear polynomial in $\mathbb{K}[x_0, \ldots, x_\alpha]$ and, using the determinantal formulas for these systems (Thm. 6.1.8), we propose an alternative to the Delta method for solving *nonsingular MEP*. The degree of our formulas is one, meanwhile the one of the Delta method is α. Moreover, our Koszul resultant matrices are structured matrices. Using eigenvalues and eigenvectors computations we recover the solutions of MEP. We illustrate our approach by a detailed example in the following section.

6.4.1 Example: Two-Parameter Eigenvalue Problem

We consider the a *nonsingular* MEP given by the matrices

$$\begin{array}{lll} M^{(1,0)} := \begin{bmatrix} -7 & -3 \\ -8 & -2 \end{bmatrix} & M^{(1,1)} := \begin{bmatrix} 12 & 2 \\ 13 & 1 \end{bmatrix} & M^{(1,2)} := \begin{bmatrix} -7 & -1 \\ -7 & -1 \end{bmatrix} \\ M^{(2,0)} := \begin{bmatrix} -11 & -3 \\ 4 & 1 \end{bmatrix} & M^{(2,1)} := \begin{bmatrix} 7 & -1 \\ 1 & 2 \end{bmatrix} & M^{(2,1)} := \begin{bmatrix} -4 & 0 \\ -1 & -1 \end{bmatrix} \end{array} . \tag{6.36}$$

This is a *two-parameter eigenvalue problem* (2EP); such MEPs appear in physics, see [GHPR12]. For simplicity, we will name the three blocks of variables as $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$, instead of $\boldsymbol{X_1}, \boldsymbol{Y_1}, \boldsymbol{Y_2}$. Following Eq. 6.35, we write the 2EP as the following bilinear system

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} -7\,x_0 + 12\,x_1 - 7\,x_2 & -3\,x_0 + 2\,x_1 - x_2 \\ -8\,x_0 + 13\,x_1 - 7\,x_2 & -2\,x_0 + x_1 - x_2 \end{bmatrix} \cdot \begin{bmatrix} y_0 \\ y_1 \end{bmatrix}$$
$$\begin{bmatrix} f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} -11\,x_0 + 7\,x_1 - 4\,x_2 & -3\,x_0 - x_1 \\ 4\,x_0 + x_1 - x_2 & x_0 + 2\,x_1 - x_2 \end{bmatrix} \cdot \begin{bmatrix} z_0 \\ z_1 \end{bmatrix} .$$

According to Lem. 6.1.11, the 2EP should have 4 different solutions. To solve this system, we will introduce a linear polynomial $f_0 \in \mathbb{K}[\boldsymbol{X}]$ that separates the MEP-eigenvalue. That is, if $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$

are different MEP-eigenvalues, then $\frac{f_0}{x_i}(\boldsymbol{\lambda}_1) \neq \frac{f_0}{x_i}(\boldsymbol{\lambda}_2)$, for $x_i \in \boldsymbol{X}$ (Def. 6.2.4). Then, we consider a Sylvester-type determinantal formula for the resultant of $(f_0, \ldots, f_4)$ (Cor. 6.1.10) and we solve the original system using eigenvalue and eigenvector computations as in sections 6.2 and 6.3.

We assume that the MEP is *nonsingular*. Hence, any generic $f_0 \in \mathbb{K}[\boldsymbol{X}]$ separates the MEP-eigenvalues. In our case, we choose $f_0 := -x_0 + 5\,x_1 - 3\,x_2$. Following Cor. 6.1.10, there is a Sylvester-type formula for the resultant of the system $\boldsymbol{f} := (f_0, \ldots, f_4)$ using the degree vector $\boldsymbol{m} := (1,1,1)$. The latter is related to the determinantal data $(\{1,2\}, \emptyset, 0)$ (Def. 6.1.6). The Weyman complex reduces to

$$0 \to \left(\begin{array}{c} S_X(0) \otimes S_Y(1) \otimes S_Z(1) \otimes \{e_0\} \\ \oplus\, S_X(0) \otimes S_Y(0) \otimes S_Z(1) \otimes \{e_1 \oplus e_2\} \\ \oplus\, S_X(0) \otimes S_Y(1) \otimes S_Z(0) \otimes \{e_3 \oplus e_4\} \end{array} \right) \xrightarrow{\delta_1(\boldsymbol{m},\boldsymbol{f})} (S_X(1) \otimes S_Y(1) \otimes S_Z(1) \otimes \mathbb{K}) \to 0,$$

where the map $\delta_1(\boldsymbol{m}, \boldsymbol{f})$ is a Sylvester map (Rem. 3.3.17). Hence, the resultant of $\boldsymbol{f}$ is the determinant of a matrix C representing this map, which has dimensions 12×12 (Case 1, Lem. 6.1.12). We split C in $\left[\begin{smallmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \end{smallmatrix}\right]$ according to Def. 6.2.1 and we have:

$$C = \left[\begin{array}{c|cccccccc|cccc} & x_2y_0z_0 & x_2y_0z_1 & x_2y_1z_0 & x_2y_1z_1 & x_1y_0z_0 & x_1y_0z_1 & x_1y_1z_0 & x_1y_1z_1 & x_0y_0z_0 & x_0y_0z_1 & x_0y_1z_0 & x_0y_1z_1 \\ \hline z_0e_1 & -7 & & -1 & & 12 & & 2 & & -7 & & -3 & \\ z_1e_1 & & -7 & & -1 & & 12 & & 2 & & -7 & & -3 \\ z_0e_2 & -7 & & -1 & & 13 & & 1 & & -8 & & -2 & \\ z_1e_2 & & -7 & & -1 & & 13 & & 1 & & -8 & & -2 \\ y_0e_3 & -4 & & & & 7 & -1 & & & -11 & -3 & & \\ y_1e_3 & & & -4 & & & & 7 & -1 & & & -11 & -3 \\ y_0e_4 & -1 & -1 & & & 1 & 2 & & & 4 & 1 & & \\ y_1e_4 & & & -1 & -1 & & & 1 & 2 & & & 4 & 1 \\ \hline y_0z_0e_0 & -3 & & & & 5 & & & & -1 & & & \\ y_0z_1e_0 & & -3 & & & & 5 & & & & -1 & & \\ y_1z_0e_0 & & & -3 & & & & 5 & & & & -1 & \\ y_1z_1e_0 & & & & -3 & & & & 5 & & & & -1 \end{array}\right].$$

By Prop. 6.2.7, as the system $(f_1, \ldots, f_4)$ has no solutions such that $x_0 = 0$, the matrix $C_{1,1}$ is nonsingular. Hence, by Lem. 6.2.6, we have an one-to-one correspondence between the MEP-eigenvalues and the eigenvalues of the Schur complement of $C_{2,2}$, $\widetilde{C_{2,2}} := C_{2,2} - C_{2,1}C_{1,1}^{-1}C_{1,2}$. Each eigenvalue of $\widetilde{C_{2,2}}$ is the evaluation of $\frac{f_0}{x_0}$ at a MEP-eigenvalue of the original 2EP. In our case

$$\widetilde{C_{2,2}} = \begin{bmatrix} \frac{7}{4} & 0 & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{3}{4} & \frac{3}{2} & \frac{9}{4} & 2 \\ -\frac{21}{4} & -3 & \frac{27}{4} & \frac{5}{2} \\ \frac{69}{4} & \frac{19}{2} & -\frac{63}{4} & -6 \end{bmatrix}.$$

Remark 6.4.1. *If the original MEP is* nonsingular*, after performing a generic linear change of coordinates in the variables $\boldsymbol{X}$, we can assume that there is no solution of $(f_1, \ldots, f_n)$ such that $x_0 = 0$.*

Let $\alpha_1, \ldots, \alpha_4$ be the four solutions of $(f_1, \ldots, f_4)$. Then, the eigenvalues of $\widetilde{C_{2,2}}$ are $\frac{f_0}{x_0}(\alpha_1) = 1$, $\frac{f_0}{x_0}(\alpha_2) = 2$, $\frac{f_0}{x_0}(\alpha_3) = 3$ and $\frac{f_0}{x_0}(\alpha_4) = -2$. As $\delta_1(\boldsymbol{m}, \boldsymbol{f})$ is a Sylvester-type map, the right

eigenspaces of $\widetilde{C_{2,2}}$ contain the vector of monomials $\boldsymbol{v} := \begin{bmatrix} x_0y_0z_0 \\ x_0y_0z_1 \\ x_0y_1z_0 \\ x_0y_1z_1 \end{bmatrix}$ evaluated at each solutions of $(f_1, \ldots, f_4)$ Prop. 5.2.4. If each eigenspace has dimension one, then we can recover some coordinates of the solutions by inverting the monomial map given by $\boldsymbol{v}$.

$\boldsymbol{v}$	α_1	α_2	α_3	α_4
$x_0y_0z_0$	1	1	1	1
$x_0y_0z_1$	-3	-1	-2	-3
$x_0y_1z_0$	1	1	-1	-3
$x_0y_1z_1$	-3	-1	2	9

To compute the remaining coordinates, either we substitute the computed coordinates of the solutions in the original system and we solve a linear system, or we extend each eigenvector $\boldsymbol{v}(\alpha_i)$ to $\boldsymbol{w}(\alpha_i)$, where $\boldsymbol{w}(\alpha_i)$ is the solution of the following linear system:

$$\begin{bmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \end{bmatrix} \boldsymbol{w}(\alpha_i) = \frac{f_0}{x_0}(\alpha_i) \begin{bmatrix} 0 \\ \boldsymbol{v}(\alpha_i) \end{bmatrix}, \text{ and so } \boldsymbol{w}(\alpha_i) = \begin{bmatrix} -C_{1,1}^{-1} \cdot C_{1,2}\, \boldsymbol{v}(\alpha_i) \\ \boldsymbol{v}(\alpha_i) \end{bmatrix}.$$

Each coordinate of $\boldsymbol{w}(\alpha_i)$ is a monomial in $\mathbb{K}[\boldsymbol{X}]_1 \otimes \mathbb{K}[\boldsymbol{Y}]_1 \otimes \mathbb{K}[\boldsymbol{Z}]_1$ evaluated at α_i. Hence, we can recover the coordinates of α_i from $\boldsymbol{w_i}(\alpha_i)$ by inverting a monomial map. In this case, the solutions to $(f_1, \ldots, f_4)$, and so MEP-eigenpairs, are

	MEP-Eigenvalues	MEP-Eigenvectors		
	x_0, x_1, x_2	y_0, y_1	z_0, z_1	
$\alpha_1 = ($	$1, -1, -3$	$1, 1$	$1, -3$	$)$
$\alpha_2 = ($	$1, 3, 4$	$1, 1$	$1, -1$	$)$
$\alpha_3 = ($	$1, 1, 1$	$1, -1$	$1, -2$	$)$
$\alpha_4 = ($	$1, 1, 2$	$1, -3$	$1, -3$	$)$

(6.37)

Chapter 7

Gröbner basis and sparse polynomial systems

Computing Gröbner bases is an intrinsically hard problem. For many "interesting" cases related to applications, the complexity of the algorithms to compute them is single exponential in the number of variables, but there are instances where the complexity is double exponential; it is an EXPSPACE complete problem [May97], see also [GG13, Sec. 21.7]. There are many practically efficient algorithms, see [CLO15, Ch. 10] and references therein, for which, under genericity assumptions, we can deduce precise complexity estimates, see Sec. 4.5. However, the polynomial systems coming from applications have some kind of structure. One of the main challenges in Gröbner basis theory is to improve the complexity and the practical performance of the related algorithms by exploiting the structure. In this chapter we focus on computing Gröbner bases for *mixed sparse polynomial systems*, that is, sparse polynomials systems defined by polynomials with different Newton polytope.

An approach to exploit the sparsity of the polynomials is to compute Gröbner bases over semigroup algebras [Stu93, FSS14]. Semigroup algebras are related to toric varieties. An affine toric variety is the spectrum of a semigroup algebra [CLS11, Thm. 1.1.17]. Hence, the variety defined by the polynomials over a semigroup algebra is a subvariety of a toric variety. This variety is different from the one defined by the polynomials over the original polynomial algebra, but they are related and in many applications the difference is irrelevant; see, for example, [EM99a]. We refer to [CLS11] for an introduction to toric varieties and to [Stu96] for their relation with Gröbner basis.

Following [Stu93], Faugère et al. considered *sparse unmixed systems* and introduced an algorithm to compute Gröbner bases over the semigroup algebra generated by their Newton polytope [FSS14]. By embedding the systems in a semigroup algebra, they can predict the structure of regular sparse unmixed systems and exploit it algorithmically. Their algorithm is a variant of the `Matrix-F5` algorithm [Fau02, BFS15]. They homogenize the polynomials and compute a Gröbner basis, degree-by-degree, by performing Gaussian elimination on various Macaulay matrices [Laz83]. They use the `F5` criterion [Fau02] to avoid redundant computations, that is, to skip rows reducing to zero after performing Gaussian elimination. Once they have computed the Gröbner basis for the homogenized system, they recover a Gröbner basis of the original system by dehomogenizing the computed basis. The efficiency of this approach relies on an incremental degree-by-degree construction which, under regularity assumptions, skips all the rows reducing to zero. One of the properties that they exploit in this work is that, for normal

Newton polytopes [CLS11, Def. 2.2.9], the homogenization of a *generic* unmixed system forms a regular sequence over the corresponding semigroup algebra [CLS11, Def. 9.2.9]. Unfortunately, this property is no longer true for *mixed systems*. So, for mixed systems, this algorithm fails to predict all rows reducing to zero during Gaussian elimination.

In Chapter 7, we study extensions of the algorithms to compute Gröbner bases for unmixed sparse systems [FSS14] to the mixed case. We present algorithms which, under regularity assumptions, perform no reductions to zero.

- Our first extension (Section 7.3) changes the *degree-by-degree* strategy (Sec. 4.4) of the algorithm proposed in [FSS14] by a *polynomial-by-polynomial* strategy, that is, we first compute a Gröbner basis for the ideal generated by the first i polynomials and then we extend this basis to a Gröbner basis for the ideal generated by first $(i+1)$ polynomials. In the language of signature-based algorithms, we change the module monomial order of `F5` from $<_{\text{d-pot}}$ (degree-by-degree) to $<_{\text{pot}}$ (polynomial-by-polynomial), see [EF17]. This strategy avoids every reduction to zero when the original (non-homogenized) system is a regular sequence. Unfortunately, this approach requires to compute using a $\mathrm{GRevLex}$ monomial orderings which, as we show in Ex. 7.2.1, we can not define for semigroup algebras. Hence, we introduce a novel Gröbner-like basis, that we call *scant Gröbner basis*. A *scant Gröbner basis* is a basis for an ideal over a semigroup algebra with similar properties to the usual Gröbner basis. Their main advantage is that they allow us to define $\mathrm{GRevLex}$-like orderings over semigroup algebras with many of the expected properties, see Sec. 4.3. To define *scant Gröbner basis*, we need to work with non-monomial-orderings and redefine the monomial division relation. Hence, a *scant Gröbner basis* is not Gröbner basis over a semigroup algebra. We introduce an algorithm to compute *scant Gröbner basis* for *unmixed sparse systems* which, under regularity assumptions, performs no reductions to zero.

- Our second extension (Section 4.6) computes a Gröbner basis over a multigraded semigroup algebra; the multigrading is related to the different polytopes of the sparse polynomials. Even though we embed the system in the multigraded semigroup algebra, the straightforward homogenization of the input polynomials never results in a regular sequence. Therefore, the existing criteria do not avoid all the trivial (expected) reductions to zero. Hence, to avoid all the trivial reductions to zero, we extend the classical `F5` criterion (Prop. 4.4.6) by using the exactness of the strands of the *Koszul complex*. We introduce the concept of *(sparse) regularity*, related to the exactness of these strands, to guarantee that all the reductions to zero are trivial. We present the first algorithm that computes Gröbner bases over these multigraded semigroup algebras which, under (sparse) regularity assumptions, performs no reductions to zero.

We emphasize that besides the similarity in their name, scant Gröbner bases and Gröbner bases over semigroup algebras are completely different objects. In particular, scant Gröbner bases *are not* Gröbner basis over semigroup algebras. Moreover, the corresponding algorithms follow different computational strategies: to compute scant Gröbner bases, we proceed polynomial-by-polynomial, while to compute Gröbner bases over semigroup algebras, degree-by-degree. The assumptions for the algorithms to perform no reductions to zero are also different, but both of them are satisfied under (sparse) regularity assumptions. Last but not least, we can use both objects to solve sparse systems, but we do not have complexity bounds when we use scant Gröbner bases.

7.1 Semigroup algebras and regularity

We introduce the notation that we use in the chapter. In addition, we use the definitions introduced in Sections 2.7 and 2.12, that is the pointed affine semigroups (Definitions 2.12.1 and 2.12.2), the semigroup algebras (Def. 2.12.7), the pointed rational polyhedral cones (Def. 2.12.4), the integer polytopes (Def. 2.12.14), the Minkowski sum (Def. 2.12.16), the Laurent polynomials (Ex. 2.12.9), the Newton polytope (Def. 2.12.17), the regular sequences (Def. 2.7.1), and the Koszul complex (Def. 2.7.5).

Let $\mathbb{K} \subset \mathbb{C}$ be a field of characteristic zero, $\boldsymbol{x} := (x_1, \ldots, x_n)$, and $\mathbb{K}[\boldsymbol{x}] := \mathbb{K}[x_1, \ldots, x_n]$. We consider $\mathbf{0} := (0, \ldots, 0)$ and $\mathbf{1} := (1, \ldots, 1)$. For each $r \in \mathbb{N}$, let $\boldsymbol{e_1}, \ldots, \boldsymbol{e_r}$ be the canonical basis of $\mathbb{R}^r$. Given $\boldsymbol{d_1}, \boldsymbol{d_2} \in \mathbb{N}^r$, we say $\boldsymbol{d_1} \geq \boldsymbol{d_2}$ when $\boldsymbol{d_1} - \boldsymbol{d_2} \in \mathbb{N}^r$. We denote $[r] = \{1, \ldots, r\}$. Let $\langle f_1, \ldots, f_m \rangle$ be the ideal generated by $f_1, \ldots, f_m$.

Semigroup algebras

Given a sparse polynomial system in $\mathbb{K}[\boldsymbol{x}]$, instead of consider it over $\mathbb{K}[\boldsymbol{x}]$ or $\mathbb{K}[\mathbb{Z}^n]$, see Rem. 2.12.10, we can embed it in a subalgebra related to the Newton polytopes of the polynomials. In this way, we exploit the sparsity of the (polynomials of the) system.

Definition 7.1.1 (Semigroup algebra of polytopes)**.** *We consider r integer polytopes $\Delta_1, \ldots, \Delta_r \subset \mathbb{R}^n$ such that their Minkowski sum, $\Delta := \sum_{i=1}^r \Delta_i$, has dimension n and $\mathbf{0}$ is its vertex; in particular, $\mathbf{0}$ is a vertex of every Newton polytope Δ_i. We also consider the polytope $\boldsymbol{\Delta} := \sum(\Delta_i \times \{\boldsymbol{e_i}\})$, which is the Cayley embedding of $\Delta_1, \ldots, \Delta_r$.*

In what follows, we consider the pointed rational polyhedral cones $\mathcal{C}_\Delta$ and $\mathcal{C}_{\boldsymbol{\Delta}}$, see Def. 2.12.4, and we work with the semigroup algebras $\mathbb{K}[S_\Delta] := \mathbb{K}[\mathcal{C}_\Delta \cap \mathbb{Z}^n]$ and $\mathbb{K}[S^h_{\boldsymbol{\Delta}}] := \mathbb{K}[\mathcal{C}_{\boldsymbol{\Delta}} \cap \mathbb{Z}^{n+r}]$. We will write the monomials in $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ as $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d})}$, where $\boldsymbol{\alpha} \in (\mathcal{C}_\Delta \cap \mathbb{Z}^n)$ and $\boldsymbol{d} \in \mathbb{N}^r$.

The algebra $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ is $\mathbb{N}^r$-multigraded as follows: for every $\boldsymbol{d} = (d_1, \ldots, d_r) \in \mathbb{N}^r$, $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$ is the $\mathbb{K}$-vector space spanned by the monomials $\{\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d})} : \boldsymbol{\alpha} \in (\sum d_i \cdot \Delta_i) \cap \mathbb{Z}^n\}$. Then, $F \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$ is homogeneous and has *multidegree* $\boldsymbol{d}$, which we denote by $\deg(F)$.

We can think $\mathbb{K}[S_\Delta]$ as the "dehomogenization" of $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$.

Definition 7.1.2 (Dehomogenization morphism)**.** *The dehomogenization morphism from $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ to $\mathbb{K}[S_\Delta]$ is the surjective ring homomorphism $\chi : \mathbb{K}[S^h_{\boldsymbol{\Delta}}] \to \mathbb{K}[S_\Delta]$ that maps the monomials $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d})} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ to $\chi(\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d})}) := \boldsymbol{X}^{\boldsymbol{\alpha}} \in \mathbb{K}[S_\Delta]$.*

If $\mathfrak{b}$ is a set of homogeneous polynomials in $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$, then we consider $\chi(\mathfrak{b}) = \{\chi(G) : G \in \mathfrak{b}\}$.

Observation 7.1.3. *As $\mathbf{0}$ is a vertex of Δ, there is a monomial $\boldsymbol{X}^{(\mathbf{0},\boldsymbol{e_i})} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$, for every $i \in [r]$. Hence, given a finite set of monomials $\boldsymbol{X}^{\boldsymbol{\alpha}_1}, \ldots, \boldsymbol{X}^{\boldsymbol{\alpha}_k} \in \mathbb{K}[S_\Delta]$, we can find a multidegree $\boldsymbol{d} \in \mathbb{N}^r$ such that $\boldsymbol{X}^{(\boldsymbol{\alpha}_1,\boldsymbol{d})}, \ldots, \boldsymbol{X}^{(\boldsymbol{\alpha}_k,\boldsymbol{d})} \in \mathbb{K}[S_\Delta]_{\boldsymbol{d}}$.*

Given a system of polynomials $f_1, \ldots, f_m \in \mathbb{K}[S_\Delta]$, we can find a multidegree $\boldsymbol{d} \in \mathbb{N}^r$ and homogeneous polynomials $F_1, \ldots, F_m \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$ so that it holds $\chi(F_i) = f_i$, for every $i \in [m]$.

Moreover, given homogeneous polynomials $F_1, \ldots, F_m \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ and an affine polynomial $g \in \langle \chi(F_1), \ldots, \chi(F_m) \rangle$, there is a homogeneous polynomial $G \in \langle F_1, \ldots, F_m \rangle$ such that $\chi(G) = g$.

Observation 7.1.4. *If we fix a multidegree $\boldsymbol{d} \in \mathbb{N}^r$, then the map χ restricted to $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$ is injective.*

Observation 7.1.5. *The dehomogenization of a homogeneous ideal $I \subset \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$, $\chi(I) := \langle\{\chi(f) : f \in I\}\rangle$, is an ideal in $\mathbb{K}[S_\Delta]$.*

Regularity

We use the Koszul complex (Def. 2.7.5) to define our notion of *regularity*. In what follows, we reintroduce only some notation about this complex and we refer the reader to Sec. 2.7 for more details.

Definition 7.1.6 (Koszul complex)**.** *For a sequence of homogeneous $F_1, \ldots, F_k \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ of multidegrees $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_k$ and a multidegree $\boldsymbol{d} \in \mathbb{N}^r$, we denote by $\mathcal{K}(F_1, \ldots, F_k)_{\boldsymbol{d}}$ the strand of the Koszul complex of $F_1, \ldots, F_k$ of multidegree $\boldsymbol{d}$, that is,*

$$\mathcal{K}(F_1, \ldots, F_k)_{\boldsymbol{d}} : 0 \to (\mathcal{K}_k)_{\boldsymbol{d}} \xrightarrow{\delta_k} \ldots \xrightarrow{\delta_1} (\mathcal{K}_0)_{\boldsymbol{d}} \to 0,$$

where, for $1 \leq t \leq k$, we have

$$(\mathcal{K}_t)_{\boldsymbol{d}} := \bigoplus_{\substack{I \subset \{1,\ldots,k\} \\ \#I = t}} \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{(\boldsymbol{d} - \sum_{i \in I} \boldsymbol{d}_i)} \otimes (e_{I_1} \wedge \cdots \wedge e_{I_t}).$$

We denote by $\mathcal{H}_t(F_1, \ldots, F_k)_{\boldsymbol{d}}$ the t-th Koszul homology of $\mathcal{K}(F_1, \ldots, F_k)_{\boldsymbol{d}}$, that is $\mathcal{H}_t(F_1, \ldots, F_k)_{\boldsymbol{d}} := (\mathtt{Ker}(\delta_t)/\mathtt{Im}(\delta_{t+1}))_{\boldsymbol{d}}$.

The 0-th Koszul homology is $\mathcal{H}_0(F_1, \ldots, F_k) \cong (\mathbb{K}[S^h_{\boldsymbol{\Delta}}]/\langle F_1, \ldots, F_k\rangle)$.

Definition 7.1.7 (Koszul and sparse regularity)**.** *A sequence of homogeneous polynomials $F_1, \ldots, F_k \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ is* Koszul regular *if for every $\boldsymbol{d} \in \mathbb{N}^r$ such that $\boldsymbol{d} \geq \sum_{i=1}^k \boldsymbol{d}_i$ and for every $t > 0$, the t-th Koszul homology vanishes at degree $\boldsymbol{d}$, that is $\mathcal{H}_t(F_1, \ldots, F_k)_{\boldsymbol{d}} = 0$. We say that the sequence is* (sparse) regular *if $F_1, \ldots, F_j$ is Koszul regular, for every $j \leq k$.*

The BBK bound (Thm. 2.12.19) is related to *Koszul regularity*. For example, Kushnirenko used this notion of regularity in his proof of the BBK bound [Kus76, Thm. 2].

Observation 7.1.8. *Note that the Koszul regularity does not depend on the order of the polynomials, as (sparse) regularity does.*

Sparse regularity is related to regular sequences over $\mathbb{K}[S_\Delta]$.

Lemma 7.1.9. *Consider a* (sparse) regular sequence *of homogeneous polynomials $F_1, \ldots, F_k \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ of degrees $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_k$, respectively. For each F_i, consider its dehomogenization $f_i = \chi(F_i) \in \mathbb{K}[S_\Delta]$. Then, the sequence $(f_1, \ldots, f_k)$ is a* regular sequence *in $\mathbb{K}[S_\Delta]$.*

Proof. We prove this lemma by induction on $k \in \mathbb{N}$. If $k = 1$, then the lemma is trivial as $\mathbb{K}[S_\Delta]$ is a domain (Prop. 2.12.11). If $k > 1$, then, by definition of (sparse) regular sequence, $(F_1, \ldots, F_{k-1})$ is sparse regular, and so $(f_1, \ldots, f_{k-1})$ is a regular sequence over $\mathbb{K}[S_\Delta]$. We only need to prove that f_k is a regular element in $\mathbb{K}[S_\Delta]/(f_1, \ldots, f_{k-1})$. Consider $g_k \in \mathbb{K}[S_\Delta]$ such that $g_k f_k = 0$ in $\mathbb{K}[S_\Delta]/(f_1, \ldots, f_{k-1})$. Using Obs. 7.1.3, we can prove that there are homogeneous polynomials $G_1, \ldots, G_{k-1}, G_k \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$, for some $\boldsymbol{d} \in \mathbb{N}^r$ such that $\sum_{i=1}^{k-1} G_i F_i = G_k F_k$, $\chi(G_k) = g_k$, and $\boldsymbol{d} \geq \sum_{i=1}^k \boldsymbol{d}_i$. Hence $(G_1, \ldots, G_{k-1}, -G_k)$ belongs to the kernel of the map δ_1 from the Koszul complex. As $\mathcal{H}(F_1, \ldots, F_k)_{\boldsymbol{d}} = 0$, then $G_k \in \langle F_1, \ldots, F_{k-1}\rangle$, and so $g_k = \chi(G_k) \in \chi(\langle F_1, \ldots, F_{k-1}\rangle) = \langle f_1, \ldots, f_{k-1}\rangle$. Therefore, f_k is a regular element in $\mathbb{K}[S_\Delta]/(f_1, \ldots, f_{k-1})$ and so $(f_1, \ldots, f_k)$ is a regular sequence on $\mathbb{K}[S_\Delta]$. □

7.2 Scant Gröbner basis

Our objective is to extend the ideas in [FSS14] to the context of sparse mixed systems. Our approach is to change their *degree-by-degree* strategy to a *polynomial-by-polynomial* strategy. While the optimality of the F5 criterion for a *degree-by-degree* strategy relies on the regularity of the homogenization of the input system (over $\mathbb{K}[S^h_{\mathbf{\Delta}}]$), the optimality of a *polynomial-by-polynomial* strategy relies on the regularity of the input system (over $\mathbb{K}[S_\Delta]$) and so it is more general, see Lem. 7.1.9. To consider this strategy, we need to work with $\mathrm{GRevLex}$-like orderings. As we show in Ex. 7.2.1, in general, we cannot define a monomial ordering for $\mathbb{K}[S_\Delta]$ that behaves like $\mathrm{GRevLex}$. Thus, we are forced to work with sparse orders which are not monomial orders. In this section, we introduce *scant Gröbner bases*. These bases extend the notion of Gröbner bases to sparse orders. We prove that working with sparse orders is not as hard as it looks like: (reduced) *scant Gröbner bases* are always finite and they are related to Gröbner bases over the standard polynomial algebra. Moreover, we introduce an algorithm to compute them. We prove that, if $f_1, \dots, f_k \in \mathbb{K}[S_\Delta]$ form a regular sequence, our algorithm performs no reductions to zero. Our regularity assumptions are more general than the ones in [FSS14] and they are compatible with *mixed sparse systems*, see Lem. 7.1.9.

7.2.1 Preliminaries

Recall the definitions from Sec. 7.1. In what follows, consider the blocks of variables $\boldsymbol{y} := (y_0, \dots, y_m)$, and the standard polynomial algebra $\mathbb{K}[\boldsymbol{y}] := \mathbb{K}[y_0, \dots, y_m]$. For each $\alpha \in \mathbb{N}^{m+1}$, consider the monomial $\boldsymbol{y}^\alpha := \prod_{i=0}^m y_i^{\alpha_i}$. Let $\{a_0, a_1, \dots, a_m\}$ be a set of generators of $S \subset \mathbb{Z}^n$. Let $e_0 \dots e_m$ be the canonical basis of $\mathbb{Z}^{m+1}$. Consider the homomorphism $\rho : \mathbb{Z}^{m+1} \to S$ such that, for each $0 \leq i \leq m$, $\rho(e_i) = a_i$). By Prop. 2.12.13, the semigroup algebra $\mathbb{K}[S]$ (Def. 2.12.7) is isomorphic to the quotient ring $\mathbb{K}[\boldsymbol{y}]/T$, where T is the lattice ideal

$$T := \langle \boldsymbol{y}^\alpha - \boldsymbol{y}^\beta | \alpha, \beta \in \mathbb{N}^{m+1}, \rho(\alpha - \beta) = 0 \rangle. \tag{7.1}$$

Moreover, the ideal T is prime and $\mathbb{K}[S]$ is an integral domain (Prop. 2.12.11). As in [FSS16], in this section we only consider pointed affine semigroups, see Definitions 2.12.1 and 2.12.2.

Given a polytope $M \subset \mathbb{R}^n$, in this section we consider the semigroup algebras $\mathbb{K}[S_M]$ and $\mathbb{K}[S^h_M]$ corresponding to the semigroup algebras in Def. 7.1.1. Note that we use the letter M to denote the semigroup instead of Δ. The algebra $\mathbb{K}[S^h_M]$ is $\mathbb{Z}$-graded and, for each $d \in \mathbb{Z}$, $\mathbb{K}[S^h_M]$ is a $\mathbb{K}$-vector space generated by the monomials $\{\boldsymbol{X}^{(\boldsymbol{\alpha},d)} : \boldsymbol{\alpha} \in (d \cdot M \cap \mathbb{Z}^n)\}$. Given $f \in \mathbb{K}[S^h_M]_d$, we define the *degree* of f as $\deg(f) := d \in \mathbb{N}$. Given an homogeneous ideal $I \subset \mathbb{K}[S^h_M]$, we denote by $\chi(I) \subset \mathbb{K}[S_M]$ the ideal obtained by dehomogenizing I, see Obs. 7.1.5.

Sparse degree and homogenization

We define the *affine degree* of $\boldsymbol{X}^s \in \mathbb{K}[S_M]$, $\delta^A(\boldsymbol{X}^s)$, as the smallest $d \in \mathbb{N}$ such that $\boldsymbol{X}^{(s,d)} \in \mathbb{K}[S^h_M]$. We extend this definition to the affine polynomials in $\mathbb{K}[S_M]$ as the maximal affine degree of each monomial. That is, for $f := \sum_{s \in S_M} c_s \boldsymbol{X}^s \in \mathbb{K}[S_M]$, the affine degree of f is

$$\delta^A(f) := \max_{s \in S_M} (\delta^A(\boldsymbol{X}^s) : c_s \neq 0).$$

Let $\mathbb{K}[S_M]_{\leq d}$ be the set of all polynomials in $\mathbb{K}[S_M]$ of affine degree at most d. The map $\chi^{-1} : \mathbb{K}[S_M] \to \mathbb{K}[S_M^h]$ defines the homogenization of $f := \sum_{s\in S_M} c_s \boldsymbol{X}^s \in \mathbb{K}[S_M]$, where $\chi^{-1}(f) := \sum_{s\in S_M} c_s \boldsymbol{X}^{(s,\delta^A(f))} \in \mathbb{K}[S_M^h]$. Note that this map is not a homomorphism. For an ideal $I \subset \mathbb{K}[S_M]$, $\chi^{-1}(I)$ is the homogeneous ideal,

$$\chi^{-1}(I) := \langle \{\chi^{-1}(f) : f \in I\}\rangle.$$

Finally, given a polynomial $f \in \mathbb{K}[S_M^h]$ we define its *sparse degree* as $\delta(f) := \delta^A(\chi(f))$. Note that, the degree is always bigger or equal to the sparse degree. Even though we use the name sparse degree, it does not give a graded structure to the $\mathbb{K}$-algebra $\mathbb{K}[S_M^h]$.

Orders for Monomials

A monomial order $<$ for $\mathbb{K}[S]$ is a well-order compatible with the multiplication on $\mathbb{K}[S]$, that is $\forall s \in S, s \neq 0 \implies \boldsymbol{X}^0 < \boldsymbol{X}^s$ and $\forall s, r, t \in S, \boldsymbol{X}^s < \boldsymbol{X}^r \implies \boldsymbol{X}^{s+t} < \boldsymbol{X}^{r+t}$. These orders exist on $\mathbb{K}[S]$ if and only if S is pointed, see Sec. 4.6.

Given any well-order $<$ for $\mathbb{K}[S_M]$, we can extend it to a well-order $<_h$ for the monomials in $\mathbb{K}[S_M^h]$ as follows:

$$\boldsymbol{X}^{(s,d)} < \boldsymbol{X}^{(r,d')} \iff \begin{cases} d < d' \\ d = d' \wedge \boldsymbol{X}^s < \boldsymbol{X}^r \end{cases} \tag{7.2}$$

We refer to the order $<_h$ as *the grading of* $<$. If $<$ is a monomial order, then $<_h$ is a monomial order too.

Given an ideal $I \subset \mathbb{K}[S_M]$, it is useful to study the vector space $I \cap \mathbb{K}[S_M]_{\leq d}$, that is the elements of I of degree smaller or equal to d. This information allow us, for example, to compute the Hilbert Series of the affine ideal. It is also important for computational reasons. For example, to maintain the invariants in the affine signature-based Gröbner basis algorithms, as the F5 algorithm [EF17].

In our setting, to compute a basis of $I \cap \mathbb{K}[S_M]_{\leq d}$, we have to work with an order for the monomials in $\mathbb{K}[S_M]$ that takes into account the sparse degree. This order, $\prec$, is such that for any $\boldsymbol{X}^s, \boldsymbol{X}^r \in \mathbb{K}[S_M]$, $\delta^A(\boldsymbol{X}^s) < \delta^A(\boldsymbol{X}^r) \implies \boldsymbol{X}^s \prec \boldsymbol{X}^r$. Unfortunately, for most of the semigroup algebras $\mathbb{K}[S_M]$, *there is no monomial order* with this property. Therefore, we are forced to work with well-orders that are not monomial orders.

Example 7.2.1. *Consider the set of integer points* $\{[0,0],[1,0],[0,1],[1,1]\} \subset \mathbb{N}^2$ *and let* M *be its convex hull (Def. 2.12.3). That is,*

$$M := \textit{ConvexHull}(\{[0,0],[1,0],[0,1],[1,1]\}) \subset \mathbb{R}^2.$$

Consider a monomial order $<$ *for* $\mathbb{K}[S_M]$*. Without loss of generality, assume* $\boldsymbol{X}^{[1,0]} < \boldsymbol{X}^{[0,1]}$*. Then,* $\boldsymbol{X}^{[2,0]} < \boldsymbol{X}^{[1,1]} < \boldsymbol{X}^{[0,2]}$*. But,* $\delta^A(\boldsymbol{X}^{[2,0]}) = 2$ *and* $\delta^A(\boldsymbol{X}^{[1,1]}) = 1$*. So, no monomial order on* $\mathbb{K}[S_M]$ *takes into account the sparse degree.*

Given a monomial order $<_M$ for $\mathbb{K}[S_M]$, we define the *sparse order* $\prec$ for $\mathbb{K}[S_M]$ as follows.

$$\boldsymbol{X}^s \prec \boldsymbol{X}^r \iff \begin{cases} \delta^A(\boldsymbol{X}^s) < \delta^A(\boldsymbol{X}^r) \\ \delta^A(\boldsymbol{X}^s) = \delta^A(\boldsymbol{X}^r) \wedge \boldsymbol{X}^s <_M \boldsymbol{X}^r \end{cases} \tag{7.3}$$

Let $\prec_h$ be the grading of the sparse order of $\mathbb{K}[S_M^h]$ (Eq. 7.2). We call this order the *graded sparse order*.

Remark 7.2.2. *By definition, these two orders are the same for monomials of the same degree. That is,*

$$\forall \boldsymbol{X}^{(s,d)}, \boldsymbol{X}^{(r,d)} \in \mathbb{K}[S_M^h],\ \boldsymbol{X}^{(s,d)} \prec_h \boldsymbol{X}^{(r,d)} \iff \boldsymbol{X}^s \prec \boldsymbol{X}^r \ .$$

Usually, this order is not compatible with the multiplication, see Ex. 7.2.1. But,

Lemma 7.2.3. *If $\boldsymbol{X}^s \prec \boldsymbol{X}^t$ and $\delta^A(\boldsymbol{X}^r) + \delta^A(\boldsymbol{X}^t) = \delta^A(\boldsymbol{X}^t \cdot \boldsymbol{X}^r)$, then $\boldsymbol{X}^s \cdot \boldsymbol{X}^r \prec \boldsymbol{X}^t \cdot \boldsymbol{X}^r$.*

Proof. Note that δ^A satisfies the triangular inequality, $\delta^A(\boldsymbol{X}^{s+r}) \leq \delta^A(\boldsymbol{X}^s) + \delta^A(\boldsymbol{X}^r)$. As $\boldsymbol{X}^s \prec \boldsymbol{X}^t$, it holds $\delta^A(\boldsymbol{X}^s) \leq \delta^A(\boldsymbol{X}^t)$. By assumption, $\delta^A(\boldsymbol{X}^t) + \delta^A(\boldsymbol{X}^r) = \delta^A(\boldsymbol{X}^{t+r})$. Then,

$$\delta^A(\boldsymbol{X}^{s+r}) \leq \delta^A(\boldsymbol{X}^s) + \delta^A(\boldsymbol{X}^r) \leq \delta^A(\boldsymbol{X}^t) + \delta^A(\boldsymbol{X}^r) \leq \delta^A(\boldsymbol{X}^{t+r}).$$

Hence, either $\delta^A(\boldsymbol{X}^{s+r}) < \delta^A(\boldsymbol{X}^{t+r})$ or the sparse degree is the same. In the second case, we conclude $\delta^A(\boldsymbol{X}^s) = \delta^A(\boldsymbol{X}^t)$, and so, as $\boldsymbol{X}^s \prec \boldsymbol{X}^t$, it holds $\boldsymbol{X}^s <_M \boldsymbol{X}^t$. As $<_M$ is a monomial order, $\boldsymbol{X}^{s+r} <_M \boldsymbol{X}^{t+r}$. Hence, in both cases, $\boldsymbol{X}^s \cdot \boldsymbol{X}^r \prec \boldsymbol{X}^t \cdot \boldsymbol{X}^r$. □

We extend this property to the homogeneous case.

Corollary 7.2.4. *If $\boldsymbol{X}^{(s,d_s)} \prec \boldsymbol{X}^{(t,d_t)}$ and $\delta(\boldsymbol{X}^{(r,d_r)}) + \delta(\boldsymbol{X}^{(t,d_t)}) = \delta(\boldsymbol{X}^{(r,d_r)} \cdot \boldsymbol{X}^{(t,d_t)})$, then*

$$\boldsymbol{X}^{(s,d_s)} \cdot \boldsymbol{X}^{(r,d_r)} \prec \boldsymbol{X}^{(t,d_t)} \cdot \boldsymbol{X}^{(r,d_r)}.$$

7.2.2 Definition of scant Gröbner basis and properties

We want to define and compute Gröbner-like bases in $\mathbb{K}[S_M]$ and $\mathbb{K}[S_M^h]$ with respect to a (graded) sparse order. As these orders are not compatible with the multiplication, not all the standard definitions of Gröbner basis are equivalent. For example, the ideal generated by set of leading monomials of an ideal in $\mathbb{K}[S_M]$ might contain monomials which are not leading monomials of a polynomial in the original ideal. We say that a set of generators G of an ideal $I \subset \mathbb{K}[S_M]$ is a scant Gröbner basis with respect to an order $\prec$, if for each $f \in I$, there is a $g \in G$ such that $\mathtt{LM}_\prec(g)$ divides $\mathtt{LM}_\prec(f)$. Similarly for $\mathbb{K}[S_M^h]$.

This definition has a drawback: The multivariate polynomial division algorithm might not terminate. This can happen when $\mathtt{LM}_\prec(f) = \boldsymbol{X}^t \cdot \mathtt{LM}_\prec(g) \prec \mathtt{LM}_\prec(\boldsymbol{X}^t \cdot g)$. Then, the reduction step "increases" the leading monomial, so that the algorithm does not necessarily terminate. We can construct examples where we have a periodic sequence of reductions. To avoid this problem, we redefine the division relation.

Definition 7.2.5 (Division relation)**.** *For any $\boldsymbol{X}^{(s,d_s)}, \boldsymbol{X}^{(r,d_r)} \in \mathbb{K}[S_M^h]$, we say that $\boldsymbol{X}^{(s,d_s)}$ divides $\boldsymbol{X}^{(r,d_r)}$, and write $\boldsymbol{X}^{(s,d_s)} || \boldsymbol{X}^{(r,d_r)}$, if there is a $\boldsymbol{X}^{(t,d_t)} \in \mathbb{K}[S_M^h]$ such that*

$$\boldsymbol{X}^{(s,d_s)} \cdot \boldsymbol{X}^{(t,d_t)} = \boldsymbol{X}^{(r,d_r)} \text{ and } \delta(\boldsymbol{X}^{(s,d_s)}) + \delta(\boldsymbol{X}^{(t,d_t)}) = \delta(\boldsymbol{X}^{(r,d_r)}).$$

Similarly, for $\boldsymbol{X}^s, \boldsymbol{X}^r \in \mathbb{K}[S_M]$, we say that $\boldsymbol{X}^s$ divides $\boldsymbol{X}^r$, and write $\boldsymbol{X}^s || \boldsymbol{X}^r$, if $\chi^{-1}(\boldsymbol{X}^s) || \chi^{-1}(\boldsymbol{X}^r)$.

Remark 7.2.6. *If $\mathtt{LM}_{\prec_h}(f) || \boldsymbol{X}^{(s,d_s)}$, then there is a $\boldsymbol{X}^{(t,d_t)} \in \mathbb{K}[S_M^h]$ such that $\boldsymbol{X}^{(s,d_s)} = \boldsymbol{X}^{(t,d_t)} \cdot \mathtt{LM}_{\prec_h}(f) = \mathtt{LM}_{\prec_h}(\boldsymbol{X}^{(t,d_t)} \cdot f)$, by Lem. 7.2.3. Similarly over $\mathbb{K}[S_M]$.*

We define the scant Gröbner bases as follows.

Definition 7.2.7 (Scant Gröbner bases)**.** *Given a (graded) sparse order $\prec$, see Eq.* (7.3)*, and an ideal $I \subset \mathbb{K}[S_M]$, respectively $I \subset \mathbb{K}[S_M^h]$, a set $G \subset I$ is a scant Gröbner basis if it generates I and for any $f \in I$ there is some $g \in G$ such that $\mathtt{LM}_\prec(g) || \mathtt{LM}_\prec(f)$.*

With this definition, each step in the division algorithm reduces the leading monomial (Rem. 7.2.6), and so the division algorithm always terminates; see, for example, [CLO15, Thm. 2.3.3 & Prop. 2.6.1].

Lemma 7.2.8. *Let $f \in \mathbb{K}[S_M]$ and G be a list of polynomials in $\mathbb{K}[S_M]$. Using our definition of division relation (Def. 7.2.5), the multivariate division algorithm (Alg. 1) for the division of f by G, with respect to the order $\prec$, terminates. Moreover, if G is a scant Gröbner basis of an ideal I with respect to $\prec$ and $f \equiv f' \mod I$, then the remainder of the division algorithm for f and f' is the same. This remainder is unique and does not depend on the scant Gröbner basis of I with respect to $\prec$ that we choose.*

Our next goal is to prove that for every ideal and sparse order, there is a finite scant Gröbner basis. A priori, it is not clear how to extend Dickson's lemma [CLO15, Thm. 2.4.5] to this setting; our division relation (Def. 7.2.7) is not compatible with the recursive proof of the lemma. Our strategy is to prove first that, over $\mathbb{K}[S_M^h]$, there is always a finite scant Gröbner basis. Then, we extend this result to $\mathbb{K}[S_M]$. We show that this scant Gröbner basis is related to a Gröbner basis over a standard polynomial algebra, so it is finite.

Finiteness of scant Gröbner Bases

Homogeneous case. Let $<_M$ be a monomial order for $\mathbb{K}[S_M]$ and $\prec$ the sparse order related to $<_M$, Eq. (7.3). Consider $\prec_h$ the graded sparse order related to $\prec$ over $\mathbb{K}[S_M^h]$, Eq. (7.2).

Consider the lattice ideal T from Eq. (7.1). This ideal T is homogeneous and the algebra $\mathbb{K}[S_M^h]$ is isomorphic to $\mathbb{K}[\boldsymbol{y}]/T$ as a graded algebra. Let $\widetilde{\psi} : \mathbb{K}[\boldsymbol{y}]/T \to \mathbb{K}[S_M^h]$ and $\widetilde{\phi} : \mathbb{K}[S_M^h] \to \mathbb{K}[\boldsymbol{y}]/T$ be the isomorphisms related to $\mathbb{K}[S_M^h] \cong \mathbb{K}[\boldsymbol{y}]/T$, such that they are inverse of each other and $\widetilde{\psi}(\boldsymbol{X}^{(0,1)}) = y_0$. We extend $\widetilde{\psi}$ to $\psi : \mathbb{K}[\boldsymbol{y}] \to \mathbb{K}[S_M^h]$, where $\psi(\boldsymbol{y}^\alpha)$ is the image, under $\widetilde{\psi}$, of $\boldsymbol{y}^\alpha$ modulo T. The map ψ is a 0-graded epimorphism.

For $\boldsymbol{y}^\alpha \in \mathbb{K}[\boldsymbol{y}]$, let $\deg(\boldsymbol{y}^\alpha, y_0)$ be the degree of $\boldsymbol{y}^\alpha$ with respect to y_0 and $\deg(\boldsymbol{y}^\alpha)$ be the total degree. Given a (standard) monomial order $\widetilde{<}$ for $\mathbb{K}[\boldsymbol{y}]$, consider the graded monomial order $<_y$ for $\mathbb{K}[\boldsymbol{y}]$ defined as follows,

$$\boldsymbol{y}^a <_y \boldsymbol{y}^b \iff \begin{cases} \deg(\boldsymbol{y}^a) < \deg(\boldsymbol{y}^b) & \\ \deg(\boldsymbol{y}^a) = \deg(\boldsymbol{y}^b) \quad \wedge \quad \deg(\boldsymbol{y}^a, y_0) > \deg(\boldsymbol{y}^b, y_0) & \\ \deg(\boldsymbol{y}^a) = \deg(\boldsymbol{y}^b) \quad \wedge \quad \deg(\boldsymbol{y}^a, y_0) = \deg(\boldsymbol{y}^b, y_0) \quad \wedge & \\ \qquad\qquad \psi(\boldsymbol{y}^a) <_M \psi(\boldsymbol{y}^b) & \\ \deg(\boldsymbol{y}^a) = \deg(\boldsymbol{y}^b) \quad \wedge \quad \deg(\boldsymbol{y}^a, y_0) = \deg(\boldsymbol{y}^b, y_0) \quad \wedge & \\ \qquad\qquad \psi(\boldsymbol{y}^a) = \psi(\boldsymbol{y}^b) \quad \wedge \quad \boldsymbol{y}^a \widetilde{<} \boldsymbol{y}^b & \end{cases} \tag{7.4}$$

This order is a monomial order, because it is a total order, $\boldsymbol{y}^0$ is the unique smallest monomial (it is the only one of degree 0), and it is compatible with the multiplication (every case is compatible).

For each $f \in \mathbb{K}[\boldsymbol{y}]$, we define η as the normal form (the remainder of the division algorithm) of f with respect to the ideal T and the monomial order $<_y$. With the notation of Prop. 4.2.1, the map η is as follows,

$$\eta(f) := \mathtt{NF}_{<_y,T}(f).$$

Recall that $\eta = \eta \circ \eta$ and $\mathtt{coker}(\eta) \cong \mathbb{K}[\boldsymbol{y}]/T$. We notice that for each poset in $\mathbb{K}[\boldsymbol{y}]/T$, η assigns the same normal form to all the elements that it contains. Therefore, we abuse notation, and we also use η to denote the map $\mathbb{K}[\boldsymbol{y}]/T \to \mathbb{K}[\boldsymbol{y}]$ that maps each poset to this unique normal form. As T is homogeneous, η is a 0-graded map. We extend $\widetilde{\phi}$ to $\phi : \mathbb{K}[S_M^h] \to \mathbb{K}[\boldsymbol{y}]$ as $\phi := \eta \circ \widetilde{\phi}$. This map is 0-graded and linear, but not a homomorphism. It holds $\psi \circ \phi = Id$ and $\phi \circ \psi = \eta$. Figure 7.1 summarizes these relations.

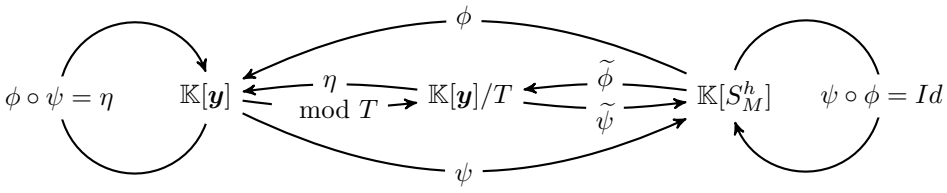


Figure 7.1: Maps between $\mathbb{K}[\boldsymbol{y}]$, $\mathbb{K}[\boldsymbol{y}]/T$, and $\mathbb{K}[S_M^h]$.

Theorem 7.2.9. *Let $I \subset \mathbb{K}[S_M^h]$ be a homogeneous ideal and consider the homogeneous ideal $J_y := \langle \phi(I) + T \rangle \subset \mathbb{K}[\boldsymbol{y}]$. If the Gröbner base of J_y with respect to $<_y$ is G_y, then $\psi(G_y)$ is a scant Gröbner base of I with respect to $\prec_h$.*

To prove the theorem we need the following lemmas.

Remark 7.2.10. *The monomial order $>_y$ behaves like* GRevLex *with respect to the variable y_0, that is,*

$$\textit{for every } f \in \mathbb{K}[\boldsymbol{y}],\ \textit{it holds } \eta(f \cdot y_0) = \eta(f) \cdot y_0.$$

Lemma 7.2.11. *For all $\boldsymbol{y}^\alpha \in \mathbb{K}[\boldsymbol{y}]$, it holds* $\deg(\eta(\boldsymbol{y}^\alpha), y_0) = \deg(\boldsymbol{y}^\alpha) - \delta(\psi(\boldsymbol{y}^\alpha))$.

Proof. Let $\boldsymbol{X}^{(s,d)} := \psi(\boldsymbol{y}^\alpha)$ and $\bar{d} = \delta(\boldsymbol{X}^{(s,d)})$. Note that $d = \deg(\boldsymbol{y}^\alpha)$, because ψ is 0-graded. We can write $\psi(\boldsymbol{y}^\alpha) = \chi^{-1}(\boldsymbol{X}^s) \cdot \boldsymbol{X}^{(0,d-\bar{d})}$. We recall that $\phi \circ \psi = \eta$ and apply $\phi = \eta \circ \widetilde{\phi}$ to the previous equality. As $\widetilde{\phi}$ is a homomorphism and by Rem. 7.2.10, we get the equality

$$\eta(\boldsymbol{y}^\alpha) = \eta(\widetilde{\phi}(\chi^{-1}(\boldsymbol{X}^s)) \cdot \widetilde{\phi}(\boldsymbol{X}^{(0,d-\bar{d})})) = \eta(\widetilde{\phi}(\chi^{-1}(\boldsymbol{X}^s)) \cdot y_0^{d-\bar{d}}) = \phi(\chi^{-1}(\boldsymbol{X}^s)) \cdot y_0^{d-\bar{d}}.$$

If $\deg(\phi(\chi^{-1}(\boldsymbol{X}^s)), y_0) = 0$, then $\deg(\eta(\boldsymbol{y}^\alpha), y_0) = 0 + d - \bar{d}$ and we proved our lemma.

If $\deg(\phi(\chi^{-1}(\boldsymbol{X}^s)), y_0) > 0$, then there is a monomial $\boldsymbol{y}^\beta$ such that $y_0 \cdot \boldsymbol{y}^\beta = \phi(\chi^{-1}(\boldsymbol{X}^s))$, and so, $\psi(y_0 \cdot \boldsymbol{y}^\beta) = \psi(\phi(\chi^{-1}(\boldsymbol{X}^s)))$. As $\psi \circ \phi = Id$ and ψ is a 0-graded epimorphism, then $\boldsymbol{X}^{(0,1)} \cdot \psi(\boldsymbol{y}^\beta) = \chi^{-1}(\boldsymbol{X}^s)$, but this is not possible by definition of homogenization, see Sec. 7.2.1. □

Corollary 7.2.12. *For all $\boldsymbol{X}^{(s,d)} \in \mathbb{K}[S_M^h]$, it holds* $\delta(\boldsymbol{X}^{(s,d)}) = d - \deg(\phi(\boldsymbol{X}^{(s,d)}), y_0)$.

As ψ and ϕ are 0-graded maps, by Lem. 7.2.11 and Cor. 7.2.12, they preserve the order.

Corollary 7.2.13. *For every $\boldsymbol{y}^\alpha, \boldsymbol{y}^\beta \in \mathbb{K}[\boldsymbol{y}]$ it holds,*

$$\eta(\boldsymbol{y}^\alpha) <_y \eta(\boldsymbol{y}^\beta) \implies \psi(\boldsymbol{y}^\alpha) \prec_h \psi(\boldsymbol{y}^\beta).$$

Recall that, as η is a normal form, if $\eta(f\, g) = f\, g$, then $\eta(f) = f$ and $\eta(g) = g$.

Lemma 7.2.14. *For every $\boldsymbol{y}^\alpha \in \mathbb{K}[\boldsymbol{y}]$ and $\boldsymbol{X}^{(s,d)} \in \mathbb{K}[S_M^h]$ it holds*

$$\boldsymbol{y}^\alpha | \phi(\boldsymbol{X}^{(s,d)}) \implies \psi(\boldsymbol{y}^\alpha) || \boldsymbol{X}^{(s,d)}.$$

Proof. Consider $\boldsymbol{y}^\beta$ such that $\boldsymbol{y}^\alpha \cdot \boldsymbol{y}^\beta = \phi(\boldsymbol{X}^{(s,d)})$. Then, $\psi(\boldsymbol{y}^\alpha) \cdot \psi(\boldsymbol{y}^\beta) = \boldsymbol{X}^{(s,d)}$. As η is a normal form, $\eta(\phi(\boldsymbol{X}^{(s,d)})) = \phi(\boldsymbol{X}^{(s,d)})$ and then, $\phi(\psi(\boldsymbol{y}^\alpha)) = \eta(\boldsymbol{y}^\alpha) = \boldsymbol{y}^\alpha$ and $\phi(\psi(\boldsymbol{y}^\beta)) = \eta(\boldsymbol{y}^\beta) = \boldsymbol{y}^\beta$. Hence, by Cor. 7.2.12 and Lem. 7.2.11, we conclude the proof by noting that,

$$\begin{aligned}\delta(\boldsymbol{X}^{(r,d)}) = \delta(\psi(\boldsymbol{y}^\alpha \cdot \boldsymbol{y}^\beta)) &= \deg(\boldsymbol{y}^\alpha \cdot \boldsymbol{y}^\beta) - \deg(\eta(\boldsymbol{y}^\alpha \cdot \boldsymbol{y}^\beta), y_0) \\ &= \deg(\boldsymbol{y}^\alpha) - \deg(\eta(\boldsymbol{y}^\alpha), y_0) + \deg(\boldsymbol{y}^\beta) - \deg(\eta(\boldsymbol{y}^\beta), y_0) \\ &= \delta(\psi(\boldsymbol{y}^\alpha)) + \delta(\psi(\boldsymbol{y}^\beta)).\end{aligned}$$

□

Corollary 7.2.15. *For all $f \in \mathbb{K}[S_M^h]$, for all $g \in \mathbb{K}[\boldsymbol{y}]$, it holds*

$$\mathtt{LM}_{<_y}(\eta(g)) | \mathtt{LM}_{<_y}(\phi(f)) \implies \mathtt{LM}_{\prec_h}(\psi(g)) || \mathtt{LM}_{\prec_h}(f).$$

Proof. By Cor. 7.2.13, $\psi(\mathtt{LM}_{<_y}(\eta(g))) = \mathtt{LM}_{\prec_h}(\psi(g))$ and $\psi(\mathtt{LM}_{<_y}(\phi(f))) = \mathtt{LM}_{<_y}(\psi(\phi(f))) = \mathtt{LM}_{\prec_h}(f)$. The proof follows from Lem. 7.2.14. □

Proof of Thm. 7.2.9. Consider $f \in I$, then $\phi(f) \in J_y$. Hence, there are $g_1, \ldots, g_k \in G_y$ and $p_1, \ldots, p_k \in \mathbb{K}[\boldsymbol{y}]$ such that $\phi(f) = \sum_{i=1}^k p_i \cdot g_i$. As $\psi \circ \phi = Id$ and ψ is an epimorphism such that $\psi(T) = 0$, then $\psi(\phi(f)) = f = \sum_{i=1}^k \psi(p_i) \cdot \psi(g_i)$ and $\psi(g_i), \ldots, \psi(g_k) \in I$. Hence, $\psi(G_y)$ generates I.

The set G_y is a Gröbner basis, then there is a $g \in G_y$ such that $\mathtt{LM}_{<_y}(g) | \mathtt{LM}_{<_y}(\phi(f))$. As $\phi(f) = \eta(\phi(f))$, it holds $\eta(\mathtt{LM}_{<_y}(\phi(f))) = \mathtt{LM}_{<_y}(\phi(f))$ and $\eta(\mathtt{LM}_{<_y}(g)) = \mathtt{LM}_{<_y}(g)$. As η is a normal form with respect to $<_y$, $\eta(\mathtt{LM}_{<_y}(g)) = \mathtt{LM}_{<_y}(\eta(g))$. By Cor. 7.2.15, $\mathtt{LM}_{\prec_h}(\psi(g)) || \mathtt{LM}_{\prec_h}(f)$. Hence, $\psi(G_y)$ is a scant Gröbner basis for I with respect to $\prec_h$. □

Corollary 7.2.16. *Given an ideal $I \subset \mathbb{K}[S_M^h]$ and a graded sparse order $\prec_h$, its scant Gröbner basis with respect to this order is finite.*

Proof. In Thm. 7.2.9 we construct a scant Gröbner basis of I with respect to $\prec_h$ from a (standard) Gröbner basis of an ideal of $\mathbb{K}[\boldsymbol{y}]$, finite as $\mathbb{K}[\boldsymbol{y}]$ is Noetherian. □

Remark 7.2.17. *As in the standard case, see Prop. 4.1.10, we can define the* reduced scant Gröbner basis *and adapt [CLO15, Prop. 2.7.6] to prove their finiteness and uniqueness.*

Non-homogeneous case. Let $\prec$ be a sparse order for $\mathbb{K}[S_M]$.

Lemma 7.2.18. *Let $I_h \subset \mathbb{K}[S_M^h]$ be a homogeneous ideal. Let $\prec_h$ be the graded sparse order for $\mathbb{K}[S_M^h]$ related to $\prec$ and consider G_h the scant Gröbner basis of I_h with respect to $\prec_h$. Then, $\chi(G_h)$ is a scant Gröbner Basis for $\chi(I_h)$ with respect to $\prec$.*

Proof. The set $\chi(G_h)$ generates $\chi(I_h)$. Note that for homogeneous polynomials, $\texttt{LM}_{\prec_h}$ commutes with the dehomogenization, that is for any homogeneous polynomial $g_h \in \mathbb{K}[S_M^h]$, $\texttt{LM}_{\prec}(\chi(g_h)) = \chi(\texttt{LM}_{\prec_h}(g_h))$. Consider $f \in \chi(I_h)$, then there is an $f_h \in I_h$ such that $f_h = \chi(f)$. In addition, there is $g_h \in G_h$ such that $\texttt{LM}_{\prec_h}(g_h)||\texttt{LM}_{\prec_h}(f_h)$. Let $\boldsymbol{X}^{(s,d)} \in \mathbb{K}[S_M^h]$ such that $\texttt{LM}_{\prec_h}(g_h) \cdot \boldsymbol{X}^{(s,d)} = \texttt{LM}_{\prec_h}(f_h)$ and $\delta(\texttt{LM}_{\prec_h}(g_h)) + \delta(\boldsymbol{X}^{(s,d)}) = \delta(\texttt{LM}_{\prec_h}(f_h))$. The sparse degree δ is independent of the homogeneous degree, so $\delta(\chi(\texttt{LM}_{\prec_h}(g_h))) + \delta(\boldsymbol{X}^s) = \delta(\chi(\texttt{LM}_{\prec_h}(f_h)))$. Hence, $\delta(\texttt{LM}_{\prec}(\chi(g_h))) + \delta(\boldsymbol{X}^s) = \delta(\texttt{LM}_{\prec}(f))$ and $\texttt{LM}_{\prec}(\chi(g_h)) \cdot \boldsymbol{X}^s = \texttt{LM}_{\prec}(f)$, so $\texttt{LM}_{\prec}(\chi(g_h))||\texttt{LM}_{\prec}(f)$ and $\chi(G_h)$ is a scant Gröbner basis of $\chi(I_h)$ with respect to $\prec$. □

Corollary 7.2.19. *There is a finite scant Gröbner basis for any $I \subset \mathbb{K}[S_M]$ with respect to $\prec$.*

Proof. Note that, for the homogenization of I, $\chi^{-1}(I)$, it holds $\chi(\chi^{-1}(I)) = I$. By Cor. 7.2.16, we can consider a finite scant Gröbner basis G_h of $\chi^{-1}(I)$ with respect to $\prec_h$. By Lem. 7.2.18, the set $\chi(G_h)$ is a finite scant Gröbner basis of I with respect to $\prec$. □

7.2.3 Algorithms

Homogeneous case. To compute a scant Gröbner basis of a homogeneous ideal $I := \langle f_1, \ldots, f_k \rangle$ with respect to $\prec_h$, we introduce the d-scant Gröbner bases, see Def. 4.4.1. A d-scant Gröbner basis of I is a finite set of polynomials $G \subset I$ such that for each homogeneous $f \in I$ with $\deg(f) \leq d$, it holds $f \in \langle G \rangle$ and there is a $g \in G$ such that $\texttt{LM}_{\prec_h}(g)||\texttt{LM}_{\prec_h}(f)$. For big enough d, for example equal to the maximal degree of a homogeneous polynomial in a reduced scant Gröbner basis, a d-scant Gröbner basis is a scant Gröbner basis. The *witness degree* of I is the minimal d such that a d-scant Gröbner basis is a scant Gröbner basis. We compute d-scant Gröbner bases using linear algebra. We follow a similar approach as Lazard's algorithm, see Sec. 4.4.

Definition 7.2.20. *Let $\prec_h$ be a monomial ordering and consider $d \in \mathbb{N}$. A Macaulay matrix $\mathcal{M}_{\prec_h,d}$ is a matrix whose columns are indexed by monomials in $\mathbb{K}[S_M^h]_d$ and whose rows are indexed by polynomials in $\mathbb{K}[S_M^h]_d$. The columns of $\mathcal{M}_{\prec_h,d}$ are sorted in decreasing order with respect to $\prec_h$. The element in $\mathcal{M}_{\prec_h,d}$ corresponding to the column indexed by $\boldsymbol{X}^{(s,d)}$ and the row indexed by $f \in \mathbb{K}[S_M^h]_d$ is the coefficient of the monomial $\boldsymbol{X}^{(s,d)}$ in f. We define $\texttt{Rows}(\mathcal{M}_{\prec_h,d})$ as the set of* non-zero *polynomials that index the rows of $\mathcal{M}_{\prec_h,d}$.*

Following the same idea as in Lazard's algorithm, see Sec. 4.4, we can compute a d-scant Gröbner basis by computing the reduced row echelon form of several Macaulay matrices. The proof of the following lemma follows from [Laz83].

Lemma 7.2.21. *Given homogeneous polynomials $f_1, \ldots, f_k \in \mathbb{K}[S_M^h]$, consider the homogeneous ideal $I := \langle f_1, \ldots, f_k \rangle \subset \mathbb{K}[S_M^h]$. Let $\mathcal{M}^{\textit{Lazard}}_{\prec_h,d}$ be the Macaulay matrix whose columns are all the monomials in $\mathbb{K}[S_M^h]_d$ sorted in decreasing order by $\prec_h$, and the rows are all the products of the form $\boldsymbol{X}^{(s,d-\deg(f_i))} \cdot f_i \in \mathbb{K}[S_M^h]_d$, for each $1 \leq i \leq k$. Let $\widetilde{\mathcal{M}}^{\textit{Lazard}}_{\prec_h,d}$ be the reduced row echelon form of $\mathcal{M}^{\textit{Lazard}}_{\prec_h,d}$. Then,*

the polynomials whose leading monomial can not be divided by the leading monomial of a polynomial obtained in smaller degree, that is

$$\bigcup_{i=1}^{d} \left\{ f \in \texttt{Rows}(\widetilde{\mathcal{M}}^{\textit{Lazard}}_{\prec_h,i}) : \left(\nexists\, g \in \bigcup_{j=1}^{i-1} \texttt{Rows}(\widetilde{\mathcal{M}}^{\textit{Lazard}}_{\prec_h,j}) \right) \texttt{LM}_{\prec_h}(g) || \texttt{LM}_{\prec_h}(f) \right\},$$

form a d-scant Gröbner basis of I.

When we compute the reduced row echelon form of $\mathcal{M}^{\text{Lazard}}_{\prec_h,d}$, many rows of the matrix reduce to zero during the Gaussian elimination procedure, and so, we perform redundant computations. We can adapt the `F5` criterion (Prop. 4.4.6) to identify these rows and avoid the reductions. We follow a *polynomial-by-polynomial* strategy: For each i, we first compute a scant Gröbner basis of $\langle f_1, \ldots, f_{i-1} \rangle$ and then we extend this basis to a scant Gröbner basis of $\langle f_1, \ldots, f_i \rangle$.

In what follows, let G be a scant Gröbner basis of a homogeneous ideal $I \subset \mathbb{K}[S^h_M]$ with respect to a order $\prec_h$ and consider a homogeneous polynomial $f \in \mathbb{K}[S^h_M]_{d_f}$ of degree d_f. Our objective is to compute a scant Gröbner basis of the ideal $I + \langle f \rangle$. We compute this scant Gröbner basis using a d-scant Gröbner basis.

For a fixed $d \in \mathbb{N}$, consider a set $\mathcal{R}_d \subset I \cap \mathbb{K}[S^h_M]_d$ such that:

- For each homogeneous $f \in I \cap \mathbb{K}[S^h_M]_d$, there is $p \in \mathcal{R}_d$ such that $\texttt{LM}_{\prec_h}(f) = \texttt{LM}_{\prec_h}(p)$.
- For each $p \in \mathcal{R}_d$, there is $g \in G$ such that $\texttt{LM}_{\prec_h}(g) || \texttt{LM}_{\prec_h}(p)$ and $p = \frac{\texttt{LM}_{\prec_h}(g)}{\texttt{LM}_{\prec_h}(p)}\, g$.
- All the polynomials in $\mathcal{R}_d$ have different leading monomials.

Moreover, consider the set of monomials $\mathfrak{b}_d \subset \mathbb{K}[S^h_M]_{d-d_f}$ such that

$$\mathfrak{b}_d := \{ \boldsymbol{X}^{(s,d-d_f)} \in \mathbb{K}[S^h_M]_{d-d_f} \ : \ \nexists\, g \in G \text{ s.t. } \texttt{LM}_{\prec_h}(g) || \boldsymbol{X}^{(s,d-d_f)} \}.$$

The proof of the following lemma follows from [BFS15, Prop. 8].

Lemma 7.2.22 (`Scant-F5` criterion)**.** *With the notation of the previous paragraphs, consider the Macaulay matrix $\mathcal{M}^{F5}_{\prec_h,d}$ with columns indexed by the monomials in $\mathbb{K}[S^h_M]_d$ in decreasing order with respect to $\prec_h$ and rows indexed by the polynomials*

$$\mathcal{R}_d \cup \{ \boldsymbol{X}^{(s,d-d_f)} \cdot f \ : \ \boldsymbol{X}^{(s,d-d_f)} \in \mathfrak{b}_d \}.$$

Let $\widetilde{\mathcal{M}}^{F5}_{\prec_h,d}$ be the reduced row echelon form of $\mathcal{M}^{F5}_{\prec_h,d}$. Then, $\texttt{Rows}(\widetilde{\mathcal{M}}^{F5}_{\prec_h,d}) = \texttt{Rows}(\widetilde{\mathcal{M}}^{\textit{Lazard}}_{\prec_h,d})$, where $\widetilde{\mathcal{M}}^{\textit{Lazard}}_{\prec_h,d}$ is the Macaulay matrix of Lem. 7.2.21 for the set $G \cup \{f\}$ (which generates the $I + \langle f \rangle$).

Moreover, if f is not a zero divisor in $\mathbb{K}[S^h_M]/I$, then $\mathcal{M}^{F5}_{\prec_h,d}$ is full-rank and we can compute $\widetilde{\mathcal{M}}^{F5}_{\prec_h,d}$ without performing reductions to zero.

Given the *witness degree* of $I + \langle f \rangle$, Lem. 7.2.22 supports an algorithm to compute a scant Gröbner basis of $I + \langle f \rangle$ (Alg. 7). If f is not a zero divisor in $\mathbb{K}[S^h_M]/I$, then the algorithm performs no reductions to zero.

Algorithm 7 `scantMatrixF5`: Scant Matrix-F5 with respect to $\prec$

Input: A sparse order $\prec_h$, a scant Gröbner basis G of the homogeneous ideal I with respect to $\prec_h$, a homogeneous polynomial $f \in \mathbb{K}[S_M^h]_{d_f}$ of degree d_f, and the witness degree d^{wit} of $I + \langle f \rangle$.

Output: The set H is a scant Gröbner basis of $I + \langle f \rangle$ with respect to $\prec_h$.

$H \leftarrow \emptyset$
for $d = 1$ to d^{wit} **do**
  $\mathcal{M}_{\prec_h,d} \leftarrow$ Macaulay matrix with columns indexed by the monomials in $\mathbb{K}[S_M^h]_d$ in decreasing order by $\prec_h$.
  for $\boldsymbol{X}^{(s,d)} \in \mathbb{K}[S_M^h]_d$ **do**
    if $\exists g \in G$ such that $\mathtt{LM}_{\prec_h}(g) || \boldsymbol{X}^{(s,d)}$ **then**
      Add to $\mathcal{M}_{\prec_h,d}$ the polynomial $\frac{\boldsymbol{X}^{(s,d)}}{\mathtt{LM}_{\prec_h}(g)} \cdot g$.
    end if
  end for
  for $\boldsymbol{X}^{(s,d-d_f)} \in \mathbb{K}[S_M^h]_{d-d_f}$ **do**
    if $\nexists g \in G$ such that $\mathtt{LM}_{\prec_h}(g) || \boldsymbol{X}^{(s,d-d_f)}$ **then**
      Add to $\mathcal{M}_{\prec_h,d}$ the polynomial $\boldsymbol{X}^{(s,d-d_f)} \cdot f$.
    end if
  end for
  $\widetilde{\mathcal{M}}_{\prec_h,d} \leftarrow$ Reduced row echelon form of $\mathcal{M}_{\prec_h,d}$.
  $H \leftarrow H \cup \{h \in \mathtt{Rows}(\widetilde{\mathcal{M}}_{\prec_h,d})$ such that $(\nexists\, g \in H) : \mathtt{LM}_{\prec_h}(g) || \mathtt{LM}_{\prec_h}(h)\}$.
end for
return H.

Non-homogeneous case. Given an ideal $I := \langle f_1 \dots f_r \rangle \subset \mathbb{K}[S_M]$, we homogenize the polynomials and use Lem. 7.2.21 to compute a scant Gröbner basis with respect to $\prec_h$. By Lem. 7.2.18, if we dehomogenize the computed basis, we obtain a scant Gröbner basis with respect to $\prec$ of I. Instead of homogenizing all polynomials f_i simultaneously, we consider an iterative approach, which, under regularity assumptions, involves only full-rank matrices, and hence avoids all reductions to zero. The following lemma allows us to compute a scant Gröbner basis in the homogeneous case, from the non-homogeneous one. This lemma is similar to Prop. 4.3.13.

Lemma 7.2.23. *If G is a scant Gröbner basis of the ideal $I \subset \mathbb{K}[S_M]$ with respect to $\prec$, then $G_h := \chi^{-1}(G)$ is a scant Gröbner basis of $\langle \chi^{-1}(I) \rangle \subset \mathbb{K}[S_M^h]$ with respect to $\prec_h$.*

Proof. First note that the homogenization commutes with the leading monomial, that is $\forall g \in \mathbb{K}[S_M]$, $\mathtt{LM}_{\prec_h}(\chi^{-1}(g)) = \chi^{-1}(\mathtt{LM}_{\prec}(g))$. Let $f_h \in \langle \chi^{-1}(I) \rangle$. Similarly to Prop. 4.3.2, we can write f_h as $\boldsymbol{X}^{(0,\deg(f_h)-\delta(f_h))} \cdot \chi^{-1}(\chi(f_h))$. Consider $g \in G$ such that $\mathtt{LM}_{\prec}(g) || \mathtt{LM}_{\prec}(\chi(f_h))$. By definition 7.2.5, $\chi^{-1}(\mathtt{LM}_{\prec}(g)) || \chi^{-1}(\mathtt{LM}_{\prec}(\chi(f_h)))$, and by commutativity, it holds that $\mathtt{LM}_{\prec_h}(\chi^{-1}(g)) || \mathtt{LM}_{\prec_h}(\chi^{-1}(\chi(f_h))$. The sparse degree and the leading monomials with respect to $\prec_h$ are invariants under the multiplication by $\boldsymbol{X}^{(0,1)}$. Hence, $\mathtt{LM}_{\prec_h}(\chi^{-1}(g)) || \mathtt{LM}_{\prec_h}(f_h)$. To conclude, we have to prove that G_h is a basis of $\langle \chi^{-1}(I) \rangle$. As for each $f_h \in \chi^{-1}(I)$ there is a $g \in G$ such that $\mathtt{LM}_{\prec_h}(\chi^{-1}(g)) || \mathtt{LM}_{\prec_h}(f_h)$. Thus, the remainder of the division algorithm (Lem. 7.2.8) is zero, and so we obtain a representation of f_h as a polynomial combination of the elements of $\chi^{-1}(G)$. □

Corollary 7.2.24. *Let $I \subset \mathbb{K}[S_M]$ be an (non-homogeneous) ideal and consider the (non-homogeneous) polynomial $f \in \mathbb{K}[S_M]$. Let G be a (non-homogeneous) scant Gröbner basis of I with respect to $\prec$ and H^h be a (homogeneous) scant Gröbner basis of $\langle \chi^{-1}(G) + \chi^{-1}(f) \rangle$ with respect to $\prec_h$. Then, $\chi(H^h)$ is a (non-homogeneous) scant Gröbner basis of $\langle I + f \rangle$ with respect to $\prec$.*

Corollary 7.2.24 supports an iterative algorithm to compute a scant Gröbner basis of $\langle f_1, \dots, f_r \rangle \subset \mathbb{K}[S_M]$ (Alg. 8):

- For each $i \leq r$, let $I_i := \langle f_1, \dots, f_i \rangle$ and consider G_i a finite scant Gröbner basis of I_i with respect to $\prec$.
- Let G_i^h be a (homogeneous) scant Gröbner basis of $\langle \chi^{-1}(G_{i-1}) \rangle + \langle \chi^{-1}(f_i) \rangle$ with respect to $\prec_h$.
- By Lem. 7.2.23,the set $\chi^{-1}(G_{i-1})$ is a (homogeneous) scant Gröbner basis of $\chi^{-1}(I_{i-1})$.
- Hence, given the witness degree of $\chi^{-1}(I_{i-1}) + \langle \chi^{-1}(f_i) \rangle$ and the (homogeneous) scant Gröbner basis $\chi^{-1}(G_{i-1})$, Alg. 7 computes the (homogeneous) scant Gröbner basis G_i^h.
- By Cor. 7.2.24, we recover the scant Gröbner basis G_i from $\chi(G_i^h)$.

By the `Scant-F5` criterion (Lem. 7.2.22), if for each i, the polynomial $\chi^{-1}(f_i)$ is not a zero divisor in $\mathbb{K}[S_M^h]/\chi^{-1}(\langle I_{i-1} \rangle)$, then Alg. 8 performs no reductions to zero. We conclude this section by proving that, if $f_1, \dots, f_k \in \mathbb{K}[S_M]$ forms a regular sequence over $\mathbb{K}[S_M]$, then the latter holds.

Lemma 7.2.25. *If $f_1, \dots, f_r \in \mathbb{K}[S_M]$ is a regular sequence, then for each $i \leq k$, $\chi^{-1}(f_i)$ is not a zero divisor of $\mathbb{K}[S_M^h]/\chi^{-1}(\langle I_{i-1} \rangle)$.*

Proof. If $\chi^{-1}(f_i)$ is a zero divisor of $\mathbb{K}[S_M^h]/\chi^{-1}(I_{i-1})$, then there is a $g \in \mathbb{K}[S_M^h]$ such that $g \notin \chi^{-1}(\langle f_1, \ldots, f_{i-1}\rangle)$ and $g \cdot \chi^{-1}(f_i) \in \chi^{-1}(\langle f_1, \ldots, f_{i-1}\rangle)$. By definition of the dehomogenization of an ideal, $\chi(g) \notin \langle f_1, \ldots, f_{i-1}\rangle$ but, as χ is a homomorphism, $\chi(g) \cdot f_i \in \langle f_1, \ldots, f_{i-1}\rangle$. So, $f_1, \ldots, f_i$ is not a regular sequence. □

Remark 7.2.26. *The* (sparse) regular sequences *(Def. 7.1.7) satisfy the assumptions of Lem. 7.2.25, see Lem. 7.1.9.*

Algorithm 8 $\mathtt{M}^2$: Mixed scant Matrix-F5 with respect to $\prec$

Input: A list of affine polynomial $f_1, \ldots, f_r \in \mathbb{K}[S_M]$, a sparse order $\prec$ and a list of numbers $d_1^{wit}, \ldots, d_r^{wit} \in \mathbb{N}$ such that d_i^{wit} is the witness degree of $\chi^{-1}(\langle f_1, \ldots, f_{i-1}\rangle) + \langle \chi^{-1}(f_i)\rangle$ with respect to $\prec_h$.

Output: The set G_r is a scant Gröbner basis of $\langle f_1, \ldots, f_r\rangle$ with respect to $\prec$.

$G_0 \leftarrow \emptyset$.
for $i = 1$ to r **do**
  $G_i^h \leftarrow \mathtt{scantMatrixF5}(\chi^{-1}(G_{i-1}), \chi^{-1}(f_i), \prec_h, d_i^{wit})$.
  $G_i \leftarrow \chi\left(G_i^h\right)$.
end for
return G_r

7.3 Gröbner basis over semigroup algebras

We follow [Stu93, FSS14] and consider Gröbner bases over semigroup algebras. We construct a semigroup algebra related to the Newton polytopes of the input polynomials and compute Gröbner bases for the ideal generated by the original polynomials in this semigroup algebra.

We embed the systems in semigroup algebras because in this place they "behave" in a predictable way that we can exploit algorithmically. Semigroup algebras are related to toric varieties. An affine toric variety is the spectrum of a semigroup algebra [CLS11, Thm. 1.1.17]. Hence, the variety defined by the polynomials over the semigroup is a subvariety of a toric variety. This variety is different from the one defined by the polynomials over the original polynomial algebra, but they are related and in many applications the difference is irrelevant, e.g., [EM99a]. We refer to [CLS11] for an introduction to toric varieties and to [Stu96] for their relation with Gröbner basis.

We extend [FSS14] to *sparse mixed systems*. We relax their regularity assumptions and we introduce an F5-like criterion (Prop. 4.4.6) that, under (sparse) regularity assumptions, see Def. 7.1.7, predicts all the rows reducing to zero during Gröbner bases computation.

7.3.1 Definitions

In this section, we follow the notation and definitions from Sec. 7.1. We emphasize that $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ is a $\mathbb{Z}^r$-graded algebra, in contrast to the algebra $\mathbb{K}[S^h_M]$ from Sec. 7.2, which is $\mathbb{Z}$-graded.

Gröbner bases

The definitions of monomial ordering (Def. 4.6.1) and Gröbner basis over semigroup algebras (Def. 4.6.4) appear in Sec. 4.6. We introduce a particular family of monomial orders for $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ that we can relate to monomial orders in $\mathbb{K}[S_\Delta]$ and $\mathbb{K}[\mathbb{N}^r]$.

Definition 7.3.1 (Multigraded monomial order)**.** *We say that a monomial order $<$ for $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ is multigraded, if there are monomial orders $<_\Delta$ for $\mathbb{K}[S_\Delta]$ and $<_h$ for $\mathbb{K}[\mathbb{N}^r]$ such that, for every $\boldsymbol{X}^{(\boldsymbol{\alpha}_1,\boldsymbol{d}_1)}, \boldsymbol{X}^{(\boldsymbol{\alpha}_2,\boldsymbol{d}_2)} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$, it holds*

$$\boldsymbol{X}^{(\boldsymbol{\alpha}_1,\boldsymbol{d}_1)} < \boldsymbol{X}^{(\boldsymbol{\alpha}_2,\boldsymbol{d}_2)} \iff \begin{cases} \boldsymbol{X}^{\boldsymbol{d}_1} <_h \boldsymbol{X}^{\boldsymbol{d}_2} \text{ or} \\ \boldsymbol{d}_1 = \boldsymbol{d}_2 \text{ and } \boldsymbol{X}^{\boldsymbol{\alpha}_1} <_\Delta \boldsymbol{X}^{\boldsymbol{\alpha}_2} \end{cases} . \tag{7.5}$$

Remark 7.3.2. *In what follows, given a multigraded monomial order $<$ for $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$, we also use the same symbol, that is $<$, for the associated monomial order of $\mathbb{K}[S_\Delta]$.*

Multigraded monomial orders are "compatible" with the dehomogenization morphism (Def. 7.1.2).

Lemma 7.3.3. *Consider a polynomial $f \in \mathbb{K}[S_\Delta]$. Let $<$ be a multigraded monomial order. For any multidegree $\boldsymbol{d}$ and any homogeneous $F \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$ such that $\chi(F) = f$, it holds $\mathtt{LM}_<(f) = \chi(\mathtt{LM}_<(F))$.*

Solutions at infinity

As we discussed in Sec. 2.12, semigroup algebras are related to toric varieties. A toric variety is an irreducible variety X that contains $(\mathbb{C}^*)^n$, see Eq. (2.7), as an open subset such that the action of $(\mathbb{C}^*)^n$ on itself extends to an algebraic action of $(\mathbb{C}^*)^n$ on X [CLS11, Def. 3.1.1]. Semigroup algebras correspond to the coordinate rings of the affine pieces of X.

Given an integer polytope Δ, we can define a projective complete normal irreducible toric variety X associated to it [CLS11, Sec. 2.3]. Likewise, given a polynomial system $(f_1, \dots, f_m)$, we can define a projective toric variety X associated to the Minkowski sum of their Newton polytopes. We can homogenize these polynomials in a way that they belong to the total coordinate ring of X [CLS11, Sec. 5.4]. This homogenization is related to the facets of the polytopes.

To be more precise, given an integer polytope $\Delta \subset \mathbb{R}^n$, we say that an integer polytope Δ_1 is a $\mathbb{N}$-*Minkowski summand* of Δ if there is a $k \in \mathbb{N}$ and another integer polytope Δ_2 such that $\Delta_1 + \Delta_2 = k \cdot \Delta$ [CLS11, Def. 6.2.11]. Every $\mathbb{N}$-Minkowski summand Δ_1 of Δ defines a torus-invariant basepoint free Cartier divisor D of the projective toric variety X associated to Δ [CLS11, Cor. 6.2.15]. This divisor defines an invertible sheaf $\mathcal{O}_X(D)$ whose global sections form the vector space of polynomials in $\mathbb{K}[\mathbb{Z}^n]$ whose Newton polytopes are contained in Δ_1 [Mas16, Lem. 1]. Therefore, to homogenize $f_1, \dots, f_r$ over X we need to choose polytopes $\Delta_1, \dots, \Delta_r$ such that all of them are $\mathbb{N}$-Minkowski summands of Δ associated to X and $\mathrm{NP}(f_i) \subset \Delta_i$. Hence, for any homogeneous polynomial $F \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{(d_1,\dots,d_r)}$, we can homogenize $\chi(F)$ (in the total coordinate ring of X) with respect to the $\mathbb{N}$-Minkowski summand $\sum_i d_i \Delta_i$ of Δ.

We alert the reader that homogeneity in $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$ is different from homogeneity in the total coordinate ring of X, see [CLS11, Sec. 5.4]. Nevertheless, there two notions are related through the degree $\boldsymbol{d}$.

Definition 7.3.4 (Solutions at infinity)**.** *Let $(f_1, \dots, f_m)$ be a system of polynomials. Let X be the projective toric variety associated to a polytope Δ such that the Newton polytope of f_i is a $\mathbb{N}$-Minkowski summand of Δ, for all i. We say that the system has no solutions at infinity with respect to X if the homogenized system with respect to their Newton polytopes has no solutions over $X \setminus (\mathbb{C}^*)^n$.*

Proposition 7.3.5. *Consider a square system $(f_1, \dots, f_n)$ having a finite number of solutions over $(\mathbb{C}^*)^n$. Let X be the projective toric variety associated to the corresponding Newton polytopes. Then, the number of solutions of the homogenized system over X, counting multiplicities, is exactly the BKK bound (Thm. 2.12.19). When the original system has no solutions at infinity, then the BKK is tight over $(\mathbb{C}^*)^n \subset X$.*

For a proof of the previous proposition see, for example, [Mas16, Thm. 3] or [Sop13, Thm. 2.6].

7.3.2 Algorithm

To compute Gröbner basis over $\mathbb{K}[S_\Delta]$ we work over $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$. We follow the Lazard's approach, see Alg. 3, adapted to the semigroup case; we "linearize" the problem by reducing the Gröbner basis computation to a linear algebra problem.

Lemma 7.3.6. *Consider homogeneous polynomials $F_1, \dots, F_m \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ and a multigraded monomial order $<$ for $\mathbb{K}[S_\Delta]$ (Def. 7.3.1). There is a multidegree $\mathbf{d}$ and homogeneous $\{G_1, \dots, G_t\} \subset$*

$\langle F_1,\dots,F_m\rangle \cap \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$ such that $\{\chi(G_1),\dots,\chi(G_t)\}$ is a Gröbner basis of the ideal $\langle\chi(F_1),\dots,\chi(F_m)\rangle$ with respect to the associated monomial order $<$ (Rem. 7.3.2).

Proof. Let $g_1,\dots,g_t \in \mathbb{K}[S_\Delta]$ be a Gröbner basis for the ideal $\langle\chi(F_1),\dots,\chi(F_m)\rangle$ with respect to $<$. By Obs. 7.1.3, there are polynomials $\bar{G}_1,\dots,\bar{G}_t \in \langle F_1,\dots,F_m\rangle$ such that $\chi(\bar{G}_i) = g_i$, for $i \in [t]$. Consider $\boldsymbol{d} \in \mathbb{N}^r$ such that $\boldsymbol{d} \geq \deg(\bar{G}_i)$, for $i \in [t]$. It suffices to consider $G_i = \boldsymbol{X}^{(0,\boldsymbol{d}-\deg(\bar{G}_i))}\,\bar{G}_i \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$, for $i \in [t]$. □

When we know a multidegree $\boldsymbol{d}$ that satisfies Lem. 7.3.6, we can compute the Gröbner basis over $\mathbb{K}[S_\Delta]$ using linear algebra. For this task we need to introduce the Macaulay matrix, see Def. 4.4.3.

Definition 7.3.7 (Macaulay matrix)**.** *A Macaulay matrix $\mathcal{M}$ of degree $\boldsymbol{d} \in \mathbb{N}^r$ with respect to a monomial order $<$ is a matrix whose columns are indexed by all monomials $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d})} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$ and the rows by polynomials in $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$. The indices of the columns are sorted in decreasing order with respect to $<$. The element of $\mathcal{M}$ whose row corresponds to a polynomial F and whose column corresponds to a monomial $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d})}$ is the coefficient of the monomial $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d})}$ of F. Let $\mathtt{Rows}(\mathcal{M})$ be the set of* non-zero *polynomials that index the rows of $\mathcal{M}$ and $\mathtt{LM}_<(\mathtt{Rows}(\mathcal{M}))$ be the set of leading monomials of these polynomials.*

Remark 7.3.8. *As the columns of the Macaulay matrices are sorted in decreasing order with respect to a monomial order, the leading monomial of a polynomial associated to a row corresponds to the index of the column of the first non-zero element in this row.*

Definition 7.3.9. *Given a Macaulay matrix $\mathcal{M}$, let $\widetilde{\mathcal{M}}$ be a new Macaulay matrix corresponding to the reduced row echelon form of $\mathcal{M}$. We can compute $\widetilde{\mathcal{M}}$ by applying Gaussian elimination to $\mathcal{M}$.*

Remark 7.3.10. *When we perform row operations (excluding multiplication by 0) to a Macaulay matrix, we do not change the ideal spanned by the polynomials corresponding to its rows.*

We use Macaulay matrices to compute a triangular basis for the vector space $\langle F_1,\dots,F_k\rangle_{\boldsymbol{d}} := \langle F_1,\dots,F_k\rangle \cap \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$ by Gaussian elimination.

Lemma 7.3.11. *Consider homogeneous polynomials $F_1,\dots,F_k \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ of multidegrees $\boldsymbol{d}_1,\dots,\boldsymbol{d}_k \in \mathbb{N}^r$, respectively, and a multigraded monomial order $<$. Let $\mathcal{M}^k_{\boldsymbol{d}}$ be the Macaulay matrix of degree $\boldsymbol{d}$ whose columns are sorted with respect to $<$ and its rows correspond to the polynomials that we obtain by considering the product of every monomial of multidegree $\boldsymbol{d}-\boldsymbol{d}_i$ and every polynomial F_i; that is*

$$\mathtt{Rows}(\mathcal{M}^k_{\boldsymbol{d}}) = \left\{ \boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d}-\boldsymbol{d}_i)} F_i : i \in [k], \boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d}-\boldsymbol{d}_i)} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}-\boldsymbol{d}_i} \right\}. \tag{7.6}$$

Let $\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}$ be the reduced row echelon form of the Macaulay matrix $\mathcal{M}^k_{\boldsymbol{d}}$ (Def. 7.3.9).

Then, the set $\mathtt{LM}_<(\mathtt{Rows}(\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}))$, see Def. 7.3.9, is the set of all the leading monomials of the ideal $\langle F_1,\dots,F_k\rangle$ at degree $\boldsymbol{d}$.

Proof. We prove that $\mathtt{LM}_<(\mathtt{Rows}(\widetilde{\mathcal{M}}^k_{\boldsymbol{d}})) = \mathtt{LM}_<(\langle F_1,\dots,F_k\rangle_{\boldsymbol{d}})$. First, we show that $\mathtt{LM}_<(\mathtt{Rows}(\widetilde{\mathcal{M}}^k_{\boldsymbol{d}})) \supseteq \mathtt{LM}_<(\langle F_1,\dots,F_k\rangle_{\boldsymbol{d}})$. Let G be a polynomial in the vector space of polynomials of degree $\boldsymbol{d}$ in $\langle F_1,\dots,F_k\rangle$. This vector space, $\langle F_1,\dots,F_k\rangle_{\boldsymbol{d}}$, is isomorphic to the row space of $\mathcal{M}^k_{\boldsymbol{d}}$, which, in turn, is the same as the row space of $\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}$, by Rem. 7.3.10. Hence, there is a vector v in the row space of $\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}$

Algorithm 9 `ComputeGB`

Input: A list of affine polynomials $f_1, \ldots, f_k \in \mathbb{K}[S_\Delta]$ and a monomial order $<$.
Output: A Gröbner basis for $\langle f_1, \ldots, f_k \rangle$ with respect to $<$.

for all f_i **do**
  Choose $F_i \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}_i}$ of multidegree $\boldsymbol{d}_i$ such that $\chi(F_i) = f_i$.
end for
Pick a big enough $\boldsymbol{d} \in \mathbb{N}^r$ that satisfies Lem. 7.3.6.
$\mathcal{M}^k_{\boldsymbol{d}} \leftarrow$ Macaulay matrix of degree $\boldsymbol{d}$ with respect to a multigraded order associated to $<$ (Def. 7.3.1)
for all F_i **do**
  for all $\boldsymbol{X}^{(\boldsymbol{\alpha}, \boldsymbol{d}-\boldsymbol{d}_i)} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}-\boldsymbol{d}_i}$ **do**
    Add the polynomial $\boldsymbol{X}^{(\boldsymbol{\alpha}, \boldsymbol{d}-\boldsymbol{d}_i)} F_i$ as row to $\mathcal{M}^k_{\boldsymbol{d}}$.
  end for
end for
$\widetilde{\mathcal{M}}^k_{\boldsymbol{d}} \leftarrow$ `GaussianElimination`$(\mathcal{M}^k_{\boldsymbol{d}})$.
return $\chi($`Rows`$(\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}))$.

that corresponds to G. Let s be the index of the first non-zero element of v. As $\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}$ is in row echelon form and v belongs to its row space, there is a row of $\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}$ such that its first non-zero element is also at the s-th position. Let F be the polynomial that corresponds to this row. Finally, the leading monomials of the polynomials F and G are the same, that is $\mathtt{LM}_<(G) = \mathtt{LM}_<(F)$, by Rem. 7.3.8.

The other direction is straightforward. □

Theorem 7.3.12. *Consider the ideal generated by homogeneous polynomials $F_1, \ldots, F_k \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ of multidegrees $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_k \in \mathbb{N}^r$, respectively. Consider a multigraded monomial order $<$ and a multidegree $\boldsymbol{d} \in \mathbb{N}^r$ that satisfy Lem. 7.3.6. Let $\mathcal{M}^k_{\boldsymbol{d}}$ and $\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}$ be the Macaulay matrices of Lem. 7.3.11.*

Then, the set $\chi($`Rows`$(\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}))$, see Def. 7.1.2, contains a Gröbner basis of the ideal $\langle \chi(F_1), \ldots, \chi(F_k) \rangle \subset \mathbb{K}[S_\Delta]$ with respect to $<$.

Proof. Let $R :=$ `Rows`$(\widetilde{\mathcal{M}}^k_{\boldsymbol{d}})$ be the set of polynomials indexing the rows of $\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}$. By Lem. 7.3.11, for every $G \in \langle F_1, \ldots, F_k \rangle_{\boldsymbol{d}}$ there is a $F \in R$ such that $\mathtt{LM}_<(G) = \mathtt{LM}_<(F)$. As $<$ is a multigraded order, it holds $\mathtt{LM}_<(\chi(G)) = \mathtt{LM}_<(\chi(F))$ (Lem. 7.3.3). As $\boldsymbol{d}$ satisfies Lem. 7.3.6, for every $h \in \langle \chi(F_1), \ldots, \chi(F_k) \rangle$ there is $G \in \langle F_1, \ldots, F_k \rangle_{\boldsymbol{d}}$ such that $\mathtt{LM}_<(\chi(G))$ divides $\mathtt{LM}_<(h)$. Hence, there is an $F \in R$ such that $\mathtt{LM}_<(\chi(F))$ divides $\mathtt{LM}_<(h)$. Therefore, R is a Gröbner basis for $\langle \chi(F_1), \ldots, \chi(F_k) \rangle$. □

Theorem 7.3.12 leads to an algorithm (Alg. 9) for computing Gröbner bases through a Macaulay matrix and Gaussian elimination.

7.3.3 Exploiting the structure of Macaulay matrices (`Koszul-F5` criterion)

If we consider all the polynomials of the set in Eq. (7.6), then many of them are linearly dependent. Hence, when we construct the Macaulay matrix of Thm. 7.3.12 and perform Gaussian elimination, many

of the rows reduce to zero; this forces Alg. 9 to perform unnecessary computations. We will extend to `F5` criterion (Prop. 4.4.6) to our setting to avoid these redundant computations.

Theorem 7.3.13 (`Koszul-F5` criterion)**.** *Consider homogeneous polynomials $F_1, \dots, F_k \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ of multidegrees $\boldsymbol{d}_1, \dots, \boldsymbol{d}_k$ and a multidegree $\boldsymbol{d} \in \mathbb{N}^r$ such that $\boldsymbol{d} \geq \boldsymbol{d}_k$, that is coordinate-wise greater than or equal to $\boldsymbol{d}_k$. Let $\mathcal{M}^{k-1}_{\boldsymbol{d}}$ and $\mathcal{M}^{k-1}_{\boldsymbol{d}-\boldsymbol{d}_k}$ be the Macaulay matrices of degrees $\boldsymbol{d}$ and $\boldsymbol{d}-\boldsymbol{d}_k$, respectively, of the polynomials $F_1, \dots, F_{k-1}$ as in Thm. 7.3.12, and let $\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}}$ and $\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}-\boldsymbol{d}_k}$ be their reduced row echelon forms.*

For every $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d}-\boldsymbol{d}_k)} \in \mathtt{LM}_<(\mathtt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}-\boldsymbol{d}_k}))$, the polynomial $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d}-\boldsymbol{d}_k)}F_k$ is a linear combination of the polynomials

$$\mathtt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}}) \cup \left\{ \boldsymbol{X}^{(\boldsymbol{\beta},\boldsymbol{d}-\boldsymbol{d}_k)}\, F_k : \begin{array}{l} \boldsymbol{X}^{(\boldsymbol{\beta},\boldsymbol{d}-\boldsymbol{d}_k)} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}-\boldsymbol{d}_k} \text{ and} \\ \boldsymbol{X}^{(\boldsymbol{\beta},\boldsymbol{d}-\boldsymbol{d}_k)} < \boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d}-\boldsymbol{d}_k)} \end{array} \right\}.$$

Proof. If $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d}-\boldsymbol{d}_k)} \in \mathtt{LM}_<(\mathtt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}-\boldsymbol{d}_k}))$, then there is $G \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}-\boldsymbol{d}_k}$ such that $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d}-\boldsymbol{d}_k)} + G \in \langle F_1, \dots, F_{k-1}\rangle_{\boldsymbol{d}-\boldsymbol{d}_k}$ and $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d}-\boldsymbol{d}_k)} > \mathtt{LM}_<(G)$. So, there are homogeneous $H_1, \dots, H_{k-1} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ such that $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d}-\boldsymbol{d}_k)} + G = \sum_i H_i F_i$. The proof follows by noticing that $\boldsymbol{X}^{(\boldsymbol{\alpha},\boldsymbol{d}-\boldsymbol{d}_k)}F_k = \sum_{i=1}^{k-1}(F_k\, H_i)F_i - G\, F_k$. □

In the following, $\mathcal{M}^k_{\boldsymbol{d}}$ is not the Macaulay matrix of Lem. 7.3.11. It contains less rows because of the `Koszul-F5` criterion. However, both matrices have the same row space, so we use the same name.

Corollary 7.3.14. *Using the notation of Thm. 7.3.13, let $\mathcal{M}^k_{\boldsymbol{d}}$ be a Macaulay matrix of degree $\boldsymbol{d}$ with respect to the order $<$ whose rows are*

$$\mathtt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}}) \cup \left\{ \boldsymbol{X}^{(\boldsymbol{\beta},\boldsymbol{d}-\boldsymbol{d}_k)}\, F_k : \begin{array}{l} \boldsymbol{X}^{(\boldsymbol{\beta},\boldsymbol{d}-\boldsymbol{d}_k)} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}-\boldsymbol{d}_k} \text{ and} \\ \boldsymbol{X}^{(\boldsymbol{\beta},\boldsymbol{d}-\boldsymbol{d}_k)} \notin \mathtt{LM}_<(\mathtt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}-\boldsymbol{d}_k})) \end{array} \right\}$$

The row space of $\mathcal{M}^k_{\boldsymbol{d}}$ and the Macaulay matrix of Lem. 7.3.11 are equal.

The correctness of Alg. 10 follows from Thm. 7.3.13.

Lemma 7.3.15. *If $\mathcal{H}_1(F_1, \dots, F_k)_{\boldsymbol{d}} = 0$ and there is a syzygy $\sum_i G_i F_i = 0$ such that $G_i \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}-\boldsymbol{d}_i}$, then $G_k \in \langle F_1, \dots, F_{k-1}\rangle_{\boldsymbol{d}-\boldsymbol{d}_k}$.*

Proof. We consider the Koszul complex $\mathcal{K}(F_1, \dots, F_k)$ (Def. 7.1.6). As $\sum_i G_i F_i = \delta_1(G_1, \dots, G_k)$, the vector of polynomials $(G_1, \dots, G_k)$ belongs to the Kernel of δ_1. As $\mathcal{H}_1(F_1, \dots, F_k)_{\boldsymbol{d}}$ vanishes, the kernel of δ_1 is generated by the image of δ_2. The latter map is

$$(H_{1,2}, \dots, H_{k-1,k}) \mapsto \sum\nolimits_{1\leq i<j\leq k} H_{i,j}(F_j \boldsymbol{e}_i - F_i \boldsymbol{e}_j),$$

where $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ are canonical bases of $\mathbb{R}^k$. Hence, there are homogeneous polynomials $(H_{1,2}, \dots, H_{k-1,k})$ such that

$$(G_1, \dots, G_k) = \sum\nolimits_{1\leq i<j\leq k} H_{i,j}(F_j \boldsymbol{e}_i - F_i \boldsymbol{e}_j).$$

Thus, $G_k = \sum_{i=1}^{k-1} H_{i,k}F_i$ and so $G_k \in \langle F_1, \dots, F_{k-1}\rangle_{\boldsymbol{d}-\boldsymbol{d}_k}$. □

Algorithm 10 `ReduceMacaulay`

Input: A list of homogeneous polynomials $F_1, \dots, F_k \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ of multidegree $\boldsymbol{d}_1, \dots, \boldsymbol{d}_k$, a multidegree $\boldsymbol{d}$, and a multigraded monomial order $<$.
Output: The Macaulay matrix of $\langle F_1, \dots, F_k \rangle_{\boldsymbol{d}} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ with respect to $<$ in row echelon form.

$\mathcal{M}^k_{\boldsymbol{d}} \leftarrow$ Macaulay matrix of degree $\boldsymbol{d}$ with respect to $<$.
if $k > 1$ **then**
  $\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}} \leftarrow \texttt{ReduceMacaulay}(\{F_1, \dots, F_{k-1}\}, \boldsymbol{d}, <)$.
  $\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}-\boldsymbol{d}_k} \leftarrow \texttt{ReduceMacaulay}(\{F_1, \dots, F_{k-1}\}, \boldsymbol{d}-\boldsymbol{d}_k, <)$.
  for $F \in \texttt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}})$ **do**
    Add the polynomial F as a row to $\mathcal{M}^k_{\boldsymbol{d}}$.
  end for
end if
for $\boldsymbol{X}^{(\boldsymbol{\alpha}, \boldsymbol{d}-\boldsymbol{d}_k)} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}-\boldsymbol{d}_k} \setminus \texttt{LM}_<(\texttt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}-\boldsymbol{d}_k}))$ **do**
  Add the polynomial $\boldsymbol{X}^{(\boldsymbol{\alpha}, \boldsymbol{d}-\boldsymbol{d}_k)} F_k$ as a row to $\mathcal{M}^k_{\boldsymbol{d}}$.
end for
$\widetilde{\mathcal{M}}^k_{\boldsymbol{d}} \leftarrow \texttt{GaussianElimination}(\mathcal{M}^k_{\boldsymbol{d}})$.
return $\widetilde{\mathcal{M}}^k_{\boldsymbol{d}}$.

The next lemma shows that, under (sparse) regularity assumptions (Def. 7.1.7), we avoid all redundant computations, that is all the rows reducing to zero during Gaussian elimination.

Lemma 7.3.16. *If $\mathcal{H}_1(F_1, \dots, F_k)_{\boldsymbol{d}} = 0$, then all the rows of the matrix $\mathcal{M}^k_{\boldsymbol{d}}$ in Alg. 10 are linearly independent.*

Proof. By construction, the rows of $\mathcal{M}^k_{\boldsymbol{d}}$ corresponding to $\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}}$ are linearly independent because the matrix is in row echelon form. Hence, if there are rows that are not linearly independent, then at least one of them corresponds to a polynomial of the form $\boldsymbol{X}^{(\boldsymbol{\alpha}, \boldsymbol{d}-\boldsymbol{d}_k)} F_k$. The right action of the Macaulay matrix $\mathcal{M}^k_{\boldsymbol{d}}$ represents a map equivalent to the map δ_1 from the strand of Koszul complex $\mathcal{K}(F_1, \dots, F_k)_{\boldsymbol{d}}$. Hence, if some of the rows of the matrix are linearly dependent, then there is an element in the kernel of δ_1. That is, there are $G_i \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}-\boldsymbol{d}_i}$ such that

- $\sum_{i=1}^{k-1} G_i F_i$ belongs to the linear span of $\texttt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}})$,
- the monomials of G_k do not belong to $\texttt{LM}_<(\texttt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}-\boldsymbol{d}_k}))$, and
- $\sum_{i=1}^{k} G_i F_i = 0$.

By Lem. 7.3.15, $G_k \in \langle F_1, \dots, F_{k-1} \rangle_{\boldsymbol{d}-\boldsymbol{d}_k}$. By Lem. 7.3.11 and Cor. 7.3.14, the leading monomials of $\texttt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}-\boldsymbol{d}_k})$ and the ideal $\langle F_1, \dots, F_{k-1} \rangle$ at degree $\boldsymbol{d} - \boldsymbol{d}_k$ are the same. Hence, we reach a contradiction because we have assumed that the leading monomial of G_k does not belong to $\texttt{LM}_<(\texttt{Rows}(\widetilde{\mathcal{M}}^{k-1}_{\boldsymbol{d}-\boldsymbol{d}_k}))$. □

Corollary 7.3.17. *If $F_1, \dots, F_k$ is a (sparse) regular sequence (Def. 7.1.7) and $\boldsymbol{d} \in \mathbb{N}^r$ is such that $\boldsymbol{d} \geq (\sum_i \boldsymbol{d}_i)$, then* `ReduceMacaulay`$(F_1, \dots, F_k, \boldsymbol{d}, <)$ *only considers matrices with linearly independent rows and avoids all redundant computations.*

To benefit from the `Koszul-F5` criterion and compute with smaller matrices during the Gröbner basis computation we should replace Lines 4 – 8 in Alg. 9 by `ReduceMacaulay`$(F_1, \ldots, F_k, \boldsymbol{d}, <)$ (Alg. 10).

Chapter 8

Solving sparse polynomial systems

We can use both *scant Gröbner basis* and *Gröbner basis over semigroup algebras* to compute normal forms and so, to solve sparse zero-dimensional system. In Chapter 8, we introduce complexity bounds for solving sparse polynomial systems using our algorithm to compute *Gröbner basis over semigroup algebras*. We do not discuss how to solve sparse polynomial systems using *scant Gröbner basis*; nevertheless, we can deduce straightforwardly an algorithm from the variant of the `FGLM` algorithm (Alg. 5) proposed in [FSS16, Sec. 4.2]. Unfortunately, we have not bounds for the arithmetic complexity of this approach, let alone bounds depending on the Newton polytopes.

We build on [ER94, Emi96, Mas16] and, under some assumptions (Ass. 8.1.1), we propose an algorithm to solve zero-dimensional square systems. Because we work with toric varieties, we only compute the solutions lying in $(\mathbb{C} \setminus \{0\})^n$. The arithmetic complexity of our algorithm is polynomial in the number of integer points in the Minkowski sum of the Newton polytopes. Our strategy is to reuse part of our algorithm to compute *Gröbner bases over semigroup algebras* (Sec. 7.3) to compute multiplication maps and, via `FGLM` [FGLM93], recover a Gröbner basis over the standard polynomial algebra $\mathbb{K}[\boldsymbol{x}]$. As we compute the solutions over $(\mathbb{C} \setminus \{0\})^n$, we do not recover a Gröbner basis for the original ideal, but for its saturation with respect to the product of all the variables. Our algorithm relies on ideas from resultant theory to avoid the computation of a Gröbner basis. Instead, we compute a part of the Gröbner basis which suffices to solve the systems. We compute with a matrix that has the same size as the one in resultant-based approaches [ER94, Emi96]. Hence, the complexity of our algorithm is similar to the one of the resultant-based approaches; however, we do not compute a mixed subdivision and we rely on weaker assumptions which, in addition, are geometric. Moreover, in general, Gröbner-type algorithms can be extended without any modification to the overdetermined case.

We introduce an algorithm to solve *mixed sparse systems* (Sec. 8.1) and two variants for two specific subfamilies: *mixed multihomogeneous systems* (Sec. 8.2) and *unmixed systems* (Sec. 8.3).

8.1 Mixed sparse polynomial systems

We introduce an algorithm that takes as input a zero-dimensional mixed sparse polynomial system $f_1, \ldots, f_n$ in $\mathbb{K}[\mathbb{Z}^n]$ and solves it by computing a lexicographical Gröbner basis for the ideal $\left(\langle f_1, \ldots, f_n\rangle : \langle \prod_j x_j\rangle^\infty\right) \subset \mathbb{K}[\boldsymbol{x}]$. The latter corresponds to the ideal associated to the intersection

of the torus $(\mathbb{C}^*)^n$ with the variety defined by I. For this, we reuse our results on computing Gröbner basis over semigroup algebras. We follow the notation of Sec. 7.3.

8.1.1 The algorithm

Let $f_1, \dots, f_n \in \mathbb{K}[\boldsymbol{x}]$ be a square zero-dimensional system. First we embed each f_i in $\mathbb{K}[\mathbb{Z}^n]$. We multiply each polynomial by an appropriate monomial, $\boldsymbol{X}^{\boldsymbol{\beta}_i} \in \mathbb{K}[\mathbb{Z}^n]$, so that $\mathbf{0}$ is a vertex for each Newton polytope, as well as, a vertex of their Minkowski sum. For each $1 \leq i \leq n$, we consider the Newton polytopes:

$$\Delta_i := \mathrm{NP}(\boldsymbol{X}^{\boldsymbol{\beta}_i} f_i). \tag{8.1}$$

Let Δ_0 be the standard n-simplex (Ex. 2.12.15); it corresponds to the Newton polytope $\mathrm{NP}(1 + \sum_i x_i)$. We consider the algebras $\mathbb{K}[S_\Delta]$ and $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ associated to the polytopes $\Delta_0, \dots, \Delta_n$, see Def. 7.1.1, and the embedding $\boldsymbol{X}^{\boldsymbol{\beta}_1} f_1, \dots, \boldsymbol{X}^{\boldsymbol{\beta}_n} f_n \in \mathbb{K}[S_\Delta]$. For each $i \geq 1$, we consider $F_i \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{e}_i}$ such that $\chi(F_i) = \boldsymbol{X}^{\boldsymbol{\beta}_i} f_i \in \mathbb{K}[S_\Delta]$.

Assumption 8.1.1. *Using the notation of Sec. 7.3.1, let X be the projective toric variety associated to $\Delta_0 + \cdots + \Delta_n$. Assume that the system $(f_1, \dots, f_n)$ has no solutions at infinity with respect to X (Def. 7.3.4). Further, assume that, for a generic linear polynomial f_0, the system $(f_0, f_1, \dots, f_n)$ has no solutions over $(\mathbb{C}^*)^n$.*

In what follows, given $\boldsymbol{d}_1, \boldsymbol{d}_2 \in \mathbb{N}^{n+1}$, we say $\boldsymbol{d}_1 \geq \boldsymbol{d}_2$ if $\boldsymbol{d}_1 - \boldsymbol{d}_2 \in \mathbb{N}^{n+1}$.

Lemma 8.1.2. *[Mas16, Thm. 3.a] Under Ass. 8.1.1, for every $\boldsymbol{d} \in \mathbb{N}^{n+1}$ such that $\boldsymbol{d} \geq \sum_{i>0} \boldsymbol{e}_i$, it holds $\mathcal{H}_0(F_1, \dots, F_n)_{\boldsymbol{d}} \cong \mathbb{K}[\mathbb{Z}^n]/\langle f_1, \dots, f_n \rangle$.*

Lemma 8.1.3. *[Mas16, Thm. 3.c] Under Ass. 8.1.1, for every homogeneous polynomial $F_0 \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}_0}$ such that the system $(f_1, \dots, f_n, \chi(F_0))$ has no solutions over $(\mathbb{C}^*)^n$, the system $(F_1, \dots, F_n, F_0)$ is Koszul regular (Def. 7.1.7) and, for every $\boldsymbol{d} \in \mathbb{N}^{n+1}$ such that $\boldsymbol{d} \geq \sum_i \boldsymbol{e}_i + \boldsymbol{d}_0$, $\langle F_1, \dots, F_n, F_0 \rangle_{\boldsymbol{d}} = \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$.*

Proof. The key observation for the proof is that the homogenization of system $(f_1, \dots, f_n, \chi(F_0))$ with respect to the toric variety X has no solutions over X (see the discussion before Def. 7.3.4). To see this, notice that, by Ass. 8.1.1, the homogenization of the system $(f_1, \dots, f_n)$ with respect to X has no solutions over $X \setminus (\mathbb{C}^*)^n$ (see also Def. 7.3.4). Moreover, we assumed that $(f_1, \dots, f_n, \chi(F_0))$ has no solutions over $(\mathbb{C}^*)^n$.

The proof follows from the argument in the proof of [Mas16, Thm. 3]. This argument is the same as in [GKZ08, Prop. 3.4.1], where the stably twisted condition is given by [Mas16, Thm. 1]. □

Corollary 8.1.4. *Under Ass. 8.1.1, for any monomial $\boldsymbol{X}^{(\boldsymbol{\alpha}, D\,\boldsymbol{e}_0)} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{D\,\boldsymbol{e}_0}$, the system $(F_1, \dots, F_n, \boldsymbol{X}^{(\boldsymbol{\alpha}, D\,\boldsymbol{e}_0)})$ is Koszul regular. For every $\boldsymbol{d} \in \mathbb{N}^{n+1}$ such that $\boldsymbol{d} \geq \sum_i \boldsymbol{e}_i + D\,\boldsymbol{e}_0$, it holds $\langle F_1 \dots F_n, \boldsymbol{X}^{(\boldsymbol{\alpha}, D\,\boldsymbol{e}_0)} \rangle_{\boldsymbol{d}} = \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$.*

We fix a graded monomial order $>$ for $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ (Def. 7.3.1). Let $\mathfrak{b}$ be the set of monomials that are not leading monomials of $\langle F_1, \dots, F_n \rangle_{\sum_{i\geq 1} \boldsymbol{e}_i}$, that is

$$\mathfrak{b} := \left\{ \boldsymbol{X}^{(\boldsymbol{\alpha}, \sum_{i\geq 1} \boldsymbol{e}_i)} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\sum_{i\geq 1} \boldsymbol{e}_i} : \left(\forall G \in \langle F_1, \dots, F_n \rangle_{\sum_{i\geq 1} \boldsymbol{e}_i} \right) \mathtt{LM}_>(G) \neq \boldsymbol{X}^{(\boldsymbol{\alpha}, \sum_{i\geq 1} \boldsymbol{e}_i)} \right\} \tag{8.2}$$

We prove that the dehomogenization of these monomials, $\chi(\mathfrak{b})$, forms a monomial basis for $\mathbb{K}[\mathbb{Z}^n]/\langle f_1,\ldots,f_n\rangle$.

Lemma 8.1.5. *The monomials in the set $\chi(\mathfrak{b})$ are $\mathbb{K}$-linearly independent in $\mathbb{K}[\mathbb{Z}^n]/\langle f_1,\ldots,f_n\rangle$.*

Proof. Assume that the lemma does not hold. Hence, there are $c_1,\ldots,c_{(\#\mathfrak{b})}\in\mathbb{K}$, not all of them 0, and $g_1,\ldots,g_n\in\mathbb{K}[\mathbb{Z}^n]$ such that $\sum_i c_i\,\chi(\mathfrak{b}_i)=\sum_i g_i\,f_i$. We can clear the denominators, introduced by the g_i's, by choosing a monomial $\boldsymbol{X}^{\boldsymbol{\alpha}}\in\mathbb{K}[\mathbb{N}^n]$ such that, for every i, $\left(\frac{\boldsymbol{X}^{\boldsymbol{\alpha}}}{\boldsymbol{X}^{\boldsymbol{\beta}_i}}\,g_i\right)\in\mathbb{K}[\mathbb{N}^n]$. Moreover, there is a degree $D\in\mathbb{N}$ and homogeneous polynomials $G_i\in\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ of multidegrees $(D\,\boldsymbol{e}_0+\sum_{j>0}\boldsymbol{e}_j-\boldsymbol{e}_i)$ such that $\chi(G_i)=\left(\frac{\boldsymbol{X}^{\boldsymbol{\alpha}}}{\boldsymbol{X}^{\boldsymbol{\beta}_i}}\,g_i\right)$ and $\boldsymbol{X}^{(\boldsymbol{\alpha},D\,\boldsymbol{e}_0)}\sum_i c_i\,\mathfrak{b}_i=\sum_i G_i\,F_i$. By Lem. 8.1.3, $(F_1,\ldots,F_n,\boldsymbol{X}^{(\boldsymbol{\alpha},D\,\boldsymbol{e}_0)})$ is Koszul regular and so, by Lem. 7.3.15, $\sum_i c_i\,\mathfrak{b}_i\in\langle F_1,\ldots,F_n\rangle_{\sum_{i>1}\boldsymbol{e}_i}$. So, a monomial in $\mathfrak{b}$ is a leading monomial of an element in $\langle F_1,\ldots,F_n\rangle_{\sum_{i>1}\boldsymbol{e}_i}$. This is a contradiction as, by construction, there is no monomial in $\mathfrak{b}$ which is a leading monomial of a polynomial in $\langle F_1,\ldots,F_n\rangle_{\sum_{i>1}\boldsymbol{e}_i}$. □

Corollary 8.1.6. *The set of monomials $\chi(\mathfrak{b})$ is a monomial basis of $\mathbb{K}[\mathbb{Z}^n]/\langle f_1,\ldots,f_n\rangle$.*

Proof. By Lem. 8.1.2, the number of elements in the set $\mathfrak{b}$ and the dimension of $\mathbb{K}[\mathbb{Z}^n]/\langle f_1,\ldots,f_n\rangle$ is the same. By Obs. 7.1.4, as all the monomials in $\mathfrak{b}$ have the same degree, the sets $\mathfrak{b}$ and $\chi(\mathfrak{b})$ have the same number of elements. By Lem. 8.1.5, the monomials in the set $\chi(\mathfrak{b})$ are linearly independent. □

Remark 8.1.7. *A way to compute the set $\mathfrak{b}$ is to compute a basis of the vector space $\langle F_1,\ldots,F_n\rangle_{\sum_{i\geq 1}\boldsymbol{e}_i}$ using Alg. 10, that is* `ReduceMacaulay` $\left((F_1,\ldots,F_n),\sum_{i\geq 1}\boldsymbol{e}_i,>\right)$.

For each $F_0\in\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{e}_0}$, we construct a Macaulay matrix (Def. 7.3.7) at multidegree $\mathbf{1}:=\sum_{i=0}^n\boldsymbol{e}_i$, say $\mathcal{M}(F_0)$, related to the ideal $\langle F_1,\ldots,F_n,F_0\rangle$. From this matrix we will recover the multiplication map of $\chi(F_0)$ in $\mathbb{K}[\boldsymbol{x}]/\langle f_1,\ldots,f_n\rangle$. The rows of $\mathcal{M}(F_0)$ correspond to polynomials of two kinds:

- the polynomials in $\texttt{Rows}(\widetilde{\mathcal{M}}^n_{\mathbf{1}})$, where $\widetilde{\mathcal{M}}^n_{\mathbf{1}}=\texttt{ReduceMacaulay}\,((F_1,\ldots,F_n),\mathbf{1},>)$, and
- the polynomials of the form $\mathfrak{b}_i\,F_0$, where $\mathfrak{b}_i\in\mathfrak{b}$.

Lemma 8.1.8. *The matrix $\mathcal{M}(F_0)$ is always square. It is full-rank if and only if $(F_1,\ldots,F_N,F_0)$ is Koszul regular.*

Proof. According to the `Koszul-F5` criterion (see Thm. 7.3.13), the row space spanned by $\mathcal{M}(F_0)$ is the same as the vector space $\langle F_1,\ldots,F_n,F_0\rangle_{\mathbf{1}}$ for any choice of $F_0\in\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{e}_0}$. By Cor. 8.1.4, we can consider an $F_0\in\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{e}_0}$ such that $(F_1,\ldots,F_n,F_0)$ is Koszul regular. Then, the rows of $\mathcal{M}(F_0)$ generate $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\mathbf{1}}$ and are linearly independent (Lem. 7.3.16). Hence, by Lem. 8.1.3, for this particular F_0, the matrix $\mathcal{M}(F_0)$ is square and full-rank. However, the matrix $\mathcal{M}(F_0)$ is square for any choice of $F_0\in\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{e}_0}$, because its number of rows does not depend on F_0. Nevertheless, it is not full-rank for any choice of $F_0\in\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{e}_0}$. If $\mathcal{M}(F_0)$ is full-rank, then $(F_1,\ldots,F_n,F_0)$ is Koszul regular because, by the sparse Nullstellensatz [Som99, Thm. 2], the homogenization of the system $(f_1,\ldots,f_n,\chi(F_0))$ has no solutions over $(\mathbb{C}^*)^n$. Consequently, the proof follows from Lem. 8.1.3. □

We reorder the columns of $\mathcal{M}(F_0)$ as shown in Eq. (8.3), such that

- the columns of the submatrix $\left[\begin{smallmatrix} M_{1,2}(F_0) \\ M_{2,2}(F_0) \end{smallmatrix}\right]$ correspond to monomials of the form $\mathfrak{b}_i\, \boldsymbol{X}^{(\mathbf{0},\boldsymbol{e}_0)}$, where $\mathfrak{b}_i \in \mathfrak{b}$, and
- the rows of $\left[\, M_{2,1}(F_0) \mid M_{2,2}(F_0) \,\right]$ are polynomials of the form $\mathfrak{b}_i\, F_0$, where $\mathfrak{b}_i \in \mathfrak{b}$.

$$\mathcal{M}(F_0) = \begin{array}{r} \\ \texttt{Rows}(\widetilde{\mathcal{M}}_{\mathbf{1}}^n) \\ \\ F_0 \cdot \mathfrak{b} \end{array} \overbrace{\left\{ \begin{array}{c} \\ \\ \\ \\ \end{array} \right.}^{} \left[\begin{array}{c|c} & \overbrace{\qquad\qquad}^{\boldsymbol{X}^{(\mathbf{0},\boldsymbol{e}_0)} \cdot \mathfrak{b}} \\ M_{1,1}(F_0) & M_{1,2}(F_0) \\ \hline\hline M_{2,1}(F_0) & M_{2,2}(F_0) \end{array}\right] \tag{8.3}$$

We prove that $M_{1,1}(F_0)$ is invertible and the Schur complement of $M_{2,2}(F_0)$, $M_{2,2}^c(F_0)$, is the multiplication map of $\chi(F_0)$ in the basis $\chi(\mathfrak{b})$ of $\mathbb{K}[\mathbb{Z}^n]/\langle f_1, \ldots, f_n\rangle$.

$$M_{2,2}^c(F_0) := (M_{2,2} - M_{2,1}\, M_{1,1}^{-1}\, M_{1,2})(F_0). \tag{8.4}$$

Lemma 8.1.9. *If $(F_1, \ldots, F_n, \boldsymbol{X}^{(\mathbf{0},\boldsymbol{e}_0)})$ is Koszul regular then, for any $F_0 \in \mathbb{K}[S_{\boldsymbol{\Delta}}^h]_{\boldsymbol{e}_0}$, the matrix $M_{1,1}(F_0)$ is invertible.*

Proof. By Lem. 8.1.8, as the system $(F_1, \ldots, F_n, \boldsymbol{X}^{(\mathbf{0},\boldsymbol{e}_0)})$ is Koszul regular, then the matrix $\mathcal{M}(\boldsymbol{X}^{(\mathbf{0},\boldsymbol{e}_0)})$ is invertible. As $M_{2,1}(\boldsymbol{X}^{(\mathbf{0},\boldsymbol{e}_0)})$ is the zero matrix and $M_{2,2}(\boldsymbol{X}^{(\mathbf{0},\boldsymbol{e}_0)})$ is the identity, then $M_{1,1}(\boldsymbol{X}^{(\mathbf{0},\boldsymbol{e}_0)})$ must be invertible. By construction, the matrices $M_{1,1}(F_0)$ and $M_{1,2}(F_0)$ are independent of the choice of F_0. Hence, for any F_0, the matrix $M_{1,1}(F_0)$ is invertible. □

Theorem 8.1.10. *The multiplication map of $\chi(F_0)$ in the monomial basis $\chi(\mathfrak{b})$ of $\mathbb{K}[\mathbb{Z}^n]/\langle f_1, \ldots, f_n\rangle$ is $M_{2,2}^c(F_0)$, that is the Schur complement of $M_{2,2}(F_0)$ (Eq. 8.4).*

Proof. Note that for every $F_0 \in \mathbb{K}[S_{\boldsymbol{\Delta}}^h]_{\boldsymbol{e}_0}$ and each element $\mathfrak{b}_i \in \mathfrak{b}$,

$$\mathfrak{b}_i\, F_0 \equiv \boldsymbol{X}^{(\mathbf{0},\boldsymbol{e}_0)} \sum_j (M_{2,2}^c(F_0))_{i,j}\, \mathfrak{b}_j \quad \text{in} \quad \mathbb{K}[S_{\boldsymbol{\Delta}}^h]/\langle F_1, \ldots, F_n\rangle,$$

where $(M_{2,2}^c(F_0))_{i,j}$ is the (i,j) element of the matrix $M_{2,2}^c(F_0)$. Hence, if we dehomogenize this relation we obtain that,

$$\chi(\mathfrak{b}_i)\chi(F_0) \equiv \sum_j (M_{2,2}^c(F_0))_{i,j}\chi(\mathfrak{b}_j) \quad \text{in} \quad \mathbb{K}[S_\Delta]/\langle \boldsymbol{X}^{\boldsymbol{\beta}_1} f_1, \ldots, \boldsymbol{X}^{\boldsymbol{\beta}_n} f_n\rangle.$$

As $\mathbb{K}[S_\Delta] \subset \mathbb{K}[\mathbb{Z}^n]$, the same relation holds in $\mathbb{K}[\mathbb{Z}^n]/\langle f_1, \ldots, f_n\rangle$. By Cor. 8.1.6, the set $\chi(\mathfrak{b})$ is a monomial basis of $\mathbb{K}[\mathbb{Z}^n]/\langle f_1, \ldots, f_n\rangle$. Therefore, $M_{2,2}^c(F_0)$ is the multiplication map of $\chi(F_0)$ in $\mathbb{K}[\mathbb{Z}^n]/\langle f_1, \ldots, f_n\rangle$. □

Using the multiplication maps in $\mathbb{K}[\mathbb{Z}^n]/\langle f_1,\dots,f_n\rangle$ and the FGLM algorithm (Alg. 5), we can compute a Gröbner basis for $\langle f_1,\dots,f_n\rangle : \langle\prod_i x_i\rangle^\infty$ over $\mathbb{K}[\boldsymbol{x}]$. The latter is the saturation over $\mathbb{K}[\mathbb{N}^n]$ of the ideal $\langle f_1,\dots,f_n\rangle$ by the product of all the variables.

Lemma 8.1.11. *Consider polynomials $f_1,\dots,f_k \subset \mathbb{K}[\mathbb{N}^n]$. We denote by $\langle f_1,\dots,f_k\rangle_{\mathbb{K}[\mathbb{Z}^n]}$ the ideal that $f_1,\dots,f_k$ generate over $\mathbb{K}[\mathbb{Z}^n]$ and by $\langle f_1,\dots,f_k\rangle_{\mathbb{K}[\mathbb{N}^n]}$ the ideal that they generate over $\mathbb{K}[\mathbb{N}^n]$. Then, the sets $\langle f_1,\dots,f_k\rangle_{\mathbb{K}[\mathbb{Z}^n]} \cap \mathbb{K}[\mathbb{N}^n]$ and $\langle f_1,\dots,f_k\rangle_{\mathbb{K}[\mathbb{N}^n]} : \langle\prod_i x_i\rangle^\infty$ are the same. The latter is an ideal over $\mathbb{K}[\mathbb{N}^n]$.*

Proof. Consider $f \in \langle f_1,\dots,f_k\rangle_{\mathbb{K}[\mathbb{Z}^n]} \cap \mathbb{K}[\mathbb{N}^n]$. Then there are $g_i \in \mathbb{K}[\mathbb{Z}^n]$ such that $f = \sum_i g_i f_i$. We can clear the denominators introduced by the g_i's by multiplying both sides by a monomial $(\prod_j x_j)^d$, where d is big enough. Then, $(\prod_j x_j)^d f = \sum_i((\prod_j x_j)^d g_i) f_i$ and $((\prod_j x_j)^d g_i) \in \mathbb{K}[\mathbb{N}^n]$. Thus, $(\prod_j x_j)^d f \in \langle f_1,\dots,f_k\rangle_{\mathbb{K}[\mathbb{N}^n]}$ and $f \in \langle f_1,\dots,f_k\rangle_{\mathbb{K}[\mathbb{N}^n]} : \langle\prod_j x_j\rangle^\infty$. The opposite direction is straightforward as $\prod_i x_i$ is a unit in $\mathbb{K}[\mathbb{Z}^n]$. □

We can perform FGLM over $\mathbb{K}[\mathbb{N}^n]$ to recover a Gröbner basis for $\langle f_1,\dots,f_n\rangle_{\mathbb{K}[\mathbb{N}^n]} : \langle\prod_i x_i\rangle^\infty$ by considering the multiplication maps of each x_i. These correspond to $M^c_{2,2}(\boldsymbol{X}^{(\boldsymbol{\alpha_i},\boldsymbol{e}_0)})$, where $\boldsymbol{\alpha_i}$ is such that $\chi(\boldsymbol{X}^{(\boldsymbol{\alpha}_i,\boldsymbol{e}_0)}) = x_i$. We note that, as $>$ is a multigraded monomial order, it holds $\boldsymbol{X}^{(\mathbf{0},\sum_{i\geq 1}\boldsymbol{e}_i)} \in \mathfrak{b}$ and so, $\chi(\boldsymbol{X}^{(\mathbf{0},\sum_{i\geq 1})}) = 1 \in \chi(\mathfrak{b})$. We omit the details of this procedure.

The following algorithm, Alg. 11, summarizes our strategy to compute Gröbner basis for zero-dimensional systems over the standard algebra.

8.1.2 Complexity

We estimate the arithmetic complexity of Algorithm 11; it is polynomial with respect to the Minkowski sum of the polytopes. We omit the cost of computing all the monomials in $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$ and we only consider the complexity of reading them. Our purpose is to highlight the dependency on the Newton polytopes. A more detailed analysis might give sharper bounds.

Definition 8.1.12. *For polytopes $\Delta_0,\dots,\Delta_n$ and for each multidegree $\boldsymbol{d} = (d_0,\dots,d_n) \in \mathbb{N}^{n+1}$ of $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$, let $P(\boldsymbol{d})$ be the number of integer points in the Minkowski sum of the polytopes given by $\boldsymbol{d}$,*

$$P(\boldsymbol{d}) = \#\Big((\textstyle\sum_{j=0}^n d_j\Delta_j) \cap \mathbb{Z}^n\Big).$$

Note that $P(\boldsymbol{d})$ equals the number of different monomials in $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{d}}$.

Lemma 8.1.13. *Let $F_1,\dots,F_k \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ be a* (sparse) regular sequence *of multidegrees $\boldsymbol{d}_1,\dots,\boldsymbol{d}_k \in \mathbb{N}^{n+1}$, respectively. Consider a multigraded monomial order $>$. For every multidegree $\boldsymbol{d} \in \mathbb{N}^{n+1}$ such that $\boldsymbol{d} \geq \sum_i \boldsymbol{d}_i$, the arithmetic complexity of computing* ReduceMacaulay $((F_1,\dots,F_k),\boldsymbol{d},>)$ *is $\mathcal{O}(2^{k+1}\,P(\boldsymbol{d})^\omega)$, where ω is the* exponent of matrix multiplication *[GG13, Sec. 12.1].*

Proof. By Cor. 7.3.17, as $F_1,\dots,F_k \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ is a (sparse) regular sequence and $\boldsymbol{d} \geq \sum_i \boldsymbol{d}_i$, all the matrices that appear during the computations of ReduceMacaulay $((F_1,\dots,F_k),\boldsymbol{d},>)$ are full-rank and their rows are linearly independent. Hence, for each matrix, their number of rows is at most their number of columns. The number of columns of a Macaulay matrix of multidegree $\boldsymbol{d}$ is $P(\boldsymbol{d})$.

Algorithm 11 `compute-0-Dim-GB`

Input: Affine system $(f_1, \ldots, f_n)$ in $\mathbb{K}[\boldsymbol{x}]$ and a monomial order $>$ for $\mathbb{K}[\boldsymbol{x}]$, such that it has a finite number of solutions over $(\mathbb{C}^*)^n$ and satisfies Ass. 8.1.1.

Output: Gröbner basis G for the ideal $\langle f_1, \ldots, f_n\rangle : \langle \prod_i x_i\rangle^\infty$ with respect to the monomial order $>$.

Consider the algebra $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ related to the polytopes of $f_1, \ldots, f_n$ and the n-simplex. (See Eq. 8.1)
For each $i \in [n]$ **do** choose $F_i \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{e}_i}$ such that $\chi(F_i) = \boldsymbol{X}^{\boldsymbol{\beta}_i} f_i$. (See Eq. 8.1)
Choose a multigraded monomial order $\bar{>}$ for $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$. (see Def. 7.3.1)
$C \leftarrow$ `Rows`(`ReduceMacaulay` $\left((F_1, \ldots, F_n), \sum_{i>0} \boldsymbol{e}_i, \bar{>}\right)$). (see Alg. 10)
$\mathfrak{b} \leftarrow \left\{ \boldsymbol{X}^{(\boldsymbol{\alpha}, \sum_{i\geq 1} \boldsymbol{e}_i)} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\sum_{i\geq 1} \boldsymbol{e}_i} : \boldsymbol{X}^{(\boldsymbol{\alpha}, \sum_{i\geq 1} \boldsymbol{e}_i)} \notin C \right\}$ (see Eq. 8.2)
$P \leftarrow$ `Rows`(`ReduceMacaulay` $((F_1, \ldots, F_n), \mathbf{1}, \bar{>})$). (see Alg. 10)
for all $x_i \in \mathbb{K}[\boldsymbol{x}]$ **do**
  Choose monomial $\boldsymbol{X}^{(\boldsymbol{\alpha}_i, \boldsymbol{e}_0)} \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{e}_0}$ such that $\chi(\boldsymbol{X}^{(\boldsymbol{\alpha}_i, \boldsymbol{e}_0)}) = x_i$.
  $\mathcal{M}(\boldsymbol{X}^{(\boldsymbol{\alpha}_i, \boldsymbol{e}_0)}) \leftarrow$ Macaulay matrix of degree $\mathbf{1}$ with respect to $\bar{>}$. (see Def. 7.3.7)
  for all $F \in P$ **do**
    Add F to $\mathcal{M}(\boldsymbol{X}^{(\boldsymbol{\alpha}_i, \boldsymbol{e}_0)})$.
  end for
  for all $\mathfrak{b}_j \in \mathfrak{b}$ **do**
    Add $\mathfrak{b}_j\, \boldsymbol{X}^{(\boldsymbol{\alpha}_i, \boldsymbol{e}_0)}$ to $\mathcal{M}(\boldsymbol{X}^{(\boldsymbol{\alpha}_i, \boldsymbol{e}_0)})$.
  end for
  Rearrange $\mathcal{M}(\boldsymbol{X}^{(\boldsymbol{\alpha}_i, \boldsymbol{e}_0)})$ as $\begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{bmatrix} (\boldsymbol{X}^{(\boldsymbol{\alpha}_i, \boldsymbol{e}_0)})$. (see Eq. 8.3)
  $\mathrm{M}_{x_i} \leftarrow (M_{2,2} - M_{2,1}\, M_{1,1}^{-1}\, M_{1,2})(\boldsymbol{X}^{(\boldsymbol{\alpha}_i, \boldsymbol{e}_0)})$. (see Thm. 8.1.10)
end for
$G \leftarrow$ `FGLM` $((\mathrm{M}_{x_1}, \ldots, \mathrm{M}_{x_n}), >)$. (See Alg. 5, recall that $1 \in \chi(\mathfrak{b})$)
return G.

Thus, in this case, the complexity of Gaussian elimination is $\mathcal{O}(P(\boldsymbol{d})^\omega)$ [GG13]. If $C(k, \boldsymbol{d})$ is cost of $\texttt{ReduceMacaulay}\,((F_1, \ldots, F_k), \boldsymbol{d}, >)$, then we have the following recursive relation

$$C(k, \boldsymbol{d}) = \begin{cases} \mathcal{O}(P(\boldsymbol{d})^\omega) & \text{if } k = 1, \\ C(k-1, \boldsymbol{d}) + C(k-1, \boldsymbol{d} - \boldsymbol{d}_k) + \mathcal{O}(P(\boldsymbol{d})^\omega) & \text{if } k > 1. \end{cases}$$

The cost $C(k-1, \boldsymbol{d})$ is greater than $C(k-1, \boldsymbol{d} - \boldsymbol{d}_k)$, as it involves bigger matrices. Hence, we obtain $C(k, \boldsymbol{d}) = \mathcal{O}(2^{k+1} P(\boldsymbol{d})^\omega)$. □

Theorem 8.1.14. *Consider an affine polynomial system $(f_1, \ldots, f_n)$ in $\mathbb{K}[\boldsymbol{x}]$ such that Ass. 8.1.1 holds. Let $(F_1, \ldots, F_n)$ be the system associated to $(f_1, \ldots, f_n)$, see discussion at the beginning of Sec. 8.1.1, where $F_i \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\boldsymbol{e}_i}$ and $\chi(F_i) = f_i$, for $i \in [n]$. Assume that $(F_1, \ldots, F_n)$ is* (sparse) regular *(Def. 7.1.7). Then, the arithmetic complexity of computing a Gröbner basis for $\langle f_1, \ldots, f_n \rangle : \langle \prod_i x_i \rangle^\infty$ with respect to any monomial ordering $>$ is upper-bounded by*

$$\mathcal{O}(2^{n+1}\, P(\boldsymbol{d})^\omega + n\, \mathrm{MV}(\Delta_1, \ldots, \Delta_n)^3) \text{ operations}.$$

Proof. In Alg. 11, we need to compute:

- The set $\texttt{Rows}(\texttt{ReduceMacaulay}\left((F_1, \ldots, F_n), \sum_{i>1} \boldsymbol{e}_i, >\right))$ to generate $\mathfrak{b}$ (Rem. 8.1.7). By Lem. 8.1.13, this costs $\mathcal{O}(2^{n+1}\, P(\sum_{i>1} \boldsymbol{e}_i)^\omega)$.
- The set $\texttt{Rows}(\texttt{ReduceMacaulay}\,((F_1, \ldots, F_n), \mathbf{1}, >))$ to generate the matrix $\mathcal{M}(F_0)$ of Lem. 8.1.8. By Lem. 8.1.13, it costs $\mathcal{O}(2^{n+1}\, P(\mathbf{1})^\omega)$.
- For each variable x_i, the Schur complement of $\mathcal{M}(F_0)$, for $\chi(F_0) = x_i$. The cost of each Schur complement computation is $\mathcal{O}(P(\mathbf{1})^\omega)$, and so the cost of this step is $\mathcal{O}(n\, P(\mathbf{1})^\omega)$.
- The complexity of `FGLM` depends on the number of solutions (Prop. 5.2.8). In this case, the number of solutions is $\mathrm{MV}(\Delta_1, \ldots, \Delta_n)$ (Prop. 7.3.5). Hence, the cost of this step is $\mathcal{O}(n\, \mathrm{MV}(\Delta_1, \ldots, \Delta_n)^3)$.

□

Note that $\mathrm{MV}(\Delta_1, \ldots, \Delta_n) < P(\mathbf{1})$. Hence, to improve the previous bound for lexicographical orders we can follow [FGHR13].

8.2 Mixed multihomogeneous systems

We consider an algorithm for solving zero-dimensional square multihomogeneous systems, see Def. 2.10.8. This algorithm relies on the same principles as our algorithm in Sec. 8.1 but we improve the complexity by exploiting the fact that we consider multihomogeneous polynomials.

8.2.1 Notation

We follow the same notation as in Sections 2.10 and 2.11 Let $n_1, \dots n_q \in \mathbb{N}$, $N := \sum_i n_i$, and $\boldsymbol{n} := (n_1, \dots, n_q) \in \mathbb{N}^q$. For $1 \leq i \leq q$, let $\boldsymbol{x_i}$ be the set of variables $\{x_{i,0}, \dots, x_{i,n_i}\}$. Let $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q] := \bigotimes_{i=1}^q \mathbb{K}[\boldsymbol{x}_i]$ be the multihomogeneous $\mathbb{Z}^q$-graded $\mathbb{K}$-algebra, such that for all $\boldsymbol{d} := (d_1, \dots, d_q) \in \mathbb{Z}^q$, it holds $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]_{\boldsymbol{d}} := \bigotimes_{i=1}^q \mathbb{K}[\boldsymbol{x}_i]_{d_i}$. Given a multihomogeneous polynomial $F \in \mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]_{\boldsymbol{d}}$, we denote its multidegree by $\deg(F) = \boldsymbol{d} \in \mathbb{N}^q$. Given a $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$-module M, we consider $(M)_{\boldsymbol{d}}$ as the graded part of M of multidegree $\boldsymbol{d}$. Given two multidegrees $\boldsymbol{d}_1, \boldsymbol{d}_2 \in \mathbb{N}^q$, we say that $\boldsymbol{d}_1 \geq \boldsymbol{d}_2$ if the inequality holds component-wise, that is, if $\boldsymbol{d}_1 - \boldsymbol{d}_2 \in \mathbb{N}^q$. We consider the multiprojective space $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$. Let $\mathbf{1} = (1, \dots, 1) \in \mathbb{Z}^q$ be the multidegree corresponding to multilinear polynomials in $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$. Consider the ideal generated by all the polynomials in $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]_{\mathbf{1}}$,

$$B = \bigcap_{i=1}^{q} \langle x_{i,0}, \dots, x_{i,n_i} \rangle.$$

Let $\mathcal{K}_\bullet(F_1, \dots, F_k)$ be the Koszul complex of $F_1, \dots, F_k$ over $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$, see Def. 2.7.5. Let $\mathcal{H}_i(F_1, \dots, F_k)$ be the i-th Koszul homology module of $\mathcal{K}_\bullet(F_1, \dots, F_k)$. Given a $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$-module M, let $H^j_B(M)$ be its j-th local cohomology module at B (Sec. 2.9).

Let $\boldsymbol{x}_h := \prod_{i=1}^q x_{i,0} \in \mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]_{\mathbf{1}}$. We say that a multihomogeneous system $(F_1, \dots, F_N)$ has *no solutions* at infinity if the system $(F_1, \dots, F_N, \boldsymbol{x}_h)$ has no solutions over $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_q}$. We dehomogenize a multihomogeneous polynomial by replacing each variable $x_{i,0}$ with 1. That is, we consider the dehomogenization homomorphism χ, see Def. 7.1.2, such that,

$$\chi(x_{i,j}) = \begin{cases} x_{i,j} & \text{if } j > 0 \\ 1 & \text{if } j = 0. \end{cases}$$

We define the algebra $\mathbb{K}[\boldsymbol{x}_{\text{aff}}]$ as the $\mathbb{K}$-algebra obtained from the dehomogenization of $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$, that is

$$\mathbb{K}[\boldsymbol{x}_{\text{aff}}] := \chi(\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]) = \mathbb{K}[x_{1,1}, \dots, x_{q,n_q}].$$

Given $F \in \mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$, we consider as $\chi(F) \in \mathbb{K}[\boldsymbol{x}_{\text{aff}}]$, its dehomogenization.

Remark 8.2.1. *For each $1 \leq i \leq q$, let $\Delta_i := \text{NP}(1 + \sum_{j=1}^{n_i} x_{i,j}) \subset \mathbb{R}^N$. Each polytope Δ_i has dimension n_i and $\Delta = \sum_{i=1}^q \Delta_i \subset \mathbb{R}^N$ is full-dimensional. The algebra $\mathbb{K}[\boldsymbol{x}_1, \dots, \boldsymbol{x}_q]$ is isomorphic to the $\mathbb{Z}^q$-graded algebra $\mathbb{K}[S^h_{\boldsymbol{\Delta}}]$ from Def. 7.1.1. Moreover, $\mathbb{K}[\boldsymbol{x}_{\text{aff}}]$ is isomorphic to $\mathbb{K}[S_\Delta]$.*

8.2.2 Multihomogeneous Macaulay bound

In this section, we construct bounds for the maximal degrees on which the homologies of the Koszul complex $\mathcal{K}_\bullet(F_1, \dots, F_k)$ do not vanish. For that, we assume that $(F_1, \dots, F_k)$ is a *regular sequence outside B*, see Def. 2.11.6. We use local cohomology (Sec. 2.9) and the multigraded Castelnuovo-Mumford regularity (Sec. 2.11). We follow the same notation as in these sections.

Proposition 8.2.2 (Multihomogeneous Macaulay bound)**.** *Let $(F_1, \dots, F_k)$ be a regular sequence outside B and consider $\boldsymbol{d} \geq \left(\sum_{i=1}^k \deg(F_i)\right) - \boldsymbol{n}$, where $\boldsymbol{n} := (n_1, \dots, n_q) \in \mathbb{N}^q$. Then, for each i, j, the j-th local cohomology of $\mathcal{H}_i(F_1, \dots, F_k)$ vanishes, that is, for all i, j it holds*

$$\left(H^j_B(\mathcal{H}_i(F_1, \dots, F_k))\right)_{\boldsymbol{d}} = 0.$$

Proof. We prove this lemma using Prop. 2.11.10 and so, we follow its notation. Fix i and j in Eq. (2.5), and consider $\alpha \subset \{1,\dots,q\}$ such that $N_\alpha + 1 + j - i \leq k$, $\#\alpha \neq \emptyset$, and $v \in \mathbf{\Sigma}(N_\alpha + 1 + j - i)$, see Eq. (2.4). We pick $t \in \alpha$. Then, the t-th coordinate of any element in $Q_\alpha + v$ is $\leq -n_t - 1 + v_t$, where v_t is the t-th coordinate of v. As all the multidegrees $\deg(F_1),\dots,\deg(F_k)$ are non-negative, $v_t \leq \sum_{i=1}^k \deg(F_i)_t$. So, $-n_t - 1 + v_t < -n_t + \sum_{i=1}^k \deg(F_i)_t \leq d_t$, where d_t is the t-coordinate of $\boldsymbol{d}$. Hence, $\boldsymbol{d} \notin Q_\alpha + v$ for any α and v as in Eq. (2.5). Therefore, by Prop. 2.11.10, $(H_B^j(H_i^k))_{\boldsymbol{d}} = 0$. □

If $(F_1,\dots,F_k)$ is a regular sequence outside B, then Prop. 8.2.2 give us a bound for the multigraded Castelnuovo-Mumford regularity of $\mathbb{K}[\boldsymbol{x}_1,\dots,\boldsymbol{x}_q]/(F_1,\dots,F_k)$. We call this bound the *multihomogeneous Macaulay bound*, as it extends the Macaulay bound (Prop. 2.8.5) to multihomogeneous polynomials. In contrast with the Macaulay bound, the *multigraded Macaulay bound* is not tight, see [ACG05, Sec. 4.4].

Corollary 8.2.3. *Let $F_1,\dots,F_k$ be regular sequence outside B and consider $\boldsymbol{d} \geq \left(\sum_{i=1}^k \deg(F_i)\right) - \boldsymbol{n}$.*

- *If $k = N$, then the dimension of $(\mathbb{K}[\boldsymbol{x}_1,\dots,\boldsymbol{x}_q]/\langle F_1,\dots,F_N\rangle)_{\boldsymbol{d}}$ is the multihomogeneous Bézout bound, that is, the number of solutions, counting multiplicities, of the system $(F_1,\dots,F_N)$ over $\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_q}$, see Prop. 2.10.9.*
- *If $k = N+1$, then $\mathbb{K}[\boldsymbol{x}_1,\dots,\boldsymbol{x}_q]_{\boldsymbol{d}} = (\langle F_1,\dots,F_{N+1}\rangle)_{\boldsymbol{d}}$.*

8.2.3 Computing graded parts of the ideals

Let $(F_1,\dots,F_k)$ be multihomogeneous system in $\mathbb{K}[\boldsymbol{x}_1,\dots,\boldsymbol{x}_q]$. Algorithm 10 computes a set of generators of the vector space $(\langle F_1,\dots,F_k\rangle)_{\boldsymbol{d}}$. If $(F_1,\dots,F_k)$ form a regular sequence outside B, and $\boldsymbol{d} \geq \left(\sum_{i=1}^k \deg(F_i)\right) - \boldsymbol{n}$, then it performs no reductions to zero.

Theorem 8.2.4. *If $(F_1,\dots,F_k)$ is a regular sequence outside B, then for every $\boldsymbol{d} \geq \left(\sum_{i=1}^k \deg(F_i)\right) - \boldsymbol{n}$, `ReduceMacaulay`$(F_1,\dots,F_k,\boldsymbol{d},<)$ (Alg. 10) only considers matrices with linearly independent rows and avoids all redundant computations.*

Proof. We proceed by induction on k. Let $\boldsymbol{D}_k := \left(\sum_{i=1}^k \deg(F_i)\right) - \boldsymbol{n}$. When $k = 1$, the ideal is principal and so the theorem holds. In step k, note that $\boldsymbol{d} \geq \boldsymbol{D}_k$ implies $\boldsymbol{d} \geq \boldsymbol{d} - \deg(F_k) \geq \boldsymbol{D}_k - \deg(F_k) = \boldsymbol{D}_{k-1}$. Hence, we have no reductions to zero in the recursive calls. By Prop. 2.11.7, it holds $H_B^0(\mathcal{H}_1(F_1,\dots,F_k)) = \mathcal{H}_1(F_1,\dots,F_k)$. As $\boldsymbol{d} \geq \boldsymbol{D}_k$, by Prop. 8.2.2, we conclude

$$(\mathcal{H}_1(F_1,\dots,F_k))_{\boldsymbol{d}} = H_B^0(\mathcal{H}_1(F_1,\dots,F_k))_{\boldsymbol{d}} = 0.$$

Hence, by Lem. 7.3.16, the rows of the matrix $\mathcal{M}_{\boldsymbol{d}}^k$ in Alg. 10 are all linearly independent and so, the algorithm performs no reductions to zero when it computes $\widetilde{\mathcal{M}}_{\boldsymbol{d}}^k$. □

8.2.4 Solving zero-dimensional multihomogeneous systems

Our solving strategy is to dehomogenize the system and to compute the multiplication maps for the affine variables. Then, we can apply `FGLM` (Alg. 5) to compute a Gröbner basis or to compute the eigenvalues/eigenvectors of the multiplication maps (Sec. 5.2.1).

Let $(F_1, \ldots, F_N)$ be a zero-dimensional system in $\mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]$ with no solutions at infinity, such that it forms regular sequence outside B. If we do not know that the system has no solutions at infinity, then we can ensure this condition by performing a generic linear change of coordinates (Def. 4.5.7) preserving the multihomogeneous structure, e.g. see [CLO06, Pg. 121]. We use Alg. 10 to construct a monomial basis and the multiplication maps over $\mathbb{K}[\boldsymbol{x}_{\text{aff}}]/\langle \chi(F_1), \ldots, \chi(F_N)\rangle$. Our strategy is similar to Sec. 8.1, but instead of introducing a linear polynomial f_0 involving every variable in $\mathbb{K}[\boldsymbol{x}_{\text{aff}}]$, we introduce a multilinear polynomial $f_0 \in \mathbb{K}[\boldsymbol{x}_{\text{aff}}]$ which is the dehomogenization of $F_0 \in \mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]_{\mathbf{1}}$. We do so because, by introducing a multihomogeneous F_0, we do not modify the underling toric variety [CLS11, Cor. 6.2.14].

Let $\boldsymbol{D}_N := \left(\sum_{i=1}^{N} \deg(F_i)\right) - \boldsymbol{n}$. Following Sec. 8.1, we fix a graded monomial order $<$ and consider $\mathfrak{b}$ as the set of monomials that are not leading monomials of $(\langle F_1, \ldots, F_N\rangle)_{\boldsymbol{D}_N}$, that is

$$\mathfrak{b} := \{\boldsymbol{x}^\alpha \in \mathbb{K}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q]_{\boldsymbol{D}_N} : (\forall G \in (\langle F_1, \ldots, F_N\rangle)_{\boldsymbol{D}_N})\ \mathtt{LM}_<(G) \neq \boldsymbol{x}^\alpha\}$$

We will prove that the dehomogenization of these monomials, $\chi(\mathfrak{b})$, forms a monomial basis for $\mathbb{K}[\boldsymbol{x}_{\text{aff}}]/\langle \chi(F_1), \ldots, \chi(F_N)\rangle$ following the same arguments as in Lem. 8.1.5 and Cor. 8.1.6.

Lemma 8.2.5. *If the system $F_1, \ldots, F_N$ has no solutions at infinity, then $\chi(\mathfrak{b})$ forms a monomial basis for $\mathbb{K}[\boldsymbol{x}_{\text{aff}}]/\langle \chi(F_1), \ldots, \chi(F_N)\rangle$.*

Proof. The set $\chi(\mathfrak{b})$ is a monomial basis if and only if all its elements are linearly independent on the quotient ring $\mathbb{K}[\boldsymbol{x}_{\text{aff}}]/\langle \chi(F_1), \ldots, \chi(F_N)\rangle$ and they generate this ring. By Cor. 8.2.3, the dimension of the quotient ring, as a $\mathbb{K}$-vector space, is the same as the cardinal of $\chi(\mathfrak{b})$. Hence, to conclude the proof we need to show that the elements of $\chi(\mathfrak{b})$ are linearly independent. Assume that there is a non-trivial linear combination $p := \sum_i c_i\, \chi(\mathfrak{b}_i)$ of the elements of $\chi(\mathfrak{b})$ such that $p \equiv 0$ in $\mathbb{K}[\boldsymbol{x}_{\text{aff}}]/\langle \chi(F_1), \ldots, \chi(F_N)\rangle$. Then, similarly to Prop. 4.3.2, there is a $\omega \in \mathbb{N}$ such that $(\boldsymbol{x}_h)^\omega \cdot P \in \langle F_1, \ldots, F_N\rangle$, where $P := \sum_i c_i\, \mathfrak{b}_i$. By construction of $\mathfrak{b}$, the set is linearly independent over $\langle F_1, \ldots, F_N\rangle$. Hence, $\omega > 0$. As $F_1, \ldots, F_n$ has no solutions at infinity, then the system $(F_1, \ldots, F_N, \boldsymbol{x}_h^\omega)$ has no solutions. Using Prop. 8.2.2, we can prove that $H_1(F_1, \ldots, F_N, \boldsymbol{x}_h^\omega)_{\boldsymbol{D_N}+\omega\,\mathbf{1}} = 0$. By Lem. 7.3.15, as $\boldsymbol{x}_h^\omega \cdot P \in (\langle F_1, \ldots, F_N\rangle)_{\boldsymbol{D_N}+\omega\,\mathbf{1}}$, then $P \in (\langle F_1, \ldots, F_N\rangle)_{\boldsymbol{D_N}}$ and so we get a contradiction. Therefore, such a linear combination $p \neq 0$ does not exists and the set $\chi(\mathfrak{b})$ is linearly independent on $\mathbb{K}[\boldsymbol{x}_{\text{aff}}]/\langle \chi(F_1), \ldots, \chi(F_N)\rangle$. □

Remark 8.2.6. *One way to compute the set $\mathfrak{b}$ is to compute a basis of the vector space $\langle F_1, \ldots, F_n\rangle_{\boldsymbol{D}_N}$ using Alg. 10, that is* $\mathtt{ReduceMacaulay}\left((F_1, \ldots, F_N), \boldsymbol{D}_N, <\right)$.

Consider $\boldsymbol{d}_{\text{reg}} := \boldsymbol{D}_N + \mathbf{1}$. We will construct a matrix similar to Eq. (8.3) and deduce the multiplications maps from its Schur complement. For each $F_0 \in \mathbb{K}[S^h_{\boldsymbol{\Delta}}]_{\mathbf{1}}$, we will construct a Macaulay matrix for $(F_1, \ldots, F_N, F_0)$ at multidegree $\boldsymbol{d}_{\text{reg}}$, say $\mathcal{M}(F_0)$. The rows of $\mathcal{M}(F_0)$ correspond to polynomials of two kinds:

- the polynomials in $\mathtt{Rows}(\widetilde{\mathcal{M}}^n_{\boldsymbol{d}_{\text{reg}}})$, where $\widetilde{\mathcal{M}}^n_{\boldsymbol{d}_{\text{reg}}} = \mathtt{ReduceMacaulay}\left((F_1, \ldots, F_N), \boldsymbol{d}_{\text{reg}}, <\right)$,

- the polynomials of the form $\mathfrak{b}_i\, F_0$, where $\mathfrak{b}_i \in \mathfrak{b}$.

Lemma 8.2.7. *The matrix $\mathcal{M}(F_0)$ is always square. It is full-rank if and only if $(F_1, \ldots, F_N, F_0)$ has no solutions.*

The proof of this lemma is the same as the one of Lem. 8.1.8 taking into account that, by Cor. 8.2.3, $H_0(F_1, \ldots, F_N, \boldsymbol{x}_h)_{\boldsymbol{d}_{\text{reg}}} = 0$.

We reorder the columns of $\mathcal{M}(F_0)$ as shown in Eq. (8.5), such that

- the columns of the submatrix $\left[\begin{smallmatrix} M_{1,2}(F_0) \\ M_{2,2}(F_0) \end{smallmatrix}\right]$ correspond to monomials of the form $\mathfrak{b}_i\, \boldsymbol{x_h}$, where $\mathfrak{b}_i \in \mathfrak{b}$, and
- the rows of $\left[\begin{smallmatrix} M_{2,1}(F_0) \mid M_{2,2}(F_0) \end{smallmatrix}\right]$ are polynomials of the form $\mathfrak{b}_i\, F_0$, where $\mathfrak{b}_i \in \mathfrak{b}$.

$$\mathcal{M}(F_0) = \begin{array}{r} \\ \texttt{Rows}(\widetilde{\mathcal{M}}^n_{\boldsymbol{d}_{\text{reg}}}) \\ \\ F_0 \cdot \mathfrak{b} \end{array} \overbrace{\left\{\begin{array}{|c||c|}\hline M_{1,1}(F_0) & M_{1,2}(F_0) \\ \hline\hline M_{2,1}(F_0) & M_{2,2}(F_0) \\ \hline \end{array}\right.}^{\boldsymbol{x_h} \cdot \mathfrak{b}} \tag{8.5}$$

Theorem 8.2.8. *If the system $F_1, \ldots, F_N$ is a regular sequence outside B and has no solutions at infinity, for any $F_0 \in \mathbb{K}[\boldsymbol{x}_{\text{aff}}]_{\mathbf{1}}$, the matrix $M_{1,1}(F_0)$ is invertible and the Schur complement of $M_{2,2}(F_0)$, $M^c_{2,2}(F_0) := (M_{2,2} - M_{2,1}\, M_{1,1}^{-1}\, M_{1,2})(F_0)$, is the multiplication map of $\chi(F_0)$ in the basis $\chi(\mathfrak{b})$ of $\mathbb{K}[\boldsymbol{x}_{\text{aff}}]/\langle \chi(F_1), \ldots, \chi(F_N)\rangle$.*

The proof of this theorem is exactly the same as the one of Lem. 8.1.9 and Thm. 8.1.10.

8.3 Unmixed sparse polynomial systems

Faugère et. al. [FSS14] introduced a variant of the `FGLM` algorithm (Alg. 5) to solve zero-dimensional *unmixed sparse systems*. Their approach needs to precompute a Gröbner basis over a semigroup algebra $\mathbb{K}[S_\Delta]$, which they use to compute normal forms. By bounding the complexity of computing such a Gröbner basis, they derive complexity bounds for the arithmetic complexity of solving *unmixed sparse systems*. In particular, under regularity assumptions, they propose a bound for the maximal degree of an element in the Gröbner basis [FSS14, Lem. 5.2]. Their bound depends on the combinatorics of the Newton polytope.

In this section, we show that their complexity bound misses some assumptions to hold; in Sec. 8.3.1, we present a counter-example to it. Moreover, it is not clear which are the assumptions, if any, for this bound to hold for arbitrary unmixed systems. With this problem as a motivation, we show how we can adapt our solving strategy from Sec. 8.1 and combine it with their `FGLM`-like algorithm to derive a novel algorithm with the complexity bound claimed in [FSS14, Thm. 5.3].

8.3.1 Counter-examples to the complexity bound of [FSS14]

Let Δ be the standard 2-simplex and consider the regular system given by two polynomials $F_1, F_2 \in \mathbb{K}[S^h_{\mathbf{\Delta}}]_2$ of degree 2, that is,

$$\begin{aligned} F_1 :=& \boldsymbol{X}^{([2,0],2)} + \boldsymbol{X}^{([1,1],2)} + \boldsymbol{X}^{([0,2],2)} + \boldsymbol{X}^{([1,0],2)} + \boldsymbol{X}^{([0,1],2)} + \boldsymbol{X}^{([0,0],2)} \\ F_2 :=& \boldsymbol{X}^{([2,0],2)} + 2\,\boldsymbol{X}^{([1,1],2)} + 3\,\boldsymbol{X}^{([0,2],2)} + 4\,\boldsymbol{X}^{([1,0],2)} + 5\,\boldsymbol{X}^{([0,1],2)} + 6\,\boldsymbol{X}^{([0,0],2)} \end{aligned}$$

Consider the graded monomial order $<$ defined as follows,

$$\boldsymbol{X}^{([x_1,y_1],d_1)} < \boldsymbol{X}^{([x_2,y_2],d_2)} \iff \begin{cases} d_1 < d_2, \text{ or} \\ d_1 = d_2 \text{ and } x_1 < x_2, \text{ or} \\ d_1 = d_2 \text{ and } x_1 = x_2 \text{ and } y_1 < y_2 \end{cases} .$$

In this case, the bound claimed in [FSS14, Lem. 5.2] is 3. However, the maximal degree of an element in the Gröbner basis of (F_1, F_2) with respect to $<$ is 4.

We also studied a more complicated semigroup algebra $\mathbb{K}[S^h_{\mathbf{\Delta}}]$, where Δ corresponds to the Newton polytope of $1+x+y+x\,y+x^2\,y^2$. We considered two generic polynomials $F_1, F_2 \in \mathbb{K}[S^h_{\mathbf{\Delta}}]_2$ of degree 2 and computed the different maximal degrees of the elements in each Gröbner bases of (F_1, F_2). For this system, we have performed an extensive search over different monomials orderings $<$, see Def. 7.3.1. We tried more than $40\,000$ monomial orderings and for none of them the bound in [FSS14, Lem. 5.2] holds.

8.3.2 Our approach

Following the same notation as in Sec. 7.3, we consider $\mathbb{K}[S_\Delta]$ and $\mathbb{K}[S^h_{\mathbf{\Delta}}]$ with respect to only one polytope Δ instead of many, and so $\mathbb{K}[S^h_{\mathbf{\Delta}}]$ is a $\mathbb{N}$-graded ring. We assume that Δ is a normal polytope.

Definition 8.3.1 (Normal polytope)**.** *[CLS11, Def. 2.2.9] We say that an integer polytope Δ is normal if for all $k, l \in \mathbb{N}$,*

$$(k \cdot \Delta) \cap \mathbb{Z}^n + (l \cdot \Delta) \cap \mathbb{Z}^n = ((k+l) \cdot \Delta) \cap \mathbb{Z}^n.$$

When Δ is normal, the algebra $\mathbb{K}[S^h_{\mathbf{\Delta}}]$ is Cohen-Macaulay [CLS11, Thm. 9.2.9] and so we can consider homogeneous regular sequences of maximal length. In particular, generic square homogeneous systems are regular sequences. This means that, in these cases, our algorithm `ReduceMacaulay` (Alg. 10) will avoid every reduction to zero. This is so because the Koszul complex of a regular sequence is exact, see Prop. 2.7.8.

Theorem 8.3.2. *Consider homogeneous polynomials $(F_1, \ldots, F_n, F_0)$ of degrees $d_1, \ldots, d_n, d_0 \in \mathbb{N}$, respectively, such that $(F_1, \ldots, F_n, F_0)$ is a homogeneous regular sequence over $\mathbb{K}[S^h_{\mathbf{\Delta}}]$ and $d_0 = 1$. Let r be the smallest integer such that $r \cdot \Delta$ contains an integer interior point. Let $D_n := (\sum_{i \geq 1} d_i) - r + 1$. Then, for all $d \geq D_n$, it holds*

- $\dim_{\mathbb{K}}([\mathbb{K}[S^h_{\mathbf{\Delta}}]/\langle F_1, \ldots, F_n\rangle]_d) = \mathrm{MV}(d_1\Delta, \ldots, d_n\,\Delta)$ *and*
- $\dim_{\mathbb{K}}([\mathbb{K}[S^h_{\mathbf{\Delta}}]/\langle F_1, \ldots, F_n, F_0\rangle]_{d+1}) = 0.$

Proof. We can prove this statement by studying the Hilbert series of $\mathbb{K}[S^h_{\mathbf{\Delta}}]$, following the first part of the argument in the proof of [FSS16, Lem. 5.2]. This series is given by the rational function $HS_{\mathbb{K}[S^h_{\mathbf{\Delta}}]}(t) = \frac{Q(t)}{(1-t)^{n+1}}$, where the degree of $Q(t)$ is $(n+1-r)$ [FSS16, Prop. 2.7]. □

We will state the following theorems without proving them. Their proofs rely in the same ideas as in Sections 8.1 and 8.2. We fix a graded monomial order $<$ and consider $\mathfrak{b}$ as the set of monomials that are not leading monomials of $[\langle F_1, \ldots, F_n\rangle]_{D_n}$, that is

$$\mathfrak{b} := \left\{ \boldsymbol{X}^{(\boldsymbol{\alpha}, D_n)} \in \mathbb{K}[S^h_{\mathbf{\Delta}}]_{D_n} : (\forall G \in [\langle F_1, \ldots, F_n\rangle]_{D_n})\ \mathtt{LM}_<(G) \neq \boldsymbol{x}^{\alpha} \right\}.$$

Lemma 8.3.3. *If the system $(F_1, \ldots, F_N)$ has no solutions at infinity, see Ass. 8.1.1, then the dehomogenization (Def. 7.1.2) of $\mathfrak{b}$, $\chi(\mathfrak{b})$, is a monomial basis of $\mathbb{K}[S_\Delta]/\langle \chi(F_1), \ldots, \chi(F_N)\rangle$.*

Let $d_{\text{reg}} := D_n + 1$. For each $F_0 \in \mathbb{K}[S^h_{\mathbf{\Delta}}]_1$, we will construct a Macaulay matrix for $(F_1, \ldots, F_n, F_0)$ at degree d_{reg}, say $\mathcal{M}(F_0)$. The rows of $\mathcal{M}(F_0)$ corresponds to polynomials of two kinds:

- the polynomials in $\mathtt{Rows}(\widetilde{\mathcal{M}}^n_{d_{\text{reg}}})$, where $\widetilde{\mathcal{M}}^n_{d_{\text{reg}}} = \mathtt{ReduceMacaulay}\left((F_1, \ldots, F_n), d_{\text{reg}}, <\right)$, and
- the polynomials of the form $\mathfrak{b}_i\, F_0$, where $\mathfrak{b}_i \in \mathfrak{b}$.

Lemma 8.3.4. *The matrix $\mathcal{M}(F_0)$ is always square. It is full-rank if and only if $(F_1, \ldots, F_N, F_0)$ is a regular sequence.*

We reorder the columns of $\mathcal{M}(F_0)$ as shown in Eq. (8.6), so that

- the columns of the submatrix $\left[\begin{smallmatrix} M_{1,2}(F_0) \\ \hline M_{2,2}(F_0) \end{smallmatrix}\right]$ correspond to monomials of the form $\mathfrak{b}_i\, \boldsymbol{X}^{(0,1)}$, where $\mathfrak{b}_i \in \mathfrak{b}$, and
- the rows of $\left[\begin{smallmatrix} M_{2,1}(F_0) \mid M_{2,2}(F_0) \end{smallmatrix}\right]$ correspond to polynomials of the form $\mathfrak{b}_i\, F_0$, where $\mathfrak{b}_i \in \mathfrak{b}$.

$$\mathcal{M}(F_0) = \begin{array}{r} \\ \mathtt{Rows}(\widetilde{\mathcal{M}}^n_{d_{\text{reg}}}) \\ \\ F_0 \cdot \mathfrak{b} \end{array} \left\{ \begin{array}{c} \overbrace{\phantom{M_{1,2}(F_0)}}^{\boldsymbol{X}^{(0,1)} \cdot \mathfrak{b}} \\ \left[\begin{array}{c|c} M_{1,1}(F_0) & M_{1,2}(F_0) \\ \hline M_{2,1}(F_0) & M_{2,2}(F_0) \end{array}\right] \end{array} \right. \tag{8.6}$$

Theorem 8.3.5. *If the system $F_1, \ldots, F_N$ is a regular sequence, for any $F_0 \in \mathbb{K}[S^h_{\mathbf{\Delta}}]_1$, the matrix $M_{1,1}(F_0)$ is invertible and the Schur complement of $M_{2,2}(F_0)$, $M^c_{2,2}(F_0) := (M_{2,2} - M_{2,1}\, M^{-1}_{1,1}\, M_{1,2})(F_0)$, is the multiplication map of $\chi(F_0)$, in the basis $\chi(\mathfrak{b})$, of $\mathbb{K}[S_\Delta]/\langle \chi(F_1), \ldots, \chi(F_N)\rangle$.*

Once we have constructed the multiplications maps over $\mathbb{K}[S_\Delta]$, we can solve the system by, first, computing a lexicographical Gröbner basis for $\langle \chi(F_1), \ldots, \chi(F_N) \rangle$ over $\mathbb{K}[S_\Delta]$ using the algorithm `Sparse-FGLM` [FSS16, Alg. 2], and then inverting a monomial map. This strategy is the same as the one suggested in [FSS16, Sec. 3], but our approach to compute the multiplication maps is different; we do not compute an intermediate Gröbner basis.

Remark 8.3.6. *We emphasize that the matrix* `ReduceMacaulay` $((F_1, \ldots, F_n), d_{reg}, <)$ *that we use to construct the set* `Rows`$(\widetilde{\mathcal{M}}^n_{d_{reg}})$ *is the same as the one computed in [FSS16, Alg. 1] with the parameters of [FSS16, Lem. 2]. The correctness of our approach relies on the fact that at degree* d_{reg} *we have enough information to compute the multiplication maps. We do not need to assume, as in [FSS16], that at this degree we had computed a Gröbner basis; which is not true in general, see discussion in Sec. 8.3.1.*

Chapter 9

Binary form decomposition

Symmetric tensor decomposition is an important problem with applications in several areas for example signal processing, statistics, data analysis and computational neuroscience. It is equivalent to Waring's problem for homogeneous polynomials, that is to write a homogeneous polynomial in n variables of degree D as a sum of D-th powers of linear forms, using the minimal number of summands. This minimal number is called the *rank* of the polynomial/tensor. We focus on decomposing binary forms, a problem that corresponds to the decomposition of symmetric tensors of dimension 2 and order D. Under this formulation, the problem finds its roots in invariant theory where the decompositions are known as canonical forms. In this context many different algorithms were proposed. In recent years, those algorithms were extended for the general symmetric tensor decomposition problem.

In chapter 9, we introduce a new *superfast* algorithm that improves the complexity of previous approaches to decompose binary forms. For that, we use results from *structured linear algebra*. It achieves a *softly linear*, and so *quasi-optimal*, arithmetic complexity bound. To the best of our knowledge, the previously known algorithms have at least quadratic complexity bounds. Our algorithm computes a symbolic decomposition in $O(\mathtt{M}(D)\log(D))$ arithmetic operations (Thm. 9.4.19), where $\mathtt{M}(D)$ is the complexity of multiplying two polynomials of degree D. It is deterministic when the decomposition is unique. When the decomposition is not unique, our algorithm is randomized. We present a Monte Carlo version of it and we show how to modify it to a Las Vegas one, within the same complexity.

From the symbolic decomposition, we approximate the terms of the decomposition with an error of $2^{-\varepsilon}$ in $O\big(D\log^2(D)\big(\log^2(D)+\log(\varepsilon)\big)\big)$ arithmetic operations (Thm. 9.4.20). Moreover, we bound the algebraic degree of the problem by $\min(rank, D-rank+1)$ (Thm. 9.4.15). We show that this bound can be tight. When the input polynomial has integer coefficients, our algorithm performs, up to poly-logarithmic factors, $\widetilde{\mathcal{O}}_B(D\ell + D^4 + D^3\tau)$ bit operations (Thm. 9.4.21), where τ is the maximum bitsize of the coefficients and $2^{-\ell}$ is the relative error of the terms in the decomposition.

This chapter is self-contained.

9.1 Introduction

The problem of decomposing a symmetric tensor consists in writing it as the sum of rank-1 symmetric tensors, using the minimal number of summands. This minimal number is known as the rank of the

symmetric tensor[1]. The symmetric tensors of rank-1 correspond to, roughly speaking, the k-th outer-product of a vector. The decomposition of symmetric tensor is a common problem which appears in divers areas, such as signal processing, statistics, data mining, computational neuroscience, computer vision, psychometrics, chemometrics, among others. For a contemporary introduction to the theory of tensor, their decompositions and applications we refer to e.g., [Com14, Lan11].

There is an equivalence between decomposing symmetric tensors and solving Waring's problem for homogeneous polynomials, e.g., [CGLM08, Hel92]. Given a symmetric tensor of dimension n and order D we can construct a homogeneous polynomial in n variables of total degree D. Then, finding the decomposition for the tensor is equivalent to write the polynomial as a sum of D-th powers of linear forms, using the minimal numbers of summands. This minimal number is the rank the polynomial/tensor.

Under this formulation, symmetric tensor decomposition dates back to the origin of modern (linear) algebra as a part of Invariant Theory. In this setting, the decomposition corresponds to canonical forms [Syl04b, Syl04a, Gun87]. Together with the theory of apolarity, this problem was of great importance because the decompositions provide information about the behavior of the polynomials under linear change of variables [KR84].

Binary Form Decomposition We study the decomposition of symmetric tensors of order D and dimension 2. In terms of homogeneous polynomials, we consider a binary form

$$f(x,y) := \sum_{i=0}^{D} \binom{D}{i} a_i x^i y^{D-i}, \tag{9.1}$$

where $a_i \in \mathbb{K} \subset \mathbb{C}$ and $\mathbb{K}$ is some field of characteristic zero. We want to compute a decomposition

$$f(x,y) = \sum_{j=1}^{r} (\alpha_j x + \beta_j y)^D, \tag{9.2}$$

where $\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_r \in \overline{\mathbb{K}}$ (the algebraic closure of $\mathbb{K}$) and r is minimal. We say that a decomposition *unique* if, for all the decompositions, the set of points $\{(\alpha_j, \beta_j) : 1 \leq j \leq r\} \subset \mathbb{P}^1(\mathbb{K})$ is unique, where $\mathbb{P}^1(\mathbb{K})$ is the projective space of $\overline{\mathbb{K}}$ [Rez13a].

Previous work The decomposition of binary forms, Eq. (9.2), has been largely studied for $\mathbb{K} = \mathbb{C}$. More than one century ago [Syl04a, Syl04b] described the necessary and sufficient conditions for a decomposition to exist (see Sec. 9.2.1). He related the decompositions to the kernel of Hankel matrices. For a modern approach of this topic, we refer to [KR84, Kun90, Rez13a, IK99]. Sylvester's work was extended to a more general kind of polynomial decompositions that we do not consider in this work, e.g., [Gun87, Rez96, IK99].

From the algorithmic point of view, Sylvester's work leads to an algorithm (Alg. 12) to decompose binary forms, see [CM96, Sec. 3.4.3]. In the case where the binary form is of odd degree, then we can compute the decompositions using Berlekamp-Massey algorithm, [Dü89]. When the decomposition is unique, the Catalecticant algorithm, which also works for symmetric tensors of bigger dimension [IK99, OO13], improves Sylvester's work. For an arbitrary binary form, [Hel92] presented a randomized algorithm based on Padé approximants and continued fractions, in which he also characterized the

[1] Some authors, e.g. [CGLM08], refer to this number as the symmetric rank of the tensor.

different possible decompositions. Unfortunately, all these algorithms have complexity at least quadratic in the degree of the binary form.

Besides the problem of computing the decomposition(s) many authors considered the subproblems of computing the rank and deciding where there exists a unique decomposition, e.g., [Syl04a, Syl04b, Hel92, CS11, BGI11]. For example, [Syl04a, Syl04b] considered generic binary forms, that is binary forms with coefficients belonging to a dense algebraic open subset of $\overline{\mathbb{K}}^{D+1}$ [CM96, Sec. 3], and proved that when the degree is $2k$ or $2k+1$, the rank is $k+1$ and that the minimal decomposition is unique only when the degree is odd. In the non-generic case, [Hel92, CS11, IK99], among others, proved that the rank is related to the kernel of a Hankel matrix and that the decomposition of a binary form of degree $2k$ or $2k-1$ and rank r, is unique if and only if $r \leq k$. With respect to the rank, different authors, e.g., [CS11, CGLM08, BGI11], proposed algorithms to compute its value. Even though the authors do not provide complexity estimates, using recent superfast algorithms for Hankel matrices [Pan01], we can deduce a nearly-optimal arithmetic complexity bound for the approach of [CS11].

For the general problem of symmetric tensor decomposition, Sylvester's work was successfully extended to cases in which the decomposition is unique [BCMT10, OO13]. Besides tensor decomposition, there are other related decompositions for binary forms and univariate polynomial that we do not treat, e.g., [Rez96, Rez13b, GKL03, GR10, GMKP17].

Formulation of the problem Instead of decomposing the binary form as in Eq. (9.2), we compute $\lambda_1 \ldots \lambda_r, \alpha_1 \ldots \alpha_r, \beta_1 \ldots \beta_r \in \overline{\mathbb{K}}$, where r is minimal, such that,

$$f(x,y) = \sum_{j=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^D. \tag{9.3}$$

Since every λ_j belongs to the algebraic closure of the field $\mathbb{K}$, the problems are equivalent. This approach allow us to control the algebraic degree [Baj88, NRS10, DHO$^+$16] of the parameters λ_j, α_j, and β_j in the decompositions (Section 9.4.1).

Note that if the field is not algebraically closed and we force the parameters to belong to the base field, that is $\lambda_j, \alpha_j, \beta_j \in \mathbb{K}$, the decompositions induced by Eq. (9.2) and Eq. (9.3) are not equivalent. We do not consider the latter case and we refer to [Hel92, Rez92, CGLM08, BCG11, Ble15], for $\mathbb{K} = \mathbb{R}$, and to [Rez96, Rez13a, RT17], $\mathbb{K} \subset \mathbb{C}$.

Main results We extend Sylvester's algorithm to achieve a nearly-optimal complexity bound in the degree of the binary form. By considering structural properties of the Hankel matrices, we restrict the possible values for the rank of the decompositions and we identify when the decompositions are unique. We build upon [Hel92] and we use the Extended Euclidean Algorithm to deduce a better complexity estimate than what was previously known. Similarly to Sylvester's algorithm, our algorithm decomposes successfully any binary form, without making any assumptions on the input.

First, we focus on *symbolic decompositions*, that is a representation of the decomposition as a sum of a rational function evaluated at the roots of a univariate polynomial (Definition 9.4.12). We introduce an algorithm to compute a symbolic decomposition of a binary form of degree D in $O(\mathtt{M}(D)\log(D))$, where $\mathtt{M}(D)$ is the arithmetic complexity of polynomial multiplication (Thm. 9.4.19). When the decomposition is unique the algorithm is deterministic and this is a worst case bound. When the decomposition is not unique, our algorithm makes some random choices to fulfill certain genericity assumptions; thus the

algorithm is a Monte Carlo one. However, we can verify if the genericity assumptions hold within the same complexity bound, that is $O(\mathtt{M}(D)\log(D))$, and hence we can also deduce a Las Vegas variant of the algorithm.

Following the standard terminology used in structured matrices [Pan01], our algorithm is *superfast* as its arithmetic complexity matches the size of the input up to poly-logarithmic factors. The symbolic decomposition allow us to approximate the terms in a decomposition, with a relative error of $2^{-\varepsilon}$, in $O\big(D\log^2(D)\big(\log^2(D)+\log(\varepsilon)\big)\big)$ arithmetic operations [Pan02, MP13]. Moreover, we can deduce for free the rank and the border rank of the tensor, see [CS11, Sec. 1].

Using results from [KY89], we bound the algebraic degree of the decompositions by $\min(\text{rank}, D-\text{rank}+1)$ (Thm. 9.4.4). Moreover, we prove lower bounds for the algebraic degree of the decomposition and we show that in certain cases the bound is tight (Sec. 9.4.1). For polynomials with integer coefficients, we bound the bit complexity, up to poly-logarithmic factors, by $\widetilde{\mathcal{O}}_B(D\ell + D^4 + D^3\tau)$, where τ is the maximum bitsize of the coefficients of the input binary form and $2^{-\ell}$ is the error of the terms in the decomposition (Thm. 9.4.21). This Boolean worst case bound holds independently of whether the decomposition is unique or not.

In this exposition we omit the presentation of our randomized algorithm, which can be found in [BFPT16]. With respect to our previous algorithm, the main algorithm that we present (Alg. 14) omits an initial linear change of coordinates as we now rely on fewer genericity assumptions. In contrast to [BFPT16] the algorithm that we present is deterministic when the decomposition is unique (Thm. 9.4.19). When the decomposition is not unique, our algorithm is still randomized but we present bounds for the number of bad choices that it could make (Prop. 9.4.5).

With respect to the algebraic degree of the problem, we propose bounds and study their tightness (Thm. 9.4.3). We introduce explicit lower bounds showing that our bounds can be tight (Sec. 9.4.1).

Organization of the chapter First we introduce the notation. In Sec. 9.2 we present the preliminaries that we need for our algorithm. We present Sylvester's algorithm (Sec. 9.2.1), we recall some properties of the structure of the kernel of the Hankel matrices (Sec. 9.2.2), we analyze its relation to rational reconstructions of series/polynomials (Sec. 9.2.3), and we present the Extended Euclidean Algorithm (Sec. 9.2.4). Later, in Sec. 9.3, we present our main algorithm to decompose binary forms (Alg. 14) and its proof of correctness (Sec. 9.3.3). This algorithm uses Alg. 15 to compute the kernel of a family of Hankel matrices, which we consider in Sec. 9.3.1. Finally, in Sec. 9.4, we study the algebraic degree of the problem (Sec. 9.4.1), we present tight bounds for it (Sec. 9.4.1), and we analyze the arithmetic (Sec. 9.4.2) and bit complexity of Alg. 14 (Sec. 9.4.3).

Notation We denote by O, respectively $\mathcal{O}_B$, the arithmetic, respectively bit, complexity and we use $\widetilde{\mathcal{O}}$, respectively $\widetilde{\mathcal{O}}_B$, to ignore (poly-)logarithmic factors. $\mathtt{M}(n)$ is the arithmetic complexity of multiplying two polynomial of degree n. Let $\mathbb{K}$ be a subfield of $\mathbb{C}$, and $\overline{\mathbb{K}}$ its algebraic closure. If $v=(v_0,\ldots,v_n)^\mathsf{T}$ then $P_v = P_{(v_0,\ldots,v_n)} := \sum_{i=0}^n v_i x^i y^{n-i}$. Given a binary form $f(x,y)$, we denote by $f(x)$ the univariate polynomial $f(x) := f(x,1)$. By $f'(x)$ we denote the derivative of $f(x)$ with respect to x. For a matrix M, $\mathtt{rk}(M)$ is its rank and $\mathtt{Ker}(M)$ its kernel.

9.2 Preliminaries

9.2.1 An algorithm based on Sylvester's theorem

Sylvester's theorem (Thm. 9.2.2) relates the minimal decomposition of a binary form to the kernel of a Hankel matrix. Moreover, it implies an (incremental) algorithm for computing the minimal decomposition. The version that we present in Alg. 12 comes from [CM96, Sec. 3.2].

Definition 9.2.1. *Given a vector $a = (a_0, \ldots, a_D)^\mathsf{T}$, we denote by $\{H_a^k\}_{1\leq k\leq D}$ the family of Hankel matrices indexed by k, where $H_a^k \in \mathbb{K}^{(D-k+1)\times(k+1)}$ and*

$$H_a^k := \begin{pmatrix} a_0 & a_1 & \cdots & a_{k-1} & a_k \\ a_1 & a_2 & \cdots & a_k & a_{k+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{D-k-1} & a_{D-k} & \cdots & a_{D-2} & a_{D-1} \\ a_{D-k} & a_{D-k+1} & \cdots & a_{D-1} & a_D \end{pmatrix}. \tag{9.4}$$

We may omit the a in H_a^k when it is clear from the context.

Theorem 9.2.2 (Sylvester, 1851)**.** *Let $f(x,y) = \sum_{i=0}^{D} \binom{D}{i} a_i x^i y^{D-i}$ with $a_i \in \mathbb{K} \subseteq \mathbb{C}$. Also, consider a non-zero $c = (c_0, \ldots, c_r)^\mathsf{T} \in \mathbb{K}^{r+1}$, such that the polynomial*

$$P_c = \sum\nolimits_{i=0}^{r} c_i\, x^i\, y^{r-i} = \prod\nolimits_{j=1}^{r} (\beta_j x - \alpha_j y)$$

is square-free and $\alpha_i, \beta_i \in \overline{\mathbb{K}}$. Then, there are $\lambda_1, \ldots \lambda_r \in \overline{\mathbb{K}}$ such that

$$f(x,y) = \sum_{j=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^D,$$

if and only if $(c_0, \ldots, c_r)^\mathsf{T} \in \mathtt{Ker}(H_a^r)$.

For a proof of Thm. 9.2.2 we refer to the work of [Rez13a, Thm. 2.1 & Cor. 2.2]. Thm. 9.2.2 implies Alg. 12. This algorithm will execute steps 2 and 3 as many times as the rank. In the i-th iteration it will compute the kernel of H^i. The dimension of this kernel is $\leq i$, and each vector in the kernel has $i+1$ coordinates. As the rank of the binary form can be as big as the degree of the binary form, a straightforward bound for the arithmetic complexity of Alg. 12 is at least cubic in the degree.

We can improve the complexity of Alg. 12 by a factor of D by noticing that the rank of the binary form is either $\mathtt{rk}(H^{\lceil \frac{D}{2} \rceil})$ or $D - \mathtt{rk}(H^{\lceil \frac{D}{2} \rceil}) + 2$ [CS11, Sec. 3] [Hel92, Thm. B]. Another way to compute the rank is by using the minors [BGI11, Alg. 2].

The bottleneck of the previous approaches is that they have to compute the kernel of a Hankel matrix. However, even if we know that the rank of the binary form is r, then the dimension of the kernel of H^r can be as big as $O(D)$; the same bound holds for the length of the vectors in the kernel. Hence, the complexity is lower bounded by $O(D^2)$.

Our approach avoids the incremental construction. We exploit the structure of the kernel of the Hankel matrices and we prove that the rank has just two possible values (Lem. 9.3.1). Moreover, we use a compact representation of the vectors in the kernel. We describe them as a combination of two polynomials of degree $O(D)$.

Algorithm 12 INCRDECOMP [CM96, Fig. 1]

1. $r := 1$
2. Get a random $c \in \texttt{Ker}(H^r)$
3. If P_c is not square-free, $r := r + 1$ and GO TO 2
4. Write P_c as $\prod_{j=1}^{r}(\beta_j x - \alpha_j y)$
5. Solve the transposed Vandermonde system:

$$\begin{pmatrix} \beta_1^D & \cdots & \beta_r^D \\ \beta_1^{D-1}\alpha_1 & \cdots & \beta_r^{D-1}\alpha_r \\ \vdots & \ddots & \vdots \\ \alpha_1^D & \cdots & \alpha_r^D \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} = \begin{pmatrix} a_0 \\ \vdots \\ a_D \end{pmatrix} \tag{9.5}$$

6. Return $\sum_{j=1}^{r} \lambda_j(\alpha_j x + \beta_j y)^D$

9.2.2 Kernel of the Hankel matrices

The Hankel matrices are one of the most well-known structured matrices [Pan01]. They are related to polynomial multiplication. We present results about the structure of their kernel. For details, we refer to [HR84, Ch. 5].

Proposition 9.2.3. *Matrix-vector multiplication of Hankel matrices is equivalent to polynomial multiplication. Given two binary forms $A := \sum_{i=0}^{D} a_i x^i y^{D-i}$ and $U := \sum_{i=0}^{k} u_i x^i y^{k-i}$, consider $R := \sum_{i=0}^{D+k} r_i x^i y^{D+k-i} = A \cdot U$. If we choose the monomial basis $\{y^{D+k}, \ldots, x^{D+k}\}$, then the equality $A \cdot U = R$ is equivalent to Eq.* (9.6)*, where the central submatrix of the left matrix is $H^k_{(a_0,\ldots,a_D)}$*

(Def. 9.2.1).

$$\left(\begin{array}{ccccc}
 & & & a_0 & a_0 \\
 & & & a_0 & a_1 \\
 & & \iddots & \iddots & \vdots \\
 & a_0 & \cdots & a_{k-2} & a_{k-1} \\
\hline
a_0 & a_1 & \cdots & a_{k-1} & a_k \\
a_1 & a_2 & \cdots & a_k & a_{k+1} \\
\vdots & \vdots & \iddots & \vdots & \vdots \\
a_{D-k} & a_{D-k+1} & \cdots & a_{D-1} & a_D \\
\hline
a_{D-k-1} & a_{D-k} & \cdots & a_D & \\
\vdots & \iddots & \iddots & & \\
a_{D-1} & a_D & & & \\
a_D & & & &
\end{array}\right)
\begin{pmatrix} u_k \\ \vdots \\ u_1 \\ u_0 \end{pmatrix}
=
\left(\begin{array}{c}
r_0 \\ r_1 \\ \vdots \\ r_{k-1} \\ \hline r_k \\ r_{k+1} \\ \vdots \\ r_D \\ \hline r_{D+1} \\ \vdots \\ r_{D+k-1} \\ r_{D+k-1}
\end{array}\right). \tag{9.6}$$

Consider a family of Hankel matrices $\{H_a^k\}_{1\leq k\leq D}$ as in Def. 9.2.1. There is a formula for the dimension of the kernel of each matrix in the family that involves two numbers, N_1^a and N_2^a. To be more specific the following holds:

Proposition 9.2.4. *For any family of Hankel matrices $\{H_a^k\}_{1\leq k\leq D}$ there are two constants, N_1^a and N_2^a, such that*

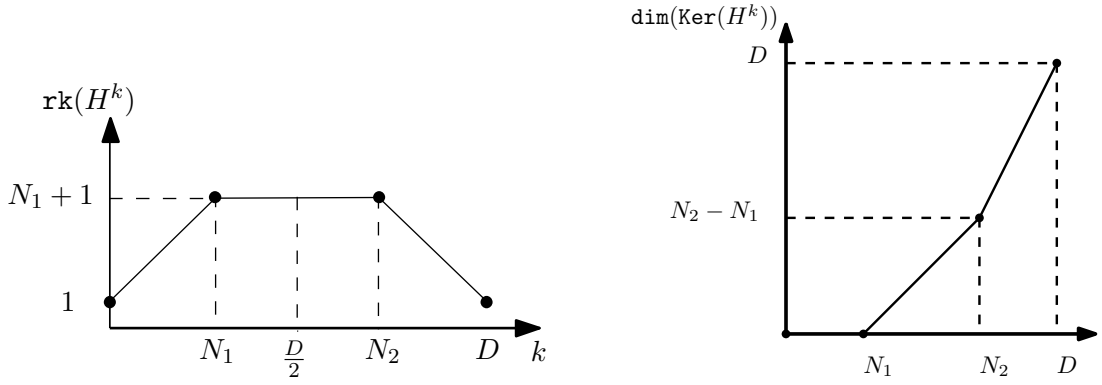
1. $0 \leq N_1^a \leq N_2^a \leq D$.
2. *For all k, $1 \leq k \leq D$, it holds* $\dim(\texttt{Ker}(H_a^k)) = \max(0; k - N_1^a) + \max(0; k - N_2^a)$.
3. $N_1^a + N_2^a = D$.

We may skip a in N_1^a and N_2^a when it is clear from the context.

Figure (9.1) illustrates Prop. 9.2.4. The dimension of the kernels and the ranks of the matrices are piecewise-linear functions in k, given by three line segments. In the case of the dimension of the kernels, it is an increasing function. For k from 1 to N_1, the kernel of the matrix is trivial, so the rank increases as the number of columns, that is, the slope of the graph of the ranks is 1 and the one of the dimension of the kernels is 0. For k from $N_1 + 1$ to N_2, the rank remains constant as for each column that we add, the dimension of the kernel increases by one. Hence, the slope of the graph of the ranks is 0 and the one of the dimension of the kernels is 1. For k from $N_2 + 1$ to D, the rank decreases because the dimension of the kernel increases by 2, and so the slope of the graph of the ranks is -1 and the one of the dimension of the kernel is 2.

If $N_1 = N_2$, the graph degenerate to two line segments. For the graph of the ranks, the first segment has slope 1 for k from 1 to $N_1 + 1$ and the second segment has slope -1 for k from $N_1 + 1$ to D. For the graph of the dimension of the kernels, the first segment has slope 0 from 1 to $N_1 + 1$, and the second one has slope 2 from N_1 to D.

The elements of the kernel of the matrices in $\{H^k\}$ are related. To express this relation from a linear algebra point of view we introduce the **U-chains**.



(a) Rank of the Hankel matrices (b) Dimension of the kernel

Figure 9.1: Relation between N_1, N_2 and D

Definition 9.2.5 ([HR84, Def. 5.1])**.** *A **U-chain** of length k of a vector $v = (v_0, \ldots, v_n)^\mathsf{T} \in \mathbb{K}^{n+1}$ is a set of vectors $\{\mathbb{U}_k^0 v, \mathbb{U}_k^1 v, \ldots, \mathbb{U}_k^{k-1} v\} \subset \mathbb{K}^{n+k}$. The i-th element, $0 \leq i \leq k-1$, is*

$$\mathbb{U}_k^i v = (\underbrace{0 \ \ldots \ 0}_{i}, \overbrace{v_0 \ \ldots \ v_n}^{n+1}, \underbrace{0 \ \ldots \ 0}_{k-1-i})$$

where $\mathbb{U}_k^i$ is a $(n+k) \times (n+1)$ i-shifting matrix [HR84, page 11].

If v is not zero, then all the elements in a U-chain of v are linearly independent. The following theorem uses the U-chains to relate the vectors of the kernels in a family of Hankel matrices.

Proposition 9.2.6 (Vectors ***v*** and ***w***)**.** *Given a family of Hankel matrices $\{H^k\}_{1 \leq k \leq D}$, let N_1 and N_2 be the constants of Prop. 9.2.4. There are two vectors, $v \in \mathbb{K}^{N_1+2}$ and $w \in \mathbb{K}^{N_2+2}$, such that,*

- *If $0 \leq k \leq N_1$, then $\mathtt{Ker}(H^k) = \{0\}$.*
- *If $N_1 < k \leq N_2$, then the U-chain of v of length $(k - N_1)$ is a basis of $\mathtt{Ker}(H^k)$, that is*

$$\mathtt{Ker}(H^k) = \langle \mathbb{U}_{k-N_1}^0 v, \ldots, \mathbb{U}_{k-N_1}^{k-N_1-1} v \rangle.$$

- *If $N_2 < k \leq D$, then the U-chain of v of length $k - N_1$ together with the U-chain of w of length $k - N_2$ is a basis of $\mathtt{Ker}(H^k)$, that is*

$$\mathtt{Ker}(H^k) = \langle \mathbb{U}_{k-N_1}^0 v, \ldots, \mathbb{U}_{k-N_1}^{k-N_1-1} v, \mathbb{U}_{k-N_2}^0 w, \ldots, \mathbb{U}_{k-N_2}^{k-N_2-1} w \rangle.$$

The vectors v and w of Prop. 9.2.6 are not unique. Vector v could be any vector in $\mathtt{Ker}(H^{N_1+1})$. Vector w could be any vector in $\mathtt{Ker}(H^{N_2+1})$ that does not belong to the vector space generated by the U-chain of v of length $N_2 - N_1 + 1$. From now on, given a family of Hankel matrices, we refer to v and w as the vectors of Prop. 9.2.6.

Let u be a vector in the kernel of H^k and P_u the corresponding polynomial (see Notation). We call P_u a **kernel polynomial**. As $P_{\mathbb{U}_k^j v} = x^j y^{k-1-j} P_v$, we can write any kernel polynomial of a family of Hankel matrices as a combination of P_v and P_w [HR84, Prop. 5.1 & 5.5]. Moreover, P_v and P_w are relative prime.

Proposition 9.2.7. *Consider any family of Hankel matrices $\{H^k\}_{1\leq k\leq D}$. Hence, the kernel polynomials P_v and P_w are relative prime. Moreover, for each k, the set of kernel polynomials of the matrix H^k is as follows:*

- *If $0 < k \leq N_1$, then it is $\{0\}$.*
- *If $N_1 < k \leq N_2$, then it is $\{P_\mu P_v : \mu \in \mathbb{K}^{k-N_1}\}$.*
- *If $N_2 < k \leq D$, then it is $\{P_\mu P_v + P_\rho P_w : \mu \in \mathbb{K}^{k-N_1}, \rho \in \mathbb{K}^{k-N_2}\}$.*

Corollary 9.2.8. *Let $\omega \in \mathtt{Ker}(H^{N_2+1})$ such that $P_\omega \notin \{P_\mu P_v : \mu \in \mathbb{K}^{N_2-N_1+1}\}$, then we can consider ω as the vector w from Prop. 9.2.6.*

9.2.3 Rational Reconstructions

A rational reconstruction for a series or a polynomial is to approximate a series/polynomial as the quotient of two polynomials. Rational reconstructions are the backbone of many problems e.g., Padé approximants, Cauchy Approximations, Linear Recurrent Sequences, Hermite Interpolation. They are related to the Hankel matrices. For details about rational reconstructions, we refer to [BCG+17, Chapter 7] and references therein.

Definition 9.2.9. *Consider $a := (a_0, \ldots, a_D)^\mathsf{T} \in \mathbb{K}^{D+1}$ and a polynomial $A := \sum_{i=0}^{D} a_i x^i \in \mathbb{K}[x]$. Given a pair of univariate polynomials (U, R), we say that they are a rational reconstruction of A modulo x^{D+1} if $A \cdot U \equiv R \mod x^{D+1}$.*

These reconstructions are not necessarily unique. We are interested in them because there is a relation between the rational reconstructions of A modulo x^{D+1} and the kernels of the family of Hankel matrices $\{H_a^k\}_k$.

Lemma 9.2.10. *Following Eq.* (9.6)*, if $\omega \in \mathtt{Ker}(H_a^k)$, then*

$$\left(\begin{array}{ccc|c} & & & a_0 \\ & & \iddots & \vdots \\ & a_0 & \cdots & a_{k-1} \\ \hline a_0 & a_1 & \cdots & a_k \\ \vdots & \vdots & \iddots & \vdots \\ a_{D-k} & a_{D-k+1} & \cdots & a_D \end{array}\right) \begin{pmatrix} \omega_0 \\ \omega_1 \\ \vdots \\ \omega_k \end{pmatrix} = \left(\begin{array}{c} r_0 \\ \vdots \\ r_{k-1} \\ \hline 0 \\ \vdots \\ 0 \end{array}\right). \tag{9.7}$$

Hence, $P_\omega(1, x) = \sum_{i=0}^{k} \omega_{k-i} x^i$ and $A \cdot P_\omega(1, x) \equiv \sum_{i=0}^{k-1} r_i x^i \mod x^{D+1}$. Therefore, $(P_\omega(1, x), \sum_{i=0}^{k-1} r_i x^i)$ is a rational reconstruction of A modulo x^{D+1}.

Lemma 9.2.11. *If (U, R) is a rational reconstruction of A of degree D, then there is a vector $\omega \in \texttt{Ker}(H_a^{\max(\deg(U),\deg(R)+1)})$ such that*

$$P_\omega = U\Big(\frac{y}{x}\Big)\, x^{\max(\deg(U),\deg(R)+1)}.$$

Proof. Let $k = \deg(U)$, $q = \deg(R)$, $U = \sum_i u_i x^i$ and $R = \sum_i r_i x^i$. Following Eq. (9.6), $A\,U \equiv R \mod x^{D+1}$ is equivalent to,

$$\begin{pmatrix} & & & a_0 \\ & & \iddots & \vdots \\ & a_0 & \cdots & a_{k-1} \\ a_0 & a_1 & \cdots & a_k \\ \vdots & \vdots & \iddots & \vdots \\ a_{D-k} & a_{D-k+1} & \cdots & a_D \end{pmatrix} \begin{pmatrix} u_k \\ \vdots \\ u_1 \\ u_0 \end{pmatrix} = \begin{pmatrix} r_0 \\ r_1 \\ \vdots \\ r_q \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{9.8}$$

If $k > q$, Eq. (9.8) reduces to Eq. (9.7), and so $\omega = (u_k, \ldots, u_0) \in \texttt{Ker}(H_a^k)$. Moreover,

$$U\Big(\frac{y}{x}\Big)\, x^k = \sum_{i=0}^{k} u_i y^i x^{k-i} =_{(j \leftrightarrow k-i)} \sum_{j=0}^{k} u_{k-j} x^j y^{j-k} = P_\omega.$$

If $q \geq k$, we extend the vector $(u_k, \ldots, u_0)$ by adding $(q+1-k)$ leading zeros. We rewrite Eq. (9.8) as Eq. (9.9). The two bottom submatrices form the matrix H_a^{q+1}, and so $\omega = (0, \ldots, 0, u_k, \ldots, u_0) \in \texttt{Ker}(H_a^{q+1})$. Also, $P_\omega = \sum_{j=0}^{k} u_j x^{q+1-j} y^j + \sum_{j=k+1}^{q+1} 0\, x^{q+1-j} y^j = U\left(\frac{y}{x}\right) x^{q+1}$.

$$\left[\begin{array}{ccc|ccc} & & & & & a_0 \\ & & & & \iddots & \vdots \\ & & & a_0 & \cdots & a_k \\ & & a_0 & a_1 & \cdots & a_{k+1} \\ & \iddots & \iddots & \vdots & \iddots & \vdots \\ \hline a_0 & \cdots & a_{q-k} & a_{q+1-k} & \iddots & a_{q+1} \\ \vdots & \iddots & \vdots & \vdots & \iddots & \vdots \\ a_{D-q-1} & \cdots & a_{D-k-1} & a_{D-k} & \cdots & a_D \end{array}\right] \left(\begin{array}{c} 0 \\ \vdots \\ 0 \\ \hline u_k \\ \vdots \\ u_0 \end{array}\right) = \left(\begin{array}{c} r_0 \\ \vdots \\ r_k \\ \vdots \\ r_q \\ \hline 0 \\ \vdots \\ 0 \end{array}\right). \tag{9.9}$$

□

Remark 9.2.12. *If (U, R) is a rational reconstruction, then the degree of the kernel polynomial $P_\omega(x, y) = U\big(\frac{y}{x}\big)\, x^{\max(\deg(U),\deg(R)+1)}$ is $\max(\deg(U), \deg(R) + 1)$. In particular, the maximum power of x that divides the kernel polynomial P_ω is $x^{\max(0,\deg(R)+1-\deg(U))}$.*

9.2.4 Greatest Common Divisor and Bézout identity

The Extended Euclidean algorithm (EGCD) is a variant of the classical Euclidean algorithm that computes the Greatest Common Divisor of two univariate polynomials A and B, $\gcd(A, B)$, together with two polynomials U and V, called *cofactors*, such that $U\,A + V\,B = gcd(A, B)$. In the process of computing these cofactors, the algorithm computes a sequence of relations between A and B that are useful to solve various problems, in particular to compute the rational reconstruction of A modulo B. For a detailed exposition we refer to [BCG+17, Ch. 6] and [GG13, Ch. 3 and 11].

Algorithm 13 Calculate the EGCD of A and B

$(U_0, V_0, R_0) \leftarrow (0, 1, B)$
$(U_1, V_1, R_1) \leftarrow (1, 0, A)$
$i \leftarrow 1$
while $R_i \neq 0$ **do**
  $Q_{i-1} \leftarrow R_{i-2}$ quo R_{i-1}
  $(U_i, V_i, R_i) \leftarrow (U_{i-2}, V_{i-2}, R_{i-2}) - Q_{i-1}\,(U_{i-1}, V_{i-1}, R_{i-1})$
  $i \leftarrow i + 1$
end while
Return $\{(U_j, V_j, R_j)\}_j$

The Extended Euclidean Algorithm (Alg. 13) computes a sequence of triples $\{(U_i, V_i, R_i)\}_i$ which form the identities

$$U_i\,A + V_i\,B = R_i, \quad \text{for all } i. \tag{9.10}$$

Following [GG13], we refer to these triplets as the rows of the Extended Euclidean algorithms of A and B. Besides Eq. (9.10), the rows are related to each other as follows: the degrees of R_i form a strictly decreasing sequence, U_i and V_i are coprime, and we can deduce the degree of U_i from the one of R_{i-1}.

Remark 9.2.13. *The degrees of the polynomials $\{R_i\}_i$ form an strictly decreasing sequence, that is $\deg(R_i) > \deg(R_{i+1})$ for every i.*

Lemma 9.2.14 ([BCG+17, Sec 7.1])**.** *For each i, $U_i\,V_{i+1} - U_{i+1}\,V_i = (-1)^i$, and so the polynomials U_i and V_i are coprime.*

Lemma 9.2.15 ([BCG+17, Lem 7.1])**.** *For each $i > 0$, the degree of U_i is the degree of B minus the degree of R_{i-1}, that is*

$$\deg(U_i) = \deg(B) - \deg(R_{i-1}), \quad \forall i > 0.$$

Moreover, every row of the Extended Euclidean Algorithm is a rational reconstruction of A modulo B.

Remark 9.2.16. *For each $i \geq 0$, (U_i, R_i) is a rational reconstruction of A modulo B.*

9.3 The Algorithm

One of the drawbacks of Alg. 12, and its variants, is that they rely on the computation of the kernels of many Hankel matrices and they ignore the particular structure that is present. Using Lem. 9.3.1, we can skip many calculations by computing only two vectors, v and w (Prop. 9.2.6). This is the main idea behind Alg. 14 that leads to a softly-linear arithmetic complexity bound (Sec. 9.4.2).

Alg. 14 performs as follows: First, step 1 computes two kernel polynomials, P_v and P_w using Prop. 9.2.7, to obtain the kernel polynomials of the Hankel matrices (see Sec. 9.3.1). Then, step 2 computes a square-free kernel polynomial of the minimum degree r (see Sec. 9.3.2). Next, step 3 computes the coefficients $\lambda_1, \ldots, \lambda_r$ (see Sec. 9.4.1). Finally, step 4 recovers a decomposition for the original binary form.

Let f be a binary form as in Eq. (9.1) and let $\{H^k\}_{1 \leq k \leq D}$ be its corresponding family of Hankel matrices (see Def. 9.2.1). The next lemma establishes the rank of f.

Lemma 9.3.1. *Assume f, $\{H^k\}_k$, N_1 and N_2 of Prop. 9.2.4, and v and w of Prop. 9.2.6. If P_v (Prop. 9.2.7) is square-free then the rank of f is $N_1 + 1$, else, it is $N_2 + 1$.*

Proof. By Prop. 9.2.4, for $k < N_1 + 1$, the kernel of H^k is trivial. Hence, by Sylvester's theorem (Thm. 9.2.2), there is no decomposition with a rank smaller than $N_1 + 1$. Recall that $v \in \mathtt{Ker}(H^{N_1+1})$. So, if P_v is square-free, by Sylvester's theorem, there is a decomposition of rank $N_1 + 1$.

Assume P_v is not square-free. For $N_1 + 1 \leq k \leq N_2$, P_v divides all the kernel polynomials of the matrices H^k (Prop. 9.2.7). Therefore, none of them is square-free, and so the rank is at least $N_2 + 1$.

By Prop. 9.2.7, P_v and P_w do not share a root. So, there is a polynomial P_μ of degree $N_2 - N_1$ such that $Q_\mu := P_v P_\mu + P_w$ is square-free. A formal proof of this appears in Thm. 9.3.6. By Prop. 9.2.7, Q_μ is a square-free kernel polynomial of degree $N_2 + 1$. Consequently, by Sylvester's theorem, there is a decomposition with rank $N_2 + 1$. □

To relate Lem. 9.3.1 with the theory of binary form decomposition, we recall that the decompositions are identified with the square-free polynomials in the annihilator of f [KR84];[IK99, Chp. 1]. All the kernel polynomials of $\{H_k\}_k$ belong to the annihilator of f, which is an ideal. If f is a binary form of degree $D = 2k$ or $2k + 1$, then this ideal is generated by two binary forms of degrees $\mathtt{rk}(H^k)$ and $D + 2 - \mathtt{rk}(H^k)$, with no common zeros [IK99, Thm. 1.44]. These are the polynomials P_v and P_w. Using this interpretation, Alg. 12, and its variants, computes a (redundant) generating set of the annihilator, while Alg. 14 computes a basis.

9.3.1 Computing the polynomials P_v and P_w

We use Lem. 9.2.10 and Lem. 9.2.11 to compute the polynomials P_v and P_w from Prop. 9.2.7 as a rational reconstruction of $A := \sum_{i=0}^{D} a_i x^i$ modulo x^{D+1}. Our algorithm exploits the Extended Euclidean Algorithm in a similar way to [CC86] for computing scaled Padé fractions.

In the following, let v be the vector of Prop. 9.2.6, consider $U_v := P_v(1, x)$ and $R_v \in \mathbb{K}[x]$ as the remainder of the division of $(A \cdot P_v(1, x))$ by x^{D+1}. Note that, the polynomial R_v is the unique polynomial of degree smaller to $N_1 + 1$ such that $A \cdot P_v(1, x) \equiv R_v \mod x^{D+1}$.

Lemma 9.3.2. *If (U, R) is a rational reconstruction of A modulo x^{D+1} such that $\max(\deg(U), \deg(R)+1) \leq N_2$, then there is a polynomial $Q \in \mathbb{K}[x]$ such that $Q \cdot P_v(x, 1) = U$ and $Q \cdot R_v = R$.*

Algorithm 14 FASTDECOMP

Input: A binary form $f(x, y)$ of degree D
Output: A decomposition for $f(x, y)$ of rank r.

1. **Compute $\mathbf{P_v}$ and $\mathbf{P_w}$ of $\{\mathbf{H_a^k}\}_\mathbf{k}$**
   We use Alg. 15 with $(a_0, \ldots, a_D)$.
2. **IF $\mathbf{P_v}(\mathbf{x}, \mathbf{y})$ is square-free**,
   $Q \longleftarrow P_v$
   $r \longleftarrow N_1 + 1$ {The *rank* of the decomposition is the degree of Q}
   ELSE
   Compute a square-free binary form Q
   We compute a vector μ of length $(N_2 - N_1 + 1)$,
   such that $(P_\mu P_v + P_w)$ is square-free (Sec. 9.4.1).
   $Q \longleftarrow P_\mu P_v + P_w$
   $r \longleftarrow N_2 + 1$ {The *rank* of the decomposition is the degree of Q}
3. **Compute the coefficients $\lambda_\mathbf{1}, \ldots, \lambda_\mathbf{r}$**
   Solve the system of Eq. (9.5) where $Q(x, y) = \prod_{j=1}^{r} (\beta_j x - \alpha_j y)$.
   For details and the representation of λ_j, see Sec. 9.4.1.
4. **Return** $f(x, y) = \sum_{j=1}^{r} \lambda_j (\alpha_j x + \beta_j y)^D$

Proof. Let $k := \deg(U)$ and $q := \deg(R)$. By Lem. 9.2.11, there is a vector nonzero $\omega \in \mathtt{Ker}(H_a^{\max(k,q+1)})$ such that the kernel polynomial P_ω is equal to $U\left(\frac{y}{x}\right) x^{\max(k,q+1)}$. Hence, $\mathtt{Ker}(H_a^{\max(k,q+1)}) \neq 0$ and so, by Prop. 9.2.6, $N_1 < \max(k, q+1)$. We assume that $\max(k, q+1) \leq N_2$, hence the degree of P_ω is smaller or equal to N_2 and, by Prop. 9.2.7, P_ω is divisible by P_v. Therefore, there is a polynomial $\bar{Q} \in \mathbb{K}[x, y]$ such that $\bar{Q} P_v = P_\omega$. Let $Q := \bar{Q}(1, x)$. By definition, $U_v = P_v(1, x)$ and $U = P_\omega(1, x)$, so $U = Q\, U_v$. Hence, $Q\, R_v \equiv R \mod x^{D+1}$, because $R_v \equiv U_v\, A \mod x^{D+1}$ and $Q\, R_v \equiv Q\, U_v\, A \equiv U A \equiv R \mod x^{D+1}$. If the degrees of $(Q\, R_v)$ and R are smaller than $D + 1$, then $Q\, R_v = R$, as we want to prove. By assumption, $\deg(R) < N_2 \leq D$ and $\deg(U_v\, Q) = \deg(U) \leq N_2$. By Lem. 9.2.10, the degree of R_v is upper bounded by N_1, and so $\deg(Q\, R_v) \leq \deg(U_v\, Q\, R_v) \leq N_2 + N_1 = D$ (Prop. 9.2.4). □

We can use this lemma to recover the polynomial P_v from certain rational reconstructions.

Corollary 9.3.3. *If (U, R) is a rational reconstructions of A of degree D such that $\max(\deg(U), \deg(R) + 1) \leq N_2$ and for every polynomial Q of degree bigger than zero that divides U and R, $(\frac{U}{Q}, \frac{R}{Q})$ is not a rational reconstruction of A, then there is a non-zero constant c such that*

$P_v = c \cdot U\left(\frac{y}{x}\right) x^{\max(\deg(U),\deg(R)+1)}$ *(Prop. 9.2.7). In particular,* $N_1 = \max(\deg(U)-1, \deg(R))$.

Proof. By Lem. 9.3.2, there is a $Q \in K[x]$ such that $Q \cdot (U_v, R_v) = (U, R)$. By Lem. 9.2.10, (U_v, R_v) is a rational reconstruction, and so $\deg(Q) = 0$. Hence, $N_1 + 1 = \deg(P_v) = \max(\deg(U), \deg(R)+1)$ and $Q \cdot P_v(1, \frac{y}{x}) x^{N_1+1} = U(\frac{y}{x}) x^{N_1+1}$. □

If (U, R) is a rational reconstruction of A modulo x^{D+1} such that $\deg(U) + \deg(R) \leq D$ and $U(0) = 1$, then $\frac{R}{U}$ is the Padé approximant of A of type $(\deg(R), \deg(U))$ [BCG$^+$17, Sec. 7.1]. When this Padé approximant exists, it is unique, meaning that for any rational reconstruction with this property the quotient $\frac{R}{U}$ is unique (we can invert $U \mod x^{D+1}$ because $U(0) = 1$). When $N_1 < N_2$, we have that $\frac{D+1}{2} \leq N_2$ (Prop. 9.2.4) and so, if the the Padé approximant of A of type $(\frac{D+1}{2} - 1, \frac{D+1}{2})$ exists, by Lem. 9.3.2, we can recover P_v from it. The existence of this Padé approximant is equivalent to the condition $U_v(0) = 1$, which means $v_{N_1+1} = 1$. In the algorithm proposed in our conference paper [BFPT16, Alg. 3], we needed to assume this condition to prove its correctness. In that version, we ensured this property with a generic linear change of coordinates in the original polynomial. In this version, we skip this assumption. Following [BCG$^+$17, Thm. 7.2], when $N_1 < N_2$, we can compute v no matter the value of v_{N_1+1}. This approach has a softly-linear arithmetic complexity and involves the computation of a row of the EGCD of A and x^{D+1}. We can compute P_w from a consecutive row.

Before going into the proof, we study the case $N_1 = N_2$. In this case, there are not rational reconstructions with the prerequisites of Lem. 9.3.2, and so we treat it in a different way.

Lemma 9.3.4. *If* $N_1 = N_2$*, there is a unique rational decomposition* (U, R) *such that* $\deg(U) \leq \frac{D}{2}$, $\deg(R) \leq \frac{D}{2}$ *and* R *is monic. In particular,* $\deg(R) = \frac{D}{2}$ *and we can consider the kernel polynomial related to* $v \in \mathtt{Ker}(H^{N_1+1})$ *(Prop. 9.2.6), as* $P_v = U\left(\frac{y}{x}\right) x^{\frac{D}{2}+1}$.

Proof. First note that, as $D = N_1 + N_2$ (Prop. 9.2.4), then $N_1 = \frac{D}{2}$. Following Eq. (9.6), if we write $U = \sum_{i=0}^{N_1} u_i x^i$ and $R = \sum_{i=0}^{N_1} u_i x^i$, then we get the linear system,

$$\left(\begin{array}{cccc} & & & a_0 \\ & & a_0 & a_1 \\ & \iddots & \iddots & \vdots \\ \hline a_0 & \cdots & a_{N_1-1} & a_{N_1} \\ a_1 & \cdots & a_{N_1} & a_{N_1+1} \\ \vdots & \iddots & \vdots & \vdots \\ a_{D-N_1} & \cdots & a_{D-1} & a_D \end{array}\right) \begin{pmatrix} u_{N_1} \\ \vdots \\ u_1 \\ u_0 \end{pmatrix} = \left(\begin{array}{c} r_0 \\ \vdots \\ r_{N_1-1} \\ \hline r_{N_1} \\ 0 \\ \vdots \\ 0 \end{array}\right)$$

The matrix $H^{N_1} \in \mathbb{K}^{(D-\frac{D}{2}+1)\times(\frac{D}{2}+1)}$ is square and, as $\mathtt{Ker}(H^{N_1}) = 0$ (Prop. 9.2.6), it is invertible. If $r_{N_1} = 0$, that is $\deg(R_v) < N_1$, then the polynomial U is zero. Hence, if R is monic, then $r_{N_1} = 1$, and so we compute the coefficients of U and R as

$$\begin{pmatrix} u_{N_1} \\ \vdots \\ u_1 \\ u_0 \end{pmatrix} = (H^{N_1+1})^{-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} r_0 \\ \vdots \\ r_{N_1-1} \\ 1 \end{pmatrix} = \begin{pmatrix} & & & a_0 \\ & & a_0 & a_1 \\ & \iddots & \iddots & \vdots \\ a_0 & \cdots & a_{N_1-1} & a_{N_1} \end{pmatrix} \begin{pmatrix} u_{N_1} \\ u_{N_1-1} \\ \vdots \\ u_0 \end{pmatrix}$$

□

Lemma 9.3.5 (Correctness of Alg. 15)**.** *Let $\{(U_j, V_j, R_j)\}_j$ be the set of triplets obtained from the Extended Euclidean Algorithm for the polynomials A and x^{D+1}. Let i be the index of the first row of the extended Euclidean algorithm such that $\deg(R_i) < \frac{D+1}{2}$. Then, we can compute the polynomials P_v and P_w of Prop. 9.2.7 as*

(A) $P_v = U_i(\frac{x}{y}) \cdot x^{\max(\deg(U_i),\deg(R_i)+1)}$.

(B) *If* $\deg(R_i) > \deg(U_i)$, $P_w = U_{i+1}(\frac{x}{y}) \cdot x^{\deg(U_{i+1})}$.

(C) *If* $\deg(R_i) \leq \deg(U_i)$, $P_w = U_{i-1}(\frac{x}{y}) \cdot x^{\deg(R_{i-1}+1)}$.

Proof. **(A).** First observe that if i is the first index such that the degree of R_i is strictly smaller than $\frac{D+1}{2}$, then the degree of R_{i-1} has to be bigger or equal to $\frac{D+1}{2}$. Hence, the degree of U_i is smaller or equal to $\frac{D+1}{2}$, because by Lem. 9.2.15, $\deg(U_i) = D+1-\deg(R_{i-1}) \leq D+1-\frac{D+1}{2} = \frac{D+1}{2}$. We can consider R_{i-1} because, as the degree of $R_0 = x^{D+1}$ is $D+1 > \frac{D+1}{2}$, $i > 0$.

If $N_1 = N_2$, then D is even and $N_1 = \frac{D}{2}$ (Prop. 9.2.4). As $\lfloor \frac{D+1}{2} \rfloor = \frac{D}{2}$, $\deg(R_i) \leq \frac{D}{2}$ and $\deg(U_i) \leq \frac{D}{2}$. By Lem. 9.3.4, $\max(\deg(U_i), \deg(R_i)+1) = N_1 + 1$ and we can consider P_v as $U_i(\frac{y}{x})x^{N_1+1}$.

If $N_1 < N_2$, assume that there is a non-zero $Q \in \mathbb{K}[x]$ such that Q divides U_i and R_i and $(\frac{U_i}{Q}, \frac{R_i}{Q})$ is a rational reconstruction of A modulo x^{D+1}. Then, $\frac{U_i}{Q}A \equiv \frac{R_i}{Q} \mod x^{D+1}$ and so there is a polynomial $\bar{V}$ such that $\bar{V}x^{D+1} + \frac{U_i}{Q}A = \frac{R_i}{Q}$. Multiplying by Q, we get the equality $Q\bar{V}x^{D+1} + U_iA = R_i$. Consider the identity $V_ix^{D+1} + U_iA = R_i$ from Eq. (9.10). Coupling the two equalities together, we conclude that $V_i = Q\bar{V}$. As Q divides U_i and V_i, which are coprime (Lem. 9.2.14), Q is a constant, $\deg(Q) = 0$. If $N_1 < N_2$, then $D < 2N_2$ (Prop. 9.2.4) and so $\max(\deg(U_i), \deg(R_i)+1) \leq \frac{D+1}{2} \leq N_2$. Hence, by Lem. 9.3.2, we can consider $U_i(\frac{y}{x})x^{\max(\deg(U_i),\deg(R_i)+1)}$ as the kernel polynomial P_v from Prop. 9.2.7, related to $\texttt{Ker}(H^{N_1+1})$.

(B). Assume that the degree of U_i is strictly bigger than the one of R_i, $\deg(U_i) > \deg(R_i)$. Then $N_1 = \deg(U_i) - 1$, as $\deg(U_i) = \deg(P_v) = N_1 + 1$ (Rem. 9.2.12). Note that in this case $i > 1$ as $U_1 = 1$ and $R_1 = A$ is a nonzero polynomial, and so $\deg(U_1) \leq \deg(R_1)$. The degree of R_{i-1} is N_2 because, by Lem. 9.2.15, $\deg(R_{i-1}) = D+1-\deg(U_i) = D+1-N_1-1 = N_2$ (Prop. 9.2.4). Consider the degree of U_{i-1}. By Lem. 9.2.15, $\deg(U_{i-1}) = D+1-\deg(R_{i-2})$. As $\deg(R_{i-2}) > \deg(R_{i-1})$ (Rem. 9.2.13), then $\deg(R_{i-2}) > N_2$. Therefore, the degree of U_{i-1} is smaller or equal to the one of R_{i-1} because

$$\deg(U_{i-1}) = D+1-\deg(R_{i-2}) < D+1-N_2 = N_1+1, \text{ and so}$$

$$\deg(U_{i-1}) \leq N_1 \leq N_2 = \deg(R_{i-1}).$$

Hence, by Rem. 9.2.16, (U_{i-1}, R_{i-1}) is a rational reconstruction of A modulo x^{D+1} such that $\deg(U_{i-1}) \leq N_1$ and $\deg(R_{i-1}) = N_2$. So, $\max(\deg(U_{i-1}), \deg(R_{i-1})+1) = N_2 + 1$ and, by Rem. 9.2.12, there is a kernel polynomial $P_\omega = U_{i-1}(\frac{y}{x})x^{N_2+1}$ of degree N_2+1 such that $x^{N_2+1-\deg(U_{i-1})}$ divides P_ω. As $\deg(U_{i-1}) \leq N_1$, $x^{N_2+1-N_1}$ divides $x^{N_2+1-\deg(U_{i-1})}$ and so, it divides P_ω. We assumed that the degree of U_i is strictly bigger than the one of R_i, and so x does not divide

P_v (Rem. 9.2.12). Hence, there is no binary form Q of degree $N_2 - N_1$ such that $x^{N_2-N_1+1}$ divides $Q\,P_v$. Therefore, by Cor. 9.2.8, we can consider $P_w = P_\omega$.

(C). Assume that the degree of R_i is bigger or equal to the one of U_i, $\deg(R_i) \geq \deg(U_i)$. Hence, $\deg(R_i) + 1 = \deg(P_v) = N_1 + 1$ (Rem. 9.2.12), and so $\deg(R_i) = N_1$. In particular, $R_i \neq 0$, and so the $(i+1)$-th row of the Extended Euclidean Algorithm, $(U_{i+1}, V_{i+1}, R_{i+1})$, is defined. The degree of U_{i+1} is $N_2 + 1$, because by Rem. 9.2.12, $\deg(U_{i+1}) = D + 1 - \deg(R_i) = N_2 + 1$ (Prop. 9.2.4). The degree of R_{i+1} is strictly smaller than the one of R_i (Rem. 9.2.13), which is N_1. Hence, the degree of R_{i+1} is smaller than the degree of U_{i+1} because $\deg(R_{i+1}) < N_1 \leq N_2 < \deg(U_{i+1})$. Therefore, $P_\omega = U_{i+1}(\frac{y}{x})x^{N_2+1}$ is a kernel polynomial in $\mathtt{Ker}(H^{N_2+1})$ (Lem. 9.2.11). By Rem. 9.2.12, as $\deg(R_{i+1}) < \deg(U_{i+1})$, x does not divide P_ω. Also, the maximal power of x that divides P_v is $x^{\deg(R_i)+1-\deg(U_i)}$, and, as we assumed $\deg(R_i) \geq \deg(U_i)$, x divides P_v. Hence, every polynomial in $\{Q\,P_v : \deg(Q) = N_2 + N_1\}$ is divisible by x, and so, by Cor. 9.2.8, we can consider $P_w = P_\omega$. □

Algorithm 15 COMPUTE_PV_AND_PW

Input: A sequence $(a_0, \ldots, a_D)$.
Output: Polynomials P_v and P_w as 9.2.7.

1. $i \leftarrow$ first row of EGCD$(x^{D+1}, \sum_{i=0} a_i x^i)$ such that $R_i < \frac{D+1}{2}$.
2. $P_v \leftarrow U_i(\frac{x}{y}) \cdot x^{\max(\deg(U_i),\deg(R_i)+1)}$.
   $N_1 \leftarrow \max(\deg(U_i) - 1, \deg(R_i))$
3. **IF** $\deg(R_i) > \deg(U_i)$,
   $P_w \leftarrow U_{i+1}(\frac{x}{y}) \cdot x^{\deg(U_{i+1})}$.
   $N_2 \leftarrow \deg(U_{i+1}) - 1$.
   ELSE
   $P_w \leftarrow U_{i-1}(\frac{x}{y}) \cdot x^{\deg(R_{i-1}+1)}$.
   $N_2 \leftarrow \deg(R_{i-1})$.
4. **Return** P_v and P_w

9.3.2 Computing a square-free polynomial Q

We can compute Q at step 2 of Alg. 14 in different ways. If P_v is square-free, then we set Q equal to P_v. If P_v is not square-free, by Lem. 9.3.1, we need to find a vector $\mu \in \mathbb{K}^{(N_2-N_1+1)}$ such that $Q_\mu := P_\mu \times P_v + P_w$ is square-free. By Prop. 9.2.7, P_v and P_w are relative prime. Thus, if we take a random vector μ, generically, Q_μ would be square-free. For this to hold, we have to prove that the discriminant of Q_μ is not identically zero. To simplify notation, in the following theorem we dehomogenize the polynomials.

Theorem 9.3.6. *Given two relative prime univariate polynomials $P_v(x)$ and $P_w(x)$ of degrees N_1+1 and N_2+1 respectively, let $Q_\mu(x) := P_\mu P_v + P_w \in \mathbb{K}[\mu_0,\ldots,\mu_{N_2-N_1}][x]$. The discriminant of $Q_\mu(x)$ with respect to x is a non-zero polynomial.*

Proof. The zeros the discriminant of $Q_\mu(x)$ with respect to x over $\mathbb{K}$ correspond to the set $\{\mu \in \mathbb{K}^{N_2-N_1+1} : Q_\mu \text{ has double roots}\}$. We want to prove that the discriminant is not zero.

A univariate polynomial is square-free if and only if it does share any root with its derivative. Hence, $(\mu_0,\ldots,\mu_{N_2-N_1})^\mathsf{T} \in \{\mu \in \mathbb{K}^{N_2-N_1+1} : Q_\mu \text{ has double roots}\}$ if and only if, there is $(\mu_0,\ldots,\mu_{N_2-N_1},\alpha) \in \mathbb{K}^{N_2-N_1+1}\times\overline{\mathbb{K}}$ such that the following equations are satisfied

$$\begin{cases}(P_\mu P_v + P_w)(x) = 0\\ (P_\mu P_v' + P_\mu' P_v + P_w')(x) = 0.\end{cases} \tag{9.11}$$

In Eq. (9.11), μ_0 only appears in P_μ with degree 1. If we eliminate it to obtain the polynomial

$$(P_v\cdot P_\mu' + P_w')P_v - P_v'\cdot P_w$$

This polynomial is not identically 0 as P_v' does not divide P_v and P_v and P_w are relative prime. Hence, for each $(\mu_1,\ldots,\mu_{N_2-N_1})$, there is a finite number of solutions for this equation, bounded by the degree of the polynomial. Moreover, as the polynomials of Eq. (9.11) are linear in μ_0, each solution of the deduced equation is extensible to a finite number of solutions of Eq. (9.11). Hence, there is a $\mu \in \mathbb{K}^{N_2-N_1+1}$, such that Q_μ is square-free. Therefore, the discriminant of $Q_\mu(x)$ is not identically zero. □

Corollary 9.3.7. *For every vector $(\mu_1,\ldots,\mu_{N_2-N_1}) \in \mathbb{K}^{N_2-N_1}$ such that there is a $\mu_0 \in \mathbb{K}$ such that y^2 does not divides Q_μ, where $\mu = (\mu_0,\ldots,\mu_{N_2-N_1})$, there are at most $2D+2$ different values for $\mu_0 \in \mathbb{K}$ such that the polynomial $Q_\mu(x,y)$ is not square-free.*

Proof. If $Q_\mu(x,y)$ is not square-free, then it has a double root in $\mathbb{P}^1(\overline{\mathbb{K}})$. This root could be of the form $(\alpha,1) \in \mathbb{P}^1(\overline{\mathbb{K}})$ or $(1,0) \in \mathbb{P}^1(\overline{\mathbb{K}})$. We analyze separately these cases

First, we consider the polynomial $Q_\mu(x,1) \in \mathbb{K}[\mu_0,x]$. By Thm. 9.3.6, the discriminant of $Q_\mu(x,1)$ with respect to x is not zero. As $Q_\mu(x,1)$ is a polynomial of degree N_2+1 with respect to x, and of degree 1 with respect to μ_0, the degree with respect to μ_0 of the discriminant of $Q_\mu(x,1)$ with respect to x is at most $(N_2+1)+N_2 \le 2D+1$. Hence, there are at most $2D+1$ values for μ_0 such that $Q_\mu(x,y)$ has a root of the form $(\alpha,1) \in \mathbb{P}^1(\overline{\mathbb{K}})$ with multiplicity bigger than one.

The polynomial $Q_\mu(x,y)$ has a root of the form $(1,0) \in \mathbb{P}^1(\mathbb{K})$ with multiplicity bigger than one, if and only if y^2 divides $Q_\mu(x,y)$. If this happens, then the coefficients of the monomials $y\cdot x^{N_2-N_1-1}$ and $x^{N_2-N_1}$ in the polynomial $Q_\mu(x,y)$ are zero. By assumption, these coefficients are not identically zero as polynomials in μ_0. As $Q_\mu(x,y)$ is a linear polynomial with respect to μ_0, there is at most one value for μ_0 such that y^2 divides $Q_\mu(x,y)$.

Therefore, there are at most $(2D+1)+1$ values such that $Q_\mu(x,y)$ is not square-free. □

Remark 9.3.8. *The previous assumption is not restrictive. If y^2 divides Q_μ, where $\mu=(\mu_0,\ldots,\mu_{N_2-N_1})$, then y^2 does not divide $Q_{(\mu_0,\ldots,\mu_{N_2-N_1}+1)} = Q_\mu + x^{N_2+1}$ nor $Q_{(\mu_0,\ldots,\mu_{N_2-N_1-1}+1,\mu_{N_2-N_1})} = Q_\mu + yx^{N_2}$. Moreover, if $N_2-N_1 \ge 2$, y^2 divides (or not) $Q_\mu(x,y)$ regardless the value of μ_0. Conversely, if $N_2-N_1<2$, there is always a μ_0 such that y^2 does not divide Q_μ.*

9.3.3 Correctness of Algorithm 14

For computing a decomposition for a binary form f, we need to compute the kernel of a Hankel matrix (Thm. 9.2.2). Alg. 15 computes correctly the polynomials P_v and P_w that characterize the kernels of the family of Hankel matrices associated to f. Once we obtain these polynomials step 2 (see Cor. 9.3.7) and step 3 computes the coefficients $\alpha_j, \beta_j, \lambda_j$ of the decomposition. Hence, we have a decomposition for f, as $f(x,y) = \sum_{j=1}^{r} \lambda_j (\alpha x + \beta y)^D$.

Example Consider $f(x,y) = y^4 + 8xy^3 + 18x^2y^2 + 16x^3y + 5x^4$. The family of Hankel matrices associated to f are related to the vector $a := (1,2,3,4,5)^\mathsf{T}$, it is denoted by $\{H_a^k\}_k$, and it contains the following matrices:

$H_a^1 = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \\ 4 & 5 \end{pmatrix}$	$H_a^2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$	$H_a^3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{pmatrix}$	$H_a^4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \end{pmatrix}$

The kernel H_a^1 is trivial, so we compute the one of H_a^2. This kernel is generated by the vector $(1,-2,1)^\mathsf{T}$, so by Prop. 9.2.6 we consider $v = (1,-2,1)^\mathsf{T}$. Also, by Prop. 9.2.4, $N_1 + 1 = 2$ and $N_2 = D - N_1 = 3$. The kernel polynomial $P_v = y^2 - 2x\,y + x^2 = (x-y)^2$ is not square-free thus, by Lem. 9.3.1, the rank of $f(x,y)$ is $N_2 + 1 = 4$ and we have to compute the kernel polynomial P_w in the kernel of H_a^4. Following Prop. 9.2.6, the kernel of H_a^4 is generated by U-chain of v given vectors $\mathbb{U}_2^0 v = (1,-2,1,0,0)^\mathsf{T}$, $\mathbb{U}_2^1 v = (0,1,-2,1,0)^\mathsf{T}$, and $\mathbb{U}_2^2 v = (0,0,1,-2,1)^\mathsf{T}$, plus a vector w linear independent with this U-chain. We consider the vector $w = (0,0,0,5,-4)$, which fulfill that assumption. Hence, $P_v = y^2 - 2x\,y + x^2$ and $P_w = 5yx^3 - 4x^4$.

We proceed by computing a square-free polynomial combination of P_v and P_w. For that, we choose

$$Q := (44y^2 + 11yx + 149x^2)\,P_v + 36P_w = (5x - 11y)(x - 2y)(x + 2y)(x + y).$$

Finally, we solve the system given by the transposed of a Vandermonde matrix,

$$\begin{pmatrix} 5^4 & 1 & 1 & 1 \\ 11 \cdot 5^3 & 2 & -2 & -1 \\ 11^2 \cdot 5^2 & 2^2 & (-2)^2 & (-1)^2 \\ 11^3 \cdot 5 & 2^3 & (-2)^3 & (-1)^3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}. \tag{9.12}$$

The unique solution of the system is $(-\frac{1}{336}, 3, \frac{1}{21}, \frac{3}{16})^\mathsf{T}$, and so we recover the decomposition

$$f(x,y) = -\frac{1}{336}(11x + 5y)^4 + 3(2x + y)^4 + \frac{1}{21}(-2x + y)^4 - \frac{3}{16}(-x + y)^4.$$

Instead of considering incrementally the matrices in the Hankel family we can compute the polynomials P_v and P_w faster by applying Alg. 15. For this, we consider the polynomial $A := 5x^4 + 4x^3 + 3x^2 + 2x + 1$, and the rows of the Extended Euclidean Algorithm for A and x^5.

j	V_j	U_j	R_j
0	1	0	x^5
1	0	1	$5x^4+4x^3+3x^2+2x+1$
2	1	$\frac{1}{25}(5x-4)$	$\frac{1}{25}(x^3+2x^2+3x+4)$
3	$-25(5x-6)$	$25(x-1)^2$	25
4	$\frac{1}{25}(5x^4+4x^3+3x^2+2x+1)$	$-\frac{1}{25}x^5$	0

We need to consider the first j such that $\deg(R_j) < \frac{5}{2}$, which is $j = 3$. Hence, $N_1 = \max(\deg(U_3) - 1, \deg(R_3)) = 1$ and

$$P_v := U_3\left(\frac{y}{x}\right) x^{\max(\deg(U_3),\deg(R_3)+1)} = 25\left(\frac{y}{x}-1\right)^2 x^2 = 25(y-x)^2.$$

As $\deg(R_3) \leq \deg(U_3)$, we consider $N_2 = \deg(R_2) = 3$ and

$$P_w := U_2\left(\frac{x}{y}\right) x^{\deg(R_2)+1} = \frac{1}{25}(5yx^3 - 4x^4). \qquad \Diamond$$

9.4 Complexity

In this section we study the algebraic degree of the parameters that appear in the decomposition of a binary form and the arithmetic and bit complexity of Alg. 14.

9.4.1 Algebraic degree of the problem

If we assume that the coefficients of the input binary form Eq. (9.1) are rational numbers then the parameters of the decompositions, α_j, β_j, and λ_j (see Eq. (9.3)), are algebraic numbers, that is roots of univariate polynomials with integer coefficients. The minimum degree of this polynomials is the algebraic degree of the problem. We refer the interested reader to [Baj88, NRS10, DHO+16] for a detailed exposition about the algebraic degree and how it address the complexity of the problem at hand at a fundamental level.

The complexity of computing Q

Recall that from Lem. 9.3.1 the rank of f could be either $N_1 + 1$ of $N_2 + 1$. When the polynomial P_v is square-free, then the rank is $N_1 + 1$ and the $Q = P_v$. Following the discussion of Sec. 9.3.2, we prove that, when the rank of the binary form is $N_2 + 1$, we can compute a square-free kernel polynomial Q of this degree such that the largest degree of its irreducible factors is N_1. Moreover, we prove that for almost all the choices of $(N_2 - N_1 + 1)$ different points in $\mathbb{P}^1(\mathbb{K})$ (the projective space of $\mathbb{K}$) there is a square-free kernel polynomial of H^{N_2+1} which vanish on these points. This will be our choice for Q.

Lemma 9.4.1. *Let f be a binary form of rank N_2+1. Given (N_2-N_1+1) different points $(\alpha_0, \beta_0), \ldots, (\alpha_{N_2-N_1+1}, \beta_{N_2-N_1+1}) \in \mathbb{P}^1(\mathbb{K})$ such that none of them is a root of P_v, then there is a unique binary form P_μ of degree $N_2 - N_1$, such that the kernel polynomial $Q_\mu := P_\mu P_v + P_w$ vanish on those points.*

Proof. Without loss of generality, we assume $\beta_i = 1$. By Prop. 9.2.7, for any polynomial P_μ of degree $N_2 - N_1$, Q_μ is a kernel polynomial. Since $Q_\mu(\alpha_i, 1) = 0$, we can interpolate P_μ by noticing that $P_\mu(\alpha_j, 1) = -\frac{P_w(\alpha_j,1)}{P_v(\alpha_j,1)}$.

The degree of P_μ is $(N_2 - N_1)$ and we interpolate it at $(N_2 - N_1 + 1)$ different points. Hence there is a unique interpolation polynomial P_μ. So, Q_μ is the unique kernel polynomial of H^{N_2+1} divisible by all those linear forms. □

Example (continuation) For the example of Sec. 9.3.3, we obtained the square-free kernel polynomial by choosing the points $(2, 1)$, $(-2, 1)$ and $(-1, 1) \in \mathbb{P}^1(\mathbb{K})$. If we choose other points such that Q_μ is square-free, we will obtain a different decomposition. Hence, f does not have a unique decomposition. This holds in general. ◊

Corollary 9.4.2. *A decomposition is unique if and only if the rank is $N_1 + 1$. A decomposition is not unique if and only if the rank is $N_2 + 1$.*

Theorem 9.4.3. *Let the rank of f be $N_2 + 1$. Then there is a square-free kernel polynomial Q such that the largest degree of its irreducible factors is at most N_1.*

Proof. If the rank of f is N_2+1, then for each set of N_2-N_1+1 different points in $\mathbb{P}^1(\overline{\mathbb{K}})$, following the assumptions of Lem. 9.4.1, there is a unique kernel polynomial. There is a rational map that realizes this relation (see the proof of Lem. 9.4.1). Let this map be $Q_{[\overline{\alpha}]}$, where $\overline{\alpha} = ((\alpha_0, \beta_0), \ldots, (\alpha_{N_2-N_1}, \beta_{N_2-N_1})) \in \mathbb{P}^1(\overline{\mathbb{K}})^{N_2-N_1+1}$. The image of the map is contained in $\{P_\mu P_v + P_w : \mu \in \overline{\mathbb{K}}^{N_2-N_1+1}\}$. This set and $\mathbb{P}^1(\overline{\mathbb{K}})^{N_2-N_1+1}$ have the same dimension, $N_2 - N_1 + 1$.

Given a kernel polynomial $\hat{Q}(x, y)$, there is a finite number of distinct points $(\alpha, \beta) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $\hat{Q}(\alpha, \beta) = 0$. Hence, the pre-image of an element in the image of $Q_{[\overline{\alpha}]}$ is a finite set. Therefore, the dimension of the image and the dimension of the domain are the same.

By Thm. 9.3.6, the non-square-free kernel polynomials form a hypersurface in the space of kernel polynomials of the shape $P_\mu P_v + P_w$. If we consider the pre-image of the intersection between this hypersurface and the image of the rational map, then its dimension is smaller than $N_2 - N_1 + 1$.

Therefore, generically, for $N_2 - N_1 + 1$ different points in $\mathbb{P}^1(\mathbb{K})$, the map $Q_{[\overline{\alpha}]}(x, y)$ results a square-free kernel polynomial. As $\overline{\mathbb{K}}$ is the algebraic closure of $\mathbb{K} \subset \mathbb{C}$, the same holds over $\mathbb{K}$. □

Theorem 9.4.4. *Given a binary form f of rank r and degree D, there is a square-free kernel polynomial of degree r such that the biggest degree of its irreducible factors is* $\min(r, D - r + 1)$.

Proof. If the rank is $r = N_2 + 1$, then $\min(r, D - r + 1) = N_1$. By Thm. 9.4.3, such a square-free kernel polynomial exists. If the rank is $r = N_1 + 1$ and $N_1 < N_2$, by Lem. 9.3.1, there is a square-free kernel polynomial of degree $\min(r, D - r + 1) = N_1 + 1$. □

The previous result is related to the decomposition of tensors of the same border rank [CS11, Thm. 2]; [BGI11, Thm. 23]; [Ble15].

We can also bound the number of possible bad choices in the proof of Thm. 9.4.3.

Proposition 9.4.5. *Let f be a binary form of rank N_2+1. For every set $S \subset \mathbb{P}^1(\mathbb{K})$ of cardinal (N_2-N_1) such that $(\forall(\alpha,\beta) \in S)\ P_v(\alpha,\beta) \neq 0$ there are at most D^2+3D+1 values $(\alpha_0,\beta_0) \in \mathbb{P}^1(\mathbb{K})$ such that $(\alpha_0,\beta_0) \notin S$, $P_v(\alpha_0,\beta_0) \neq 0$ and the unique kernel polynomial $Q_\mu := P_\mu P_v + P_w$ that vanish over S and (α_0,β_0) (Lem. 9.4.1) is not square-free.*

To prove this proposition we use Lagrange polynomials to construct the maps and varieties of the proof of Thm. 9.4.3.

Let $S = \{(\alpha_1,\beta_1),\ldots,(\alpha_{N_2-N_1},\beta_{N_2-N_1})\} \subset \mathbb{P}^1(\mathbb{K})$ be the set of Prop. 9.4.5. For each $(\alpha_0,\beta_0) \in \mathbb{P}^1(\mathbb{K})$ such that $(\alpha_0,\beta_0) \notin S$ and $P_v(\alpha_0,\beta_0) \neq 0$ we consider the unique kernel polynomial Q^{α_0,β_0} which vanishes as S and (α_0,β_0) (see Lem. 9.4.1). Using Lagrange polynomial, we can write this polynomial as

$$Q^{\alpha_0,\beta_0}(x,y) = \left(-\frac{P_w(\alpha_0,\beta_0)}{P_v(\alpha_0,\beta_0)} \frac{M(x,y)}{M(\alpha_0,\beta_0)} + \sum_{i=1}^{N_2-N_1} \frac{\beta_0 x - \alpha_0 y}{\alpha_0\beta_i - \alpha_i\beta_0} E_i(x,y) \right) P_v(x,y) + P_w(x,y)$$

Where $E_i(x,y) := -\frac{P_w(\alpha_i,\beta_i)}{P_v(\alpha_i,\beta_i)} \prod_{j\notin\{0,i\}} \frac{\beta_j x-\alpha_j y}{\alpha_i\beta_j-\alpha_j\beta_i}$ and $M(x,y) := \prod_{j=1}^{N_2-N_1}(\beta_j x - \alpha_j y)$.[2]

For each (α_j,β_j), we characterize the possible $(\alpha_0,\beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that (α_j,β_j) is a root of Q^{α_0,β_0} of multiplicity bigger than one. Then, we study the $(\alpha_0,\beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that (α_0,β_0) is a root of Q^{α_0,β_0} with multiplicity bigger than one. Finally, we reduce every case to the previous ones.

To study the multiplicities of the roots, we use the fact that (α_0,β_0) is a double root of an binary form P if and only if $P(\alpha_0,\beta_0) = \frac{\partial P}{\partial x}(\alpha_0,\beta_0) = \frac{\partial P}{\partial y}(\alpha_0,\beta_0) = 0$. Hence, for each $(\alpha_0,\beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$, we consider $\frac{\partial Q^{\alpha_0,\beta_0}}{\partial x}$ and $\frac{\partial Q^{\alpha_0,\beta_0}}{\partial y}$, where

$$\begin{aligned}\frac{\partial Q^{\alpha_0,\beta_0}}{\partial x} = &-\frac{P_w(\alpha_0,\beta_0)}{P_v(\alpha_0,\beta_0)} \frac{1}{M(\alpha_0,\beta_0)} \left(\frac{\partial M}{\partial x} P_v + M \frac{\partial P_v}{\partial x} \right)(x,y) + \\ &\sum_{i=1}^{N_2-N_1} \frac{1}{\beta_0\alpha_i - \alpha_0\beta_i} \frac{\partial((\beta_0 x - \alpha_0 y)E_i P_v)}{\partial x}(x,y) + \frac{\partial P_w}{\partial x}(x,y)\end{aligned} \tag{9.13}$$

Let $O_x^{\alpha_0,\beta_0}(x,y)$ be the product between the last line of Eq. (9.13) and $M(\alpha_0,\beta_0)$, that is

$$O_x^{\alpha_0,\beta_0}(x,y) := \sum_{i=1}^{N_2-N_1} \frac{M(\alpha_0,\beta_0)}{\beta_0\alpha_i - \alpha_0\beta_i} \frac{\partial((\beta_0 x - \alpha_0 y)E_i P_v)}{\partial x}(x,y) + M(\alpha_0,\beta_0)\frac{\partial P_w}{\partial x}(x,y)$$

Note that for every $(\alpha_i,\beta_i) \in S$, $(\beta_0\alpha_i - \alpha_0\beta_i)$ divides $M(\alpha_0,\beta_0)$, as polynomials in $\overline{\mathbb{K}}[\alpha_0,\beta_0]$, so $O_x^{\alpha_0,\beta_0}(x,y)$ is a polynomial in $\overline{\mathbb{K}}[\alpha_0,\beta_0][x,y]$. The derivative of Q^{α_0,β_0} with respect to x is a rational function in $\overline{\mathbb{K}}(\alpha_0,\beta_0)[x,y]$, that we can write as $\frac{\partial Q^{\alpha_0,\beta_0}}{\partial x} = \frac{T^{\alpha_0,\beta_0}(x,y)}{P_v(\alpha_0,\beta_0)M(\alpha_0,\beta_0)}$ where

$$T^{\alpha_0,\beta_0}(x,y) := -P_w(\alpha_0,\beta_0)\left(\frac{\partial M}{\partial x} P_v + M \frac{\partial P_v}{\partial x} \right)(x,y) + O_x^{\alpha_0,\beta_0}(x,y)P_v(\alpha_0,\beta_0) \in \overline{\mathbb{K}}[\alpha_0,\beta_0][x,y]$$

[2] For each $0 \leq i \leq N_2 - N_1$, $Q^{\alpha_0,\beta_0}(x,y)$ is a rational function of degree 0 with respect to (α_i,β_i). Hence, it is well defined the evaluation of the variables (α_i,β_i) in $Q^{\alpha_0,\beta_0}(x,y)$ at points of $\mathbb{P}^1(\mathbb{K})$.

Lemma 9.4.6. *For each $(\alpha_i, \beta_i) \in S$, there are at most $N_2 + 1$ possible $(\alpha_0, \beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\alpha_0, \beta_0) \notin S$, $P_v(\alpha_0, \beta_0) \neq 0$ and that (α_i, β_i) is a root of multiplicity bigger than 1 in Q^{α_0,β_0}.*

Proof. If (α_i, β_i) is a root of multiplicity bigger than 1 in Q^{α_0,β_0}, then $\frac{\partial Q^{\alpha_0,\beta_0}}{\partial x}(\alpha_i, \beta_i) = 0$. Hence, we are looking for the (α_0, β_0) such that $T^{\alpha_0,\beta_0}(\alpha_i, \beta_i) = 0$. The polynomial $T^{\alpha_0,\beta_0}(\alpha_i, \beta_i)$ belongs to $\overline{\mathbb{K}}[\alpha_0, \beta_0]$, so if it is not identically zero, there is a finite number of points $(\alpha_0, \beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $T^{\alpha_0,\beta_0}(\alpha_i, \beta_i) = 0$. Moreover, the degree of the polynomial $T^{\alpha_0,\beta_0}(\alpha_i, \beta_i)$ is at most $\max(\deg(P_w), \deg(O_x^{\alpha_0,\beta_0}(\alpha_i, \beta_i)) + \deg(P_v)) = N_2 + 1$, hence, if the polynomial is not zero, this finite number is at most $N_2 - N_1$.

The polynomial $T^{\alpha_0,\beta_0}(\alpha_i, \beta_i) \in \overline{\mathbb{K}}[\alpha_0, \beta_0]$ is not zero. Observe that as M is square-free, $M(\alpha_i, \beta_i) = 0$ and $P_v(\alpha_i, \beta_i) \neq 0$, then $\left(\frac{\partial M}{\partial x} P_v + M \frac{\partial P_v}{\partial x}\right)(\alpha_i, \beta_i) \neq 0$. Hence, as P_w and P_v are coprime, and so $T^{\alpha_0,\beta_0}(\alpha_i, \beta_i)$ does not vanish in the roots of P_v. □

Lemma 9.4.7. *There are at most $2N_2 + 1$ possible $(\alpha_0, \beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\alpha_0, \beta_0) \notin S$, $P_v(\alpha_0, \beta_0) \neq 0$ and (α_0, β_0) is a root of multiplicity bigger than 1 in Q^{α_0,β_0}.*

Proof. Following the proof of Lem. 9.4.6, we study $T^{\alpha_0,\beta_0}(\alpha_0, \beta_0) \in \overline{\mathbb{K}}[\alpha_0, \beta_0]$.

$$\begin{aligned} T^{\alpha_0,\beta_0}(\alpha_0, \beta_0) &= -P_w(\alpha_0, \beta_0)\left(\frac{\partial M}{\partial x} P_v + M \frac{\partial P_v}{\partial x}\right)(\alpha_0, \beta_0) + O_x^{\alpha_0,\beta_0}(\alpha_0, \beta_0) P_v(\alpha_0, \beta_0) \\ &= \left(-P_w\, M\, \frac{\partial P_v}{\partial x}\right)(\alpha_0, \beta_0) + \left(O_x^{\alpha_0,\beta_0} - P_w \frac{\partial M}{\partial x}\right)(\alpha_0, \beta_0) P_v(\alpha_0, \beta_0) \end{aligned}$$

The polynomial $T^{\alpha_0,\beta_0}(\alpha_0, \beta_0)$ is not zero because, as both P_w and M are coprime to P_v, P_v does not divides $P_w\, M\, \frac{\partial P_v}{\partial x}$. We conclude the proof by noting that the degree of $T^{\alpha_0,\beta_0}(\alpha_0, \beta_0)$ is bounded by $2N_2 + 1$. □

Lemma 9.4.8. *Let $(\bar{\alpha}_0, \bar{\beta}_0), (\alpha_0, \beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\bar{\alpha}_0, \bar{\beta}_0), (\alpha_0, \beta_0) \notin S$, $P_v(\bar{\alpha}_0, \bar{\beta}_0) \neq 0$. Hence, $Q^{\bar{\alpha}_0,\bar{\beta}_0}(\alpha_0, \beta_0) = 0$ if and only if $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x, y) = Q^{\alpha_0,\beta_0}(x, y)$.*

Proof. Assume that $Q^{\bar{\alpha}_0,\bar{\beta}_0}(\alpha_0, \beta_0) = 0$. Following Lem. 9.4.1, we write $Q^{\bar{\alpha}_0,\bar{\beta}_0} = P_{\bar{\mu}} P_v + P_w$ and $Q^{\alpha_0,\beta_0} = P_\mu P_v + P_w$. Consider $Q^{\bar{\alpha}_0,\bar{\beta}_0} - Q^{\alpha_0,\beta_0} = P_v(P_{\bar{\mu}} - P_\mu)$. This polynomial vanishes over $\mathbb{P}^1(\overline{\mathbb{K}})$ at the $N_1 + 1$ roots of P_v, at the $N_2 - N_1$ points on S, and at $(\alpha_0, \beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$. Hence, $Q^{\bar{\alpha}_0,\bar{\beta}_0} - Q^{\alpha_0,\beta_0} = 0$ as it is a binary form of degree at most N_2+1 with N_2+2 different roots over $\mathbb{P}^1(\overline{\mathbb{K}})$. Therefore, if $Q^{\bar{\alpha}_0,\bar{\beta}_0}(\alpha_0, \beta_0) = 0$, then $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x, y) = Q^{\alpha_0,\beta_0}(x, y)$. By definition, $Q^{\alpha_0,\beta_0}(\alpha_0, \beta_0) = 0$. Hence, if $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x, y) = Q^{\alpha_0,\beta_0}(x, y)$, then we have $Q^{\bar{\alpha}_0,\bar{\beta}_0}(\alpha_0, \beta_0) = 0$. □

Proof of Prop. 9.4.5. We want to bound the number of different points $(\alpha_0, \beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $Q^{\alpha_0,\beta_0}(x, y)$ is not a square-free binary form over $\overline{\mathbb{K}}[x, y]$. If the binary form $Q^{\alpha_0,\beta_0}(x, y)$ is not square-free, then it has a root over $\mathbb{P}^1(\overline{\mathbb{K}})$ with multiplicity bigger than one. If such a root is $(\alpha_i, \beta_i) \in S$, we can bound the possible number of different values for $(\alpha_0, \beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ by $(N_2 + 1)$ (Lem. 9.4.7). Hence, if there is a i such that $(\alpha_i, \beta_i) \in S$ has multiplicity bigger than one as a root of $Q^{\alpha_0,\beta_0}(x, y)$, we can bound the possible number of different values for $(\alpha_0, \beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ by $\#S \cdot (N_2 + 1) = (N_2 - N_1)(N_2 + 1)$.

If Q^{α_0,β_0} is not square-free and the multiplicity of every root $(\alpha_i, \beta_i) \in S$ is one, then there must be a root $(\bar{\alpha}_0, \bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\bar{\alpha}_0, \bar{\beta}_0) \notin S$ and its multiplicity as a root of Q^{α_0,β_0} is bigger than one.

By Lem. 9.4.8, $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y) = Q^{\alpha_0,\beta_0}(x,y)$, and so $(\bar{\alpha}_0,\bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ has multiplicity bigger than one as a root of $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y)$. Hence, $P_v(\bar{\alpha}_0,\bar{\beta}_0) \neq 0$ and, by Lem. 9.4.7, we can bound the possible number of different values for $(\bar{\alpha}_0,\bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ by $2N_2+1$. As $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y)$ has N_1+1 roots over $\mathbb{P}^1(\overline{\mathbb{K}}) \setminus S$ then, by Lem. 9.4.8, there are N_1+1 different $(\alpha_0,\beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y) = Q^{\alpha_0,\beta_0}(x,y)$. Hence, for each $(\bar{\alpha}_0,\bar{\beta}_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\bar{\alpha}_0,\bar{\beta}_0)$ has multiplicity bigger than one as a root of $Q^{\bar{\alpha}_0,\bar{\beta}_0}(x,y)$, there are N_1+1 points $(\alpha_0,\beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $(\bar{\alpha}_0,\bar{\beta}_0)$ has multiplicity bigger than one as a root of $Q^{\alpha_0,\beta_0}(x,y)$. Therefore, the number of different values for $(\alpha_0,\beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that $Q^{\alpha_0,\beta_0}(x,y)$ has a root in $\mathbb{P}^1(\overline{\mathbb{K}}) \setminus S$ with multiplicity bigger than one is bounded by $(N_1+1)(2N_2+1)$.

Joining these bounds, we deduce that there are at most $(N_2-N_1)(N_2+1)+(N_1+1)(2N_2+1)$ different $(\alpha_0,\beta_0) \in \mathbb{P}^1(\overline{\mathbb{K}})$ such that Q_{α_0,β_0} is not square-free. Recalling that $N_1 = D-N_2$ and $N_2 \leq D$ (Prop. 9.2.4), we can bound $(N_2-N_1)(N_2+1)+(N_1+1)(2N_2+1)$, by D^2+3D+1. □

Complexity of computing λ

We compute the coefficients λ_j of the decomposition by solving a linear system involving a transposed Vandermonde matrix (Step 3 of Alg. 14). We follow [KY89] to write the solution of Eq. (9.5) as the evaluation of a rational function over the roots of a univariate polynomial.

Definition 9.4.9. *Given a polynomial $P(x) := \sum_{i=0}^{n} a_i x^i$ and $0 < k \leq n$, let*

$$Quo(P(x), x^k) := \sum_{i=k}^{n} a_i x^{i-k}.$$

Proposition 9.4.10 ([KY89, Sec. 5])**.** *If $\alpha_j \neq \alpha_i$, for all $i \neq j$, then there is a unique solution to the system of Eq.* (9.14)*.*

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_r \\ \vdots & \vdots & & \vdots \\ \alpha_1^{r-1} & \alpha_2^{r-1} & \cdots & \alpha_r^{r-1} \end{pmatrix} \lambda = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{r-1} \end{pmatrix} \tag{9.14}$$

Moreover, if the solution is $\lambda = (\lambda_1,\ldots,\lambda_r)^{\mathsf{T}}$ then, $\lambda_j = \frac{T}{Q'}(\alpha_j)$ where $Q'(x)$ is the derivative of $Q(x) := \prod_{i=1}^{r}(x-\alpha_i)$, $R(x) := \sum_{i=1}^{r} a_{r-i}x^{i-1}$ and $T(x) := Quo\big(Q(x)R(x), x^r\big)$.

Lemma 9.4.11. *Given a binary form $f(x,y) := \sum_{i=0}^{D} \binom{D}{i} a_i x^i y^{D-i}$, let Q be a square-free kernel polynomial of degree r, obtained after step 3 of Alg. 14. Assume that y does not divide Q. Let α_j be the j-th roots of $Q(x)$, $Q'(x)$ be the derivative of $Q(x)$ and the polynomial $T(x) := Quo\big(Q(x)R(x), x^r\big)$, with $R(x) := \sum_{i=1}^{r} a_{r-i}x^{i-1}$. Then, each λ_j from step 3 in Alg. 15 can be written as $\lambda_j = \frac{T}{Q'}(\alpha_j)$.*

Proof. As y does not divide Q, we can write it as $Q(x,y) = \prod_i (x-\alpha_i y)$, where all the α_i are different. Hence, as the $r \times r$ leading principal submatrix of Eq. (9.5) is invertible, we can restrict the problem to solve that $r \times r$ leading principal subsystem. This system is Eq. (9.14). Therefore, the proof follows from Prop. 9.4.10. □

Proposition-Definition 9.4.12 (Symbolic decomposition)**.** *Let Q be a square-free kernel polynomial related to a minimal decomposition of a binary form f of degree D, such that y does not divide Q. In this case, we can write f as*

$$f(x,y) = \sum_{\{\alpha \in \overline{\mathbb{K}} | Q(\alpha)=0\}} \frac{T}{Q'}(\alpha)(\alpha x + y)^D.$$

Remark 9.4.13. *If the square-free kernel polynomial related to a decomposition of rank r is divisible by y, we can compute $\{\lambda_j\}_{j<r}$ of Eq. (9.5) as in Lem. 9.4.11, by taking $\frac{Q}{y}$ as the kernel polynomial. It is without loss of generality to consider $Q = P_{(u_0,\dots,-1,0)^\intercal}$, because Q is square-free, and so y^2 can not divide it. Hence, $\lambda_r = a_D - \sum_{i=0}^{r-2} u_i a_{D-r+i+1}$ [Rez13a, Eq. 2.12].*

To summarize the section, given a binary form f of rank r, there is a square-free kernel polynomial Q of the degree r, such that the largest degree of its irreducible factors is bounded by $\min(r, D-r+1)$ (Proposition-Definition 9.4.12). If $Q(x,y)$ is not divisible by y, the decomposition is

$$f(x,y) = \sum\nolimits_{\{\alpha \in \overline{\mathbb{K}} | Q(\alpha)=0\}} \frac{T}{Q'}(\alpha)(\alpha x + y)^D,$$

where T and Q' are polynomials whose coefficients belong to $\mathbb{K}$ and whose degrees are bounded by r, defined in Lem. 9.4.11. When y divides Q, the form is similar.

Lower bounds on the algebraic degree

In this section we analyze the tightness of the bound of Thm. 9.4.4. To do so, we construct families of examples where the bound is tight. We present two families of examples. In the first one, the decomposition is unique. In the second one, it is not.

Proposition 9.4.14 ([HR84, Theorem 5.2])**.** *For every pair of relatively prime binary forms, $\bar{P}_v$ and $\bar{P}_w$, of degrees $\bar{N}_1+1$ and $\bar{N}_2+1$, $\bar{N}_1 \leq \bar{N}_2$, there is a sequence $a = (a_0, \dots, a_{\bar{N}_1+\bar{N}_2})$ such that $N_1^a = \bar{N}_1$, $N_2^a = \bar{N}_2$, and we can consider the polynomials $\bar{P}_v$ and $\bar{P}_w$ as the kernel polynomials P_v and P_w from Prop. 9.2.7 with respect to the family of Hankel matrices $\{H_a^k\}_k$.*

Theorem 9.4.15. *If there is an irreducible binary form of degree $\bar{N}_1 + 1$ in $\mathbb{K}[x,y]$, then for every $D > 2\,\bar{N}_1$, there is a binary form f of degree D such that its decomposition is unique, its rank $\bar{N}_1 + 1$, and the degree of the biggest irreducible factor of the polynomial Q from Alg. 14 in the decomposition is $\min(\bar{N}_1 + 1, D - \bar{N}_1) = \bar{N}_1 + 1$. That is, the algebraic degree of the minimal decomposition over $\mathbb{K}$ is $\bar{N}_1 + 1$ and the bound of Thm. 9.4.4 is tight.*

Proof. Let $\bar{P}_v$ be a irreducible binary form of degree $\bar{N}_1 + 1$. Let $\bar{P}_w$ be any binary form of degree $\bar{N}_2 + 1 := D - \bar{N}_1 + 1$ relatively prime with $\bar{P}_v$. Consider the sequence $a = (a_0, \dots, a_{\bar{N}_1+\bar{N}_2})$ of Prop. 9.4.14 with respect to $\bar{P}_v$ and $\bar{P}_w$, and the binary form $f(x,y) := \sum_{i=0}^{D} \binom{D}{i} a_i x^i y^{D-i}$. As $\mathbb{K}$ is of characteristic 0, $\mathbb{K}$ is a perfect field, and so, as $\bar{P}_v$ is irreducible, it is square-free. Then, by Lem. 9.3.1, the rank of the decomposition is $N_1^a + 1 = \bar{N}_1 + 1$, and by Cor. 9.4.2 the decomposition is unique. Following Alg. 14, the polynomial Q is equal to $\bar{P}_v$, which is an irreducible polynomial of degree $\bar{N}_1+1$. As $D > 2\bar{N}_1$, then $\min(\bar{N}_1 + 1, D - \bar{N}_1) = \bar{N}_1 + 1$ and the bound of Thm. 9.4.4 is tight. □

Lemma 9.4.16. *Let $\mathbb{K} = \mathbb{Q}$ and $p \in \mathbb{N}$ a prime number. Then, there is a binary form f of degree $2(p-1)$ whose decomposition is not unique and the bound of Thm. 9.4.4 is tight.*

Proof. Consider the polynomial $f(x,y) := \binom{2(p-1)}{p-1} x^{p-1}\, y^{p-1}$. Using Alg. 15, we obtain $P_v = -y^p$ and $P_w = x^p$, $N_1 = N_2 = p-1$. The polynomial P_v is not square-free, so we have to consider a square-free kernel polynomial in $\mathtt{Ker}(H^{N_2+1})$. Moreover, the rank of the decomposition is $N_2+1 = p$. Every kernel polynomial in $\mathtt{Ker}(H^{N_2+1})$ in $\mathbb{Q}[x,y]$ can be written as $\mu_w x^p - \mu_v y^p$ for some $\mu_w, \mu_v \in \mathbb{Q}$. We are interested in the zeros of these polynomials (step 3 of Alg. 14), thus we consider coprime $\mu_w, \mu_v \in \mathbb{Z}$, as the zeros do not change. As we want to consider square-free kernel polynomials, neither μ_w nor μ_v can be zero, and so $(1,0) \in \mathbb{P}^1(\mathbb{Q})$ is not a root of any of these polynomials. Hence, we rewrite our polynomial as $\frac{1}{\mu_v y^p}(\frac{\mu_w x^p}{\mu_v y^p} - 1)$, and so we look for the factorization over $\mathbb{Q}[z]$ of $\frac{\mu_w}{\mu_v} z^p - 1$, where $z = \frac{x}{y}$. We can use the Newton's polygon criterion, e.g., [Cas86, Chp. 6.3], to show that, if $\sqrt[p]{\frac{\mu_w}{\mu_v}} \notin \mathbb{Q}$, then $\frac{\mu_w}{\mu_v} z^p - 1$ is irreducible over $\mathbb{Q}[x,y]$ and so the degree of its biggest irreducible factor is $p > \min(p, 2\,(p-1) - p + 1)$. If this is not the case, then $\sqrt[p]{|\frac{\mu_w}{\mu_v}|} \in \mathbb{Q}$, and so we can factor it as

$$\left(\sqrt[p]{|\frac{\mu_w}{\mu_v}|} \cdot z\right)^p - 1 = \left(\sqrt[p]{|\frac{\mu_w}{\mu_v}|} \cdot z - 1\right) \left(\sum_{i=0}^{p-1} \left(\sqrt[p]{|\frac{\mu_w}{\mu_v}|} \cdot z\right)^i\right).$$

The second factor is irreducible because there is an automorphism in $\mathbb{Q}[x]$ (given by $z \mapsto \sqrt[p]{|\frac{\mu_v}{\mu_w}|}z$) that transforms it into the p-th cyclotomic polynomial, which is irreducible as p is prime. Hence, the biggest irreducible factor of this polynomial has degree $p - 1 = \min(p, 2\,(p-1) - p + 1)$ and the bound of Thm. 9.4.4 is tight. □

9.4.2 Arithmetic complexity

Lemma 9.4.17 (Complexity of Alg. 15)**.** *Given a binary form $f = \sum_{i=0}^{D} \binom{D}{i} a_i x^i y^{D-i}$ of degree D, Alg. 15 computes P_v and P_w in $O(\mathtt{M}(D)\log(D))$.*

Proof. The complexity of the algorithm is the complexity of computing the rows $(i+1)$, i and $(i-1)$ of the Extended Euclidean algorithm between $\sum_{i=0} a_i x^i$ and x^{D+1}, where the i-th row is the first row i such that $\deg(R_i) < \frac{D}{2}$ (Lem. 9.3.5). This can be done using the *Half-GCD algorithm*, which computes these rows in $O(\mathtt{M}(D)\log(D))$. For a detailed reference of how this algorithm works see [BCG$^+$17, Ch. 6.3] or [GG13, Ch. 11]. □

Lemma 9.4.18 (Complexity of computing Q)**.** *Given the kernel polynomials P_v and P_w from Prop. 9.2.7, we compute a square-free polynomial $Q_\mu := P_\mu P_v + P_w$ with the algebraic degree of Thm. 9.4.4 in $O(\mathtt{M}(D)\log(D))$.*

Proof. To compute the vector μ, we choose randomly $N_2 - N_1 + 1$ linear forms and we proceed as in Lem. 9.4.1. The complexity bound is due to multi-point evaluation and interpolation of a univariate polynomial [GG13, Ch. 10]. □

Theorem 9.4.19. *When the decomposition is unique, that is when the rank is $N_1 + 1$, then Alg. 14 computes deterministically a symbolic decomposition (Proposition-Definition 9.4.12) of a binary form in $O(\mathtt{M}(D)\ \log(D))$.*

When the decomposition is not unique, that is when the rank is $N_2 + 1$, then Alg. 14 is a Monte Carlo algorithm that computes a symbolic decomposition of a binary form in $O(\mathtt{M}(D)\ \log(D))$.

Proof. The first step of the algorithm, in both cases, is to compute the kernel polynomials P_v and P_w of Prop. 9.2.7 using Alg. 15. By Lem. 9.4.17, we compute them deterministically in $O(\mathtt{M}(D)\log(D))$.

If P_v is square-free, which means that the decomposition is unique, then $Q = P_v$. Otherwise, we need to choose some random values to construct the polynomial square-free polynomial Q from the kernel polynomials P_v and P_w, step 2 using (Thm. 9.4.3), in $O(\mathtt{M}(D)\log(D))$ (Lem. 9.4.18). This is the step that makes the algorithm a Monte Carlo one, as we might fail to produce a square-free polynomial Q.

In both cases, at step 3 we compute the rational function that describes the solution of the system in Eq. (9.5), in $O(\mathtt{M}(D)\log(D))$ [KY89]. At step 4 of the algorithm we return the decomposition. □

We can bound the probability of error of Alg. 14 using Prop. 9.4.5, which bounds the number of bad values that lead us to a non square-free polynomial Q. Moreover, we can introduce a Las Vegas version of Alg. 14 by checking if the values that we choose to construct a polynomial Q result indeed a square-free polynomial. We recall that this check can be done in $O(\mathtt{M}(D)\log(D))$ by computing the GCD between the Q and its derivatives.

Theorem 9.4.20. *If we want to output an approximation of the terms of the minimal decomposition, with a relative error of $2^{-\epsilon}$, we can use Pan's algorithm [Pan02] [MP13, Thm. 15.1.1] to approximate the roots of Q. In this case the complexity becomes $O\big(D\log^2(D)\big(\log^2(D) + \log(\epsilon)\big)\big)$.*

9.4.3 Bit complexity

Let $f \in \mathbb{Z}[x, y]$ be a binary form as in Eq. (9.1), of degree D and let τ be the maximum bitsize of the coefficients a_i. We study the bit complexity of computing suitable approximations of the α_j's, β_j's, and λ_j's of Eq. (9.3), say $\widetilde{\alpha}_j$, $\widetilde{\beta}_j$ and $\widetilde{\lambda}_j$ respectively, that induce an approximate decomposition correct up to ℓ bits. That is a decomposition such that $\|f - \sum_j \widetilde{\lambda}_j(\widetilde{\alpha}_j x + \widetilde{\beta}_j y)^D\|_\infty \leq 2^{-\ell}$. We need to estimate an upper bound on the number of bits that are necessary to perform all the operations of the algorithm.

The first step of the algorithm is to compute P_v and P_w, via the computation of three rows of the Extended GCD of two polynomials of degree D and $D + 1$ with coefficients of maximal sized τ. This can be achieved in $\widetilde{\mathcal{O}}_B(D^2\tau)$ bit operations [GG13, Cor. 11.14.B], and the maximal bit size of P_v and P_w is $\widetilde{\mathcal{O}}(D\tau)$. We check if P_v is a square-free polynomial in $\widetilde{\mathcal{O}}(D^2\tau)$, via the computation of the GCD of $P_v(x, 1)$ and its derivative [GG13, Cor. 11.14.A], and by checking if y^2 divides it.

If P_v is square-free polynomial, then $Q = P_v$. If P_v is not square-free, then we can compute Q by assigning values to the coefficients of P_μ. We assume that y^2 does not divide P_w, if this does not hold, we replace P_w by the kernel polynomial $x^{N_2-N_1}P_v + P_w$, which is coprime to P_v, and so not divisible by y, as P_v and the original P_w are coprime (Prop. 9.2.7). We set all the coefficients of P_μ to zero, except the constant term. Then $Q = \mu_0 y^{N_2-N_1}P_v + P_w$. Now we have to choose μ_0 so that Q is square-free. As $y^{N_2-N_1}P_v$ and P_w are coprime, there are at most $2D + 2$ forbidden values for μ_0 such that Q is

not square-free (Cor. 9.3.7), thus at least one of the first $2D + 3$ integer fits our requirements. We test them all. Each test corresponds to a GCD computation, that costs $\widetilde{\mathcal{O}}_B(D^2\tau)$ and so the overall cost is $\widetilde{\mathcal{O}}_B(D^3\tau)$.

Let $\sigma = \widetilde{\mathcal{O}}(D\tau)$ be the maximal bit size of Q. By Rem. 9.4.13, we can assume that y does not divide Q, consider $y = 1$ and treat Q as an univariate polynomial.

Let $\{\alpha_j\}_j$ be the roots of Q. We isolate them in $\widetilde{\mathcal{O}}_B(D^2\sigma)$ [Pan02]. For the (aggregate) separation bound of the roots it holds that $-\lg \prod_j \Delta(\alpha_j) = O(D\sigma + D\lg(D))$. We approximate all the roots up to accuracy $2^{-\ell_1}$ in $\widetilde{\mathcal{O}}_B(D^2\sigma + D\ell_1)$ [PT17a]. That is, we compute absolute approximations of α_j, say $\widetilde{\alpha}_j$, such that $|\alpha_j - \widetilde{\alpha}_j| \leq 2^{-\ell_1}$.

The next step consists in solving the (transposed) Vandermonde system, $V(\widetilde{\alpha})^{\mathsf{T}}\lambda = a$, where $V(\widetilde{\alpha})$ is the Vandermonde matrix we construct with the roots of Q, λ is a vector contains the coefficients of decomposition, and a is a vector containing the coefficients of F, see also Eq. (9.5). We know the entries of $V(\widetilde{\alpha})$ up to ℓ_1 bits. Therefore, we can compute the elements of the solution vector λ with an absolute approximation correct up to $\ell_2 = \ell_1 - O(D\lg(D)\sigma - \lg\prod_j \Delta(\alpha_j)) = \ell_1 - O(D\lg(D)\sigma)$ bits [PT17b, Thm. 29]. That is, we compute $\widetilde{\lambda}_j$'s such that $|\lambda_j - \widetilde{\lambda}_j| \leq 2^{-\ell_2}$. At this point we have obtained the approximate decomposition

$$\widetilde{f}(x, y) := \sum_{j=1}^{r} \widetilde{\lambda}_j(\widetilde{\alpha}_j x + y)^D.$$

To estimate the accuracy of $\widetilde{f}$ we need to expand the approximate decomposition and consider it as a polynomial in x. We do not actually perform this operation; we only estimate the accuracy as if we were. First, we expand each $(\widetilde{\alpha}_j x + y)^D$. This results polynomials with coefficients correct up $\ell_3 = \ell_2 - O(D\sigma) = \ell_1 - O(D\lg(D)\sigma) - O(D\sigma) = \ell_1 - O(D\lg(D)\sigma)$ bits [PT17b, Lemma 19]. Next, we multiply each such polynomial with $\widetilde{\lambda}_i$, and we collect the coefficients for the various powers of x. Each coefficient is the sum of $r \leq D$ terms. The last two operations do not affect, asymptotically, the precision. Therefore, the polynomial $\widetilde{f} = \sum_{j=1}^{r} \widetilde{\lambda}_j(\widetilde{\alpha}_j x + (1 - \widetilde{\alpha}_j\, t)y)^D$ that corresponds to the approximate decomposition has an absolute approximation such that $\|f - \widetilde{f}\| \leq 2^{-\ell_1 + O(D\lg(D)\sigma)}$. To achieve an accuracy of $2^{-\ell}$ in the decomposition, such that $\|f - \widetilde{f}\| \leq 2^{-\ell}$, we should choose $\ell_1 = \ell + O(D\lg(D)\sigma)$. Thus, all the computations should be performed with precision of $\ell + O(D\lg(D)\sigma)$ bits. The bit complexity of computing the decomposition of f up to ℓ bits is dominated by the solving and refining process and it is $\widetilde{\mathcal{O}}_B(D\ell + D^2\sigma)$. If we substitute the value for σ, then we arrive at the complexity bound of $\widetilde{\mathcal{O}}_B(D\ell + D^4 + D^3\tau)$.

Theorem 9.4.21. *Let $f \in \mathbb{Z}[x, y]$ be a homogeneous polynomial of degree D and maximum coefficient bitsize τ. We compute an approximate decomposition of accuracy $2^{-\ell}$ in $\widetilde{\mathcal{O}}_B(D\ell + D^4 + D^3\tau)$ bit operations.*

Bibliography

[ACG05] Azzouz Awane, Abdelouahab Chkiriba, and Michel Goze. Formes d'inertie et complexe de Koszul associés à des polynômes plurihomogènes. *Revista Matemática Complutense*, 18(1):243–260, 2005.

[AM69] Michael F. Atiyah and Ian G. Macdonald. *Introduction to commutative algebra*. Number 361 in Addison-Wesley Series in Mathematics. Addison-Wesley, Reading, MA, 1969. Russian translation published as *Vvedenie v kommutativnuju algebru* (1972). MR:0242802. Zbl:0175.03601.

[AM10] Frederick V. Atkinson and Angelo Mingarelli. *Multiparameter Eigenvalue Problems: Sturm-Liouville Theory*. CRC Press, December 2010.

[AMS09] Aleksandra S. Anokhina, Aleksei Yu. Morozov, and Shamil' R. Shakirov. Resultant as the determinant of a Koszul complex. *Theoretical and Mathematical Physics*, 160(3):1203, October 2009.

[AS88] Winfried Auzinger and Hans J. Stetter. An Elimination Algorithm for the Computation of All Zeros of a System of Multivariate Polynomial Equations. In Ravi P. Agarwal, Y. M. Chow, and S. J. Wilson, editors, *Numerical Mathematics Singapore 1988: Proceedings of the International Conference on Numerical Mathematics held at the National University of Singapore, May 31–June 4, 1988*, International Series of Numerical Mathematics / Internationale Schriftenreihe zur Numerischen Mathematik / Série internationale d'Analyse numérique, pages 11–30. Birkhäuser Basel, Basel, 1988.

[Atk72] Frederick V. Atkinson. *Multiparameter Eigenvalue Problems*. Academic Press, 1972.

[Baj88] Chanderjit Bajaj. The algebraic degree of geometric optimization problems. *Discrete & Computational Geometry*, 3(1):177–191, 1988.

[BBF17] Jérémy Berthomieu, Brice Boyer, and Jean-Charles Faugère. Linear algebra for computing Gröbner bases of linear recursive multidimensional sequences. *Journal of Symbolic Computation*, 83:36–67, November 2017.

[BC17] Nicolás Botbol and Marc Chardin. Castelnuovo Mumford regularity with respect to multigraded ideals. *Journal of Algebra*, 474:361–392, March 2017.

[BCG11] Mats Boij, Enrico Carlini, and Anthony Geramita. Monomials as sums of powers: the real binary case. *Proceedings of the American Mathematical Society*, 139(9):3039–3043, 2011.

[BCG+17] Alin Bostan, Frédéric Chyzak, Marc Giusti, Romain Lebreton, Grégoire Lecerf, Bruno Salvy, and Éric Schost. *Algorithmes efficaces en calcul formel*. published by the Authors, 2017.

[BCMT10] Jerome Brachat, Pierre Comon, Bernard Mourrain, and Elias Tsigaridas. Symmetric tensor decomposition. *Linear Algebra and its Applications*, 433(11):1851–1872, December 2010.

[BEF+16] Brice Boyer, Christian Eder, Jean-Charles Faugère, Sylvian Lachartre, and Fayssal Martani. GBLA: Gröbner Basis Linear Algebra Package. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '16, pages 135–142, New York, NY, USA, 2016. ACM.

[BEM00] Laurent Busé, Mohamed Elkadi, and Bernard Mourrain. Generalized Resultants over Unirational Algebraic Varieties. *Journal of Symbolic Computation*, 29(4):515–526, May 2000.

[BEM01] Laurent Busé, Mohamed Elkadi, and Bernard Mourrain. Resultant over the residual of a complete intersection. *Journal of Pure and Applied Algebra*, 164(1):35–57, October 2001.

[Ber75] David N. Bernshtein. The number of roots of a system of equations. *Functional Analysis and Its Applications*, 9(3):183–185, July 1975.

[BFMT18] Matías R. Bender, Jean-Charles Faugère, Angelos Mantzaflaris, and Elias Tsigaridas. Bilinear Systems with Two Supports: Koszul Resultant Matrices, Eigenvalues, and Eigenvectors. In *Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*, ISSAC '18, pages 63–70, New York, NY, USA, 2018. ACM.

[BFMT19] Matías R. Bender, Jean-Charles Faugère, Angelos Mantzaflaris, and Elias Tsigaridas. Determinantal formulas for families of mixed multilinear systems. In preparation, 2019.

[BFPT16] Matías R. Bender, Jean-Charles Faugère, Ludovic Perret, and Elias Tsigaridas. A Superfast Randomized Algorithm to Decompose Binary Forms. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '16, pages 79–86, New York, NY, USA, 2016. ACM.

[BFPT18] Matías R. Bender, Jean-Charles Faugère, Ludovic Perret, and Elias Tsigaridas. A nearly optimal algorithm to decompose binary forms. Submitted, 2018.

[BFS15] Magali Bardet, Jean-Charles Faugère, and Bruno Salvy. On the complexity of the F5 Gröbner basis algorithm. *Journal of Symbolic Computation*, 70:49–70, September 2015.

[BFSS13] Magali Bardet, Jean-Charles Faugère, Bruno Salvy, and Pierre-Jean Spaenlehauer. On the complexity of solving quadratic Boolean systems. *Journal of Complexity*, 29(1):53–75, February 2013.

[BFT18] Matías R. Bender, Jean-Charles Faugère, and Elias Tsigaridas. Towards Mixed Gröbner Basis Algorithms: The Multihomogeneous and Sparse Case. In *Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*, ISSAC '18, pages 71–78, New York, NY, USA, 2018. ACM.

[BFT19] Matías Bender, Jean-Charles Faugère, and Elias Tsigaridas. Gröbner Basis over Semigroup Algebras: Algorithms and Applications for Sparse Polynomial Systems. *arXiv:1902.00208 [cs]*, February 2019. Submitted.

[BGI11] Alessandra Bernardi, Alessandro Gimigliano, and Monica Idà. Computing symmetric rank for symmetric tensors. *Journal of Symbolic Computation*, 46(1):34–53, January 2011.

[BH98] Winfried Bruns and Jürgen Herzog. *Cohen-Macaulay Rings*. Cambridge University Press, June 1998.

[Ble15] Grigoriy Blekherman. Typical real ranks of binary forms. *Foundations of Computational Mathematics*, 15(3):793–798, 2015.

[BM92] Dave Bayer and David Mumford. What can be computed in algebraic geometry? In *Computational Algebraic Geometry and Commutative Algebra*, pages 1–28. University Press, 1992.

[BMMT94] Eberhard Becker, Teo Mora, Maria Grazia Marinari, and Carlo Traverso. The shape of the Shape Lemma. In *Proceedings of the international symposium on Symbolic and algebraic computation - ISSAC '94*, pages 129–133, Oxford, United Kingdom, 1994. ACM Press.

[BMT17] Laurent Busé, Angelos Mantzaflaris, and Elias Tsigaridas. Matrix formulae for Resultants and Discriminants of Bivariate Tensor-product Polynomials. Submitted. HAL Id : hal-01654263, version 1., December 2017.

[Bot11] Nicolas Botbol. *Implicitization of rational maps*. PhD thesis, Université de Paris VI and Univesidad de Buenos Aires, September 2011. arXiv: 1109.1423 [math].

[BS87a] David Bayer and Michael Stillman. A criterion for detectingm-regularity. *Inventiones Mathematicae*, 87(1):1–11, February 1987.

[BS87b] David Bayer and Michael Stillman. A theorem on refining division orders by the reverse lexicographic order. *Duke Mathematical Journal*, 55(2):321–328, 1987.

[BS13] Markus P. Brodmann and Rodney Y. Sharp. *Local Cohomology: An Algebraic Introduction with Geometric Applications*. Cambridge University Press, 2013.

[BT06] Jean-Daniel Boissonnat and Monique Teillaud, editors. *Effective Computational Geometry for Curves and Surfaces*. Mathematics and Visualization. Springer-Verlag, Berlin Heidelberg, 2006.

[Buc79] Bruno Buchberger. A criterion for detecting unnecessary reductions in the construction of Gröbner-bases. In Edward W. Ng, editor, *Symbolic and Algebraic Computation*, Lecture Notes in Computer Science, pages 3–21. Springer Berlin Heidelberg, 1979.

[Buc06] Bruno Buchberger. Bruno Buchberger's PhD thesis 1965: An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal. *Journal of Symbolic Computation*, 41(3-4):475–511, March 2006.

[Bus01] Laurent Busé. Residual Resultant over the Projective Plane and the Implicitization Problem. In *Proceedings of the 2001 International Symposium on Symbolic and Algebraic Computation*, ISSAC '01, pages 48–55, New York, NY, USA, 2001. ACM.

[Bus04] Laurent Busé. Resultants of determinantal varieties. *Journal of Pure and Applied Algebra*, 193(1):71–97, October 2004.

[Bus06] Laurent Busé. Elimination theory in codimension one and applications. Research report, INRIA, 2006. Notes of lectures given at the CIMPA-UNESCO-IRAN school in Zanjan, Iran, July 9-22 2005.

[BW98] Bruno Buchberger and Franz Winkler. *Gröbner Bases and Applications*. Cambridge University Press, February 1998.

[Can88] John F. Canny. *The Complexity of Robot Motion Planning*. MIT Press, Cambridge, MA, USA, 1988.

[Cas86] John William Scott Cassels. *Local fields*, volume 3. Cambridge University Press Cambridge, 1986.

[CC86] Stanley Cabay and Dong-Koo Choi. Algebraic computations of scaled Padé fractions. *SIAM Journal on Computing*, 15(1):243–270, 1986.

[CDS98] Eduardo Cattani, Alicia Dickenstein, and Bernd Sturmfels. Residues and Resultants. *Journal of mathematical sciences, the University of Tokyo*, 5(1):119–148, 1998.

[CE93] John F. Canny and Ioannis Z. Emiris. An efficient algorithm for the sparse mixed resultant. In Gérard Cohen, Teo Mora, and Oscar Moreno, editors, *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, Lecture Notes in Computer Science, pages 89–104. Springer Berlin Heidelberg, 1993.

[CE00] John F. Canny and Ioannis Z. Emiris. A Subdivision-based Algorithm for the Sparse Resultant. *J. ACM*, 47(3):417–451, May 2000.

[CGLM08] Pierre Comon, Gene Golub, Lek-Heng Lim, and Bernard Mourrain. Symmetric Tensors and Symmetric Tensor Rank. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1254–1279, January 2008.

[Cha91] Marc Chardin. Differential resultants and subresultants. In L. Budach, editor, *Fundamentals of Computation Theory*, Lecture Notes in Computer Science, pages 180–189. Springer Berlin Heidelberg, 1991.

[Cha93] Marc Chardin. The Resultant via a Koszul Complex. In *Computational Algebraic Geometry*, Progress in Mathematics, pages 29–39. Birkhäuser, Boston, MA, 1993.

[Cha95] Marc Chardin. Multivariate subresultants. *Journal of Pure and Applied Algebra*, 101(2):129–138, June 1995.

[Cha03] Marc Chardin. Bounds for Castelnuovo-Mumford Regularity in Terms of Degrees of Defining Equations. In Jürgen Herzog and Victor Vuletescu, editors, *Commutative Algebra, Singularities and Computer Algebra*, pages 67–73. Springer Netherlands, Dordrecht, 2003.

[CK00] Arthur D. Chtcherba and Deepak Kapur. Conditions for Exact Resultants Using the Dixon Formulation. In *Proceedings of the 2000 International Symposium on Symbolic and Algebraic Computation*, ISSAC '00, pages 62–70, New York, NY, USA, 2000. ACM.

[CK04] Arthur D. Chtcherba and Deepak Kapur. Constructing Sylvester-type resultant matrices using the Dixon formulation. *Journal of Symbolic Computation*, 38(1):777–814, July 2004.

[CLO06] David A. Cox, John Little, and Donal O'shea. *Using algebraic geometry*, volume 185. Springer Science & Business Media, 2006.

[CLO15] David A. Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*. Undergraduate texts in mathematics. Springer, Cham, 4. ed edition, 2015.

[CLS11] David A. Cox, John B. Little, and Henry K. Schenck. *Toric Varieties*. American Mathematical Society, 2011.

[CM96] Pierre Comon and Bernard Mourrain. Decomposition of quantics in sums of powers of linear forms. *Signal Processing*, 53(2):93–107, September 1996.

[CN08] Gemma Colomé Nin. *Multigraded Structures and the Depth of Blow-up Algebras*. PhD thesis, Universitat de Barcelona, July 2008.

[Com14] Pierre Comon. Tensors: a Brief Introduction. *IEEE Signal Processing Magazine*, 31(3):44–53, May 2014.

[Cox05] David A. Cox. Solving equations via algebras. In Alicia Dickenstein and Ioannis Z. Emiris, editors, *Solving Polynomial Equations: Foundations, Algorithms, and Applications*, chapter 2, pages 63–123. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[CP16] Diego Cifuentes and Pablo Parrilo. Exploiting Chordal Structure in Polynomial Ideals: A Gröbner Bases Approach. *SIAM Journal on Discrete Mathematics*, 30(3):1534–1570, January 2016.

[CS02] Felipe Cucker and Steve Smale. On the mathematical foundations of learning. *Bulletin of the American Mathematical Society*, 39(1):1–49, 2002.

[CS11] Gonzalo Comas and Malena Seiguer. On the Rank of a Binary Form. *Foundations of Computational Mathematics*, 11(1):65–78, February 2011.

[D'A02] Carlos D'Andrea. Macaulay style formulas for sparse resultants. *Transactions of the American Mathematical Society*, 354(7):2595–2629, 2002.

[DD01] Carlos D'Andrea and Alicia Dickenstein. Explicit formulas for the multivariate resultant. *Journal of Pure and Applied Algebra*, 164(1):59–86, October 2001.

[DE03] Alicia Dickenstein and Ioannis Z. Emiris. Multihomogeneous resultant formulae by means of complexes. *Journal of Symbolic Computation*, 36(3):317–342, September 2003.

[DE05] Alicia Dickenstein and Ioannis Z. Emiris, editors. *Solving Polynomial Equations*, volume 14 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin/Heidelberg, 2005.

[DHO+16] Jan Draisma, Emil Horobeţ, Giorgio Ottaviani, Bernd Sturmfels, and Rekha R Thomas. The euclidean distance degree of an algebraic variety. *Foundations of computational mathematics*, 16(1):99–149, 2016.

[DKS13] Carlos D'Andrea, Teresa Krick, and Martín Sombra. Heights of varieties in multiprojective spaces and arithmetic nullstellensatze. *Ann. Sci. École Norm. Sup.*, 46:549–627, 2013.

[DLK14] Jesús De Loera and Edward Kim. Combinatorics and geometry of transportation polytopes: An update. In Alexander Barg and Oleg Musin, editors, *Contemporary Mathematics*, volume 625, pages 37–76. American Mathematical Society, Providence, Rhode Island, 2014.

[DMB08] Leonardo De Moura and Nikolaj Bjørner. Z3: An Efficient SMT Solver. In *Proceedings of the Theory and Practice of Software, 14th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, TACAS'08/ETAPS'08, pages 337–340, Berlin, Heidelberg, 2008. Springer-Verlag.

[DS15] Carlos D'Andrea and Martín Sombra. A poisson formula for the sparse resultant. *Proceedings of the London Mathematical Society*, 110(4):932–964, 2015.

[DSS09] Mathias Drton, Bernd Sturmfels, and Seth Sullivant. *Lectures on Algebraic Statistics*. Oberwolfach Seminars. Birkhäuser Basel, 2009.

[DYY16] Bo Dong, Bo Yu, and Yan Yu. A Homotopy Method for Finding All Solutions of a Multiparameter Eigenvalue Problem. *SIAM Journal on Matrix Analysis and Applications*, 37(2):550–571, January 2016.

[Dü89] Arne Dür. On computing the canonical form for a binary form of odd degree. *Journal of Symbolic Computation*, 8(4):327–333, October 1989.

[EC95] Ioannis Z. Emiris and John F. Canny. Efficient Incremental Algorithms for the Sparse Resultant and the Mixed Volume. *Journal of Symbolic Computation*, 20(2):117–149, August 1995.

[Ede12] Christian Eder. *Signature–based algorithms to compute standard bases*. PhD thesis, Technische Universität Kaiserslautern, 2012.

[EF17] Christian Eder and Jean-Charles Faugère. A survey on signature-based algorithms for computing Gröbner bases. *Journal of Symbolic Computation*, 80:719–784, May 2017.

[Eis04] David Eisenbud. *Commutative Algebra: with a View Toward Algebraic Geometry*. Graduate Texts in Mathematics. Springer-Verlag, New York, 2004.

[Eis05] David Eisenbud. *The Geometry of Syzygies: A Second Course in Algebraic Geometry and Commutative Algebra*. Graduate Texts in Mathematics. Springer-Verlag, New York, 2005.

[EM98] Mohamed Elkadi and Bernard Mourrain. Some Applications of Bezoutians in Effective Algebraic Geometry. report, INRIA, December 1998.

[EM99a] Ioannis Z. Emiris and Bernard Mourrain. Computer Algebra Methods for Studying and Computing Molecular Conformations. *Algorithmica*, 25(2):372–402, June 1999.

[EM99b] Ioannis Z. Emiris and Bernard Mourrain. Matrices in Elimination Theory. *Journal of Symbolic Computation*, 28(1):3–44, July 1999.

[EM12] Ioannis Z. Emiris and Angelos Mantzaflaris. Multihomogeneous resultant formulae for systems with scaled support. *Journal of Symbolic Computation*, 47(7):820–842, July 2012.

[Emi96] Ioannis Z. Emiris. On the Complexity of Sparse Elimination. *Journal of Complexity*, 12(2):134–166, June 1996.

[Emi05] Ioannis Z. Emiris. Toric resultants and applications to geometric modelling. In Alicia Dickenstein and Ioannis Z. Emiris, editors, *Solving Polynomial Equations: Foundations, Algorithms, and Applications*, chapter 7, pages 63–123. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[EMT16] Ioannis Z. Emiris, Angelos Mantzaflaris, and Elias Tsigaridas. On the Bit Complexity of Solving Bilinear Polynomial Systems. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '16, pages 215–222, New York, NY, USA, 2016. ACM.

[EP02] Ioannis Z. Emiris and Victor Y. Pan. Symbolic and Numeric Methods for Exploiting Structure in Constructing Resultant Matrices. *Journal of Symbolic Computation*, 33(4):393–413, April 2002.

[ER94] Ioannis Z. Emiris and Ashutosh Rege. Monomial bases and polynomial system solving (extended abstract). In *Proceedings of the international symposium on Symbolic and algebraic computation - ISSAC '94*, pages 114–122, Oxford, United Kingdom, 1994. ACM Press.

[Fau99] Jean-Charles Faugère. A new efficient algorithm for computing Gröbner bases (F4). *Journal of Pure and Applied Algebra*, 139(1):61–88, June 1999.

[Fau02] Jean-Charles Faugère. A New Efficient Algorithm for Computing GrÖBner Bases Without Reduction to Zero (F5). In *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation*, ISSAC '02, pages 75–83, New York, NY, USA, 2002. ACM.

[Fau10] Jean-Charles Faugère. FGb: A Library for Computing Gröbner Bases. In Komei Fukuda, Joris Hoeven, Michael Joswig, and Nobuki Takayama, editors, *Mathematical Software - ICMS 2010*, volume 6327 of *Lecture Notes in Computer Science*, pages 84–87, Berlin, Heidelberg, September 2010. Springer Berlin / Heidelberg.

[FEDS12] Jean-Charles Faugère, Mohab Safey El Din, and Pierre-Jean Spaenlehauer. Critical points and grÖbner bases: The unmixed case. In *Proceedings of the 37th International Symposium on Symbolic and Algebraic Computation*, ISSAC '12, pages 162–169, New York, NY, USA, 2012. ACM.

[FGHR13] Jean-Charles Faugère, Pierrick Gaudry, Louise Huot, and Guénaël Renault. Polynomial Systems Solving by Fast Linear Algebra. *arXiv:1304.6039 [cs]*, April 2013.

[FGLM93] Jean-Charles Faugère, Patrizia M. Gianni, Daniel Lazard, and Teo Mora. Efficient Computation of Zero-dimensional Gröbner Bases by Change of Ordering. *Journal of Symbolic Computation*, 16(4):329–344, October 1993.

[FJ03] Jean-Charles Faugère and Antoine Joux. Algebraic cryptanalysis of hidden field equation (hfe) cryptosystems using gröbner bases. In *Annual International Cryptology Conference*, pages 44–60. Springer, 2003.

[FM17] Jean-Charles Faugère and Chenqi Mou. Sparse FGLM algorithms. *Journal of Symbolic Computation*, 80:538–569, May 2017.

[FS12] Jean-Charles Faugère and Jules 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*, ISSAC '12, pages 170–178, New York, NY, USA, 2012. ACM.

[FS13] Jean-Charles Faugère and Jules Svartz. Gröbner bases of ideals invariant under a commutative group: The non-modular case. In *Proceedings of the 38th International Symposium on Symbolic and Algebraic Computation*, ISSAC '13, pages 347–354, New York, NY, USA, 2013. ACM.

[FSEDS11] Jean-Charles Faugère, Mohab Safey El Din, and Pierre-Jean Spaenlehauer. Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1,1): Algorithms and complexity. *Journal of Symbolic Computation*, 46(4):406–437, April 2011.

[FSEDS13] Jean-Charles Faugère, Mohab Safey El Din, and Pierre-Jean Spaenlehauer. On the complexity of the generalized MinRank problem. *Journal of Symbolic Computation*, 55:30–58, August 2013.

[FSEDV13] Jean-Charles Faugère, Mohab Safey El Din, and Thibaut Verron. On the complexity of computing gröbner bases for quasi-homogeneous systems. In *Proceedings of the 38th International Symposium on Symbolic and Algebraic Computation*, ISSAC '13, pages 189–196, New York, NY, USA, 2013. ACM.

[FSS14] Jean-Charles Faugère, Pierre-Jean Spaenlehauer, and Jules Svartz. Sparse Gröbner bases: the unmixed case. *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation - ISSAC '14*, pages 178–185, 2014. Version 3 in arXiv: 1402.7205 [cs].

[FSS16] Jean-Charles Faugère, Pierre-Jean Spaenlehauer, and Jules Svartz. Computing Small Certificates of Inconsistency of Quadratic Fewnomial Systems. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '16, pages 223–230, New York, NY, USA, 2016. ACM.

[Gal74] André Galligo. A propos du théorème de préparation de weierstrass. In François Norguet, editor, *A propos du théorème de préparation de weierstrass*, volume 409 of *Lecture Notes in Mathematics*, pages 543–579. Springer Berlin Heidelberg, Berlin, Heidelberg, 1974.

[Gal79] André Galligo. Théorème de division et stabilité en géométrie analytique locale. *Annales de l'institut Fourier*, 29(2):107–184, 1979.

[GG13] Joachim von zur Gathen and Jürgen Gerhard. *Modern computer algebra*. Cambridge University Press, Cambridge, 2013.

[GHJZ14] Sergio Galeani, Didier Henrion, Alain Jacquemard, and Luca Zaccarian. Design of Marx generators as a structured eigenvalue assignment. *Automatica*, 50(10):2709–2717, October 2014.

[GHPR12] Călin-Ioan Gheorghiu, Michiel E. Hochstenbach, Bor Plestenjak, and Joost Rommes. Spectral collocation solutions to multiparameter Mathieu's system. *Applied Mathematics and Computation*, 218(24):11990–12000, August 2012.

[Gil84] Robert Gilmer. *Commutative Semigroup Rings*. University of Chicago Press, March 1984.

[Giu84] Marc Giusti. Some effectivity problems in polynomial ideal theory. In John Fitch, editor, *EUROSAM 84*, volume 174, pages 159–171. Springer-Verlag, Berlin/Heidelberg, 1984.

[GKL03] Mark Giesbrecht, Erich Kaltofen, and Wen-shin Lee. Algorithms for computing sparsest shifts of polynomials in power, chebyshev, and pochhammer bases. *Journal of Symbolic Computation*, 36(3-4):401–424, 2003.

[GKZ08] Israel M. Gelfand, Mikhail Kapranov, and Andrei Zelevinsky. *Discriminants, resultants, and multidimensional determinants*. Springer Science & Business Media, 2008.

[GLR05] Israel Gohberg, Peter Lancaster, and Leiba Rodman, editors. *Matrix Polynomials*. Birkhäuser Basel, Basel, 2005.

[GM88] Rüdiger Gebauer and H. Michael Möller. On an installation of Buchberger's algorithm. *Journal of Symbolic Computation*, 6(2):275–286, October 1988.

[GMKP17] Ignacio García-Marco, Pascal Koiran, and Timothee Pecatte. Reconstruction algorithms for sums of affine powers. In *Proceedings of the 2017 ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '17, pages 317–324, New York, NY, USA, 2017. ACM.

[GR10] Mark Giesbrecht and Daniel S Roche. Interpolation of shifted-lacunary polynomials. *Computational Complexity*, 19(3):333–354, 2010.

[Gun87] Sigmund Gundelfinger. Zur theorie der binären formen. *Journal für die reine und angewandte Mathematik*, 100:413–424, 1887.

[GV91] Laureano González-Vega. A Subresultant Theory for Multivariate Polynomials. In *Proceedings of the 1991 International Symposium on Symbolic and Algebraic Computation*, ISSAC '91, pages 79–85, New York, NY, USA, 1991. ACM.

[Har77] Robin Hartshorne. *Algebraic Geometry*. Graduate Texts in Mathematics. Springer-Verlag, New York, 1977.

[Hel92] Uwe Helmke. Waring's problem for binary forms. *Journal of Pure and Applied Algebra*, 80(1):29–45, June 1992.

[HJS13] María Isabel Herrero, Gabriela Jeronimo, and Juan Sabia. Affine solution sets of sparse polynomial systems. *Journal of Symbolic Computation*, 51:34–54, April 2013.

[HJSS02] Joos Heintz, Gabriela Jeronimo, Juan Sabia, and Pablo Solernó. Intersection theory and deformation algorithms: the multi-homogeneous case, 2002. Manuscript.

[HKP04] Michiel E. Hochstenbach, Tomaž Kosir, and Bor Plestenjak. A Jacobi–Davidson Type Method for the Two-Parameter Eigenvalue Problem. *SIAM Journal on Matrix Analysis and Applications*, 26(2):477–497, January 2004.

[HM93] Joos Heintz and Jacques Morgenstern. On the Intrinsic Complexity of Elimination Theory. *Journal of Complexity*, 9(4):471–498, December 1993.

[Hoc16] Melvin Hochster. Depth, Cohen-Macaulay rings, and flatness, 2016. Lecture notes for Math 615, Winter, 2016, University of Michigan.

[HR84] Georg Heinig and Karla Rost. *Algebraic methods for Toeplitz-like matrices and operators*. Springer, 1984.

[HR17] Jonathan D Hauenstein and Jose Israel Rodriguez. Multiprojective witness sets and a trace test. *arXiv:1507.07069 [math]*, 2017.

[HS95] Birkett Huber and Bernd Sturmfels. A polyhedral method for solving sparse polynomial systems. *Mathematics of Computation*, 64(212):1541–1555, 1995.

[HVT04] Huy Tài Hà and Adam Van Tuyl. The regularity of points in multi-projective spaces. *Journal of Pure and Applied Algebra*, 187(1-3):153–167, March 2004.

[IK99] Anthony Iarrobino and Vassil Kanev. *Power Sums, Gorenstein Algebras, and Determinantal Loci*. Lecture Notes in Mathematics. Springer-Verlag, Berlin Heidelberg, 1999.

[Jou91] Jean-Pierre Jouanolou. Le formalisme du résultant. *Advances in Mathematics*, 90(2):117–263, December 1991.

[JS07] Gabriela Jeronimo and Juan Sabia. Computing multihomogeneous resultants using straight-line programs. *Journal of Symbolic Computation*, 42(1):218–235, January 2007.

[Jud98] Kenneth L. Judd. *Numerical Methods in Economics*. MIT Press, 1998.

[Khe02] Amit Khetan. Determinantal Formula for the Chow Form of a Toric Surface. In *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation*, ISSAC '02, pages 145–150, New York, NY, USA, 2002. ACM.

[Kho78] Askold G. Khovanskii. Newton polyhedra and toroidal varieties. *Functional Analysis and Its Applications*, 11(4):289–296, 1978.

[KR84] Joseph P. S. Kung and Gian-Carlo Rota. The invariant theory of binary forms. *Bulletin of the American Mathematical Society*, 10(1):27–86, January 1984.

[KS97] Deepak Kapur and Tushar Saxena. Extraneous Factors in the Dixon Resultant Formulation. In *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation*, ISSAC '97, pages 141–148, New York, NY, USA, 1997. ACM.

[KS99] Aviad Kipnis and Adi Shamir. Cryptanalysis of the HFE Public Key Cryptosystem by Relinearization. In Michael Wiener, editor, *Advances in Cryptology — CRYPTO' 99*, Lecture Notes in Computer Science, pages 19–30. Springer Berlin Heidelberg, 1999.

[KSY94] Deepak Kapur, Tushar Saxena, and Lu Yang. Algebraic and Geometric Reasoning Using Dixon Resultants. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation*, ISSAC '94, pages 99–107, New York, NY, USA, 1994. ACM.

[Kun90] Joseph P. S. Kung. Canonical forms of binary forms: variations on a theme of Sylvester. *Institute for Mathematics and Its Applications*, 19:49–58, 1990.

[Kus76] Anatoli G. Kushnirenko. Newton polytopes and the Bezout theorem. *Functional Analysis and Its Applications*, 10(3):233–235, July 1976.

[KY89] Erich Kaltofen and Lakshman Yagati. Improved sparse multivariate polynomial interpolation algorithms. In G. Goos, J. Hartmanis, D. Barstow, W. Brauer, P. Brinch Hansen, D. Gries, D. Luckham, C. Moler, A. Pnueli, G. Seegmüller, J. Stoer, N. Wirth, and P. Gianni, editors, *Symbolic and Algebraic Computation*, volume 358, pages 467–474. Springer Berlin Heidelberg, Berlin, Heidelberg, 1989.

[Lan02] Serge Lang. *Algebra*. Springer-Verlag, New York,, 2002.

[Lan11] Joseph M. Landsberg. *Tensors: Geometry and Applications*. American Mathematical Society, December 2011.

[Las00] Jean B. Lasserre. Global Optimization with Polynomials and the Problem of Moments. *SIAM J. on Optimization*, 11(3):796–817, March 2000.

[Laz77] Daniel Lazard. Algèbre linéaire sur $K[X_1,\dots ,X_n]$ et élimination. *Bulletin de la Société Mathématique de France*, 105:165–190, 1977.

[Laz81] Daniel Lazard. Resolution des systemes d'equations algebriques. *Theoretical Computer Science*, 15(1):77–110, January 1981.

[Laz83] Daniel Lazard. Gröbner-Bases, Gaussian Elimination and Resolution of Systems of Algebraic Equations. In *Proceedings of the European Computer Algebra Conference on Computer Algebra*, EUROCAL '83, pages 146–156, London, UK, UK, 1983. Springer-Verlag.

[LGY11] Wei Li, Xiao-Shan Gao, and Cum-Ming Yuan. Sparse Differential Resultant. In *Proceedings of the 36th International Symposium on Symbolic and Algebraic Computation*, ISSAC '11, pages 225–232, New York, NY, USA, 2011. ACM.

[Mac02] Francis Sowerby Macaulay. Some formulae in elimination. *Proceedings of the London Mathematical Society*, 1(1):3–27, 1902.

[Mas16] César Massri. Solving a sparse system using linear algebra. *Journal of Symbolic Computation*, 73:157–174, March 2016.

[May97] Ernst W. Mayr. Some Complexity Results for Polynomial Ideals. *Journal of Complexity*, 13(3):303–325, September 1997.

[MB18] Chenqi Mou and Yang Bai. On the Chordality of Polynomial Sets in Triangular Decomposition in Top-Down Style. In *Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*, ISSAC '18, pages 287–294, New York, NY, USA, 2018. ACM.

[Mey08] Carl D. Meyer. *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 2008. OCLC: 836394643.

[MM82] Ernst W. Mayr and Albert R. Meyer. The complexity of the word problems for commutative semigroups and polynomial ideals. *Advances in Mathematics*, 46(3):305–329, December 1982.

[Mor03] Teo Mora. *Solving Polynomial Equation Systems II: Macaulay's Paradigm and Gröbner Technology*. Cambridge University Press, 2003.

[Mor09] Alexander Morgan. *Solving Polynomial Systems Using Continuation for Engineering and Scientific Problems*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, January 2009.

[Mou17] Bernard Mourrain. Fast Algorithm for Border Bases of Artinian Gorenstein Algebras. In *Proceedings of the 2017 ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '17, pages 333–340, New York, NY, USA, 2017. ACM.

[MP00] Bernard Mourrain and Victor Y. Pan. Multivariate Polynomials, Duality, and Structured Matrices. *Journal of Complexity*, 16(1):110–180, March 2000.

[MP10] Andrej Muhič and Bor Plestenjak. On the quadratic two-parameter eigenvalue problem and its linearization. *Linear Algebra and its Applications*, 432(10):2529–2542, May 2010.

[MP13] John M. McNamee and Victor Y. Pan. *Numerical methods for roots of polynomials (II)*. Elsevier, 2013.

[MS87] Alexander Morgan and Andrew Sommese. A homotopy for solving general polynomial systems that respects m-homogeneous structures. *Applied Mathematics and Computation*, 24(2):101–113, November 1987.

[MS04] Diane Maclagan and Gregory G. Smith. Multigraded Castelnuovo-Mumford Regularity. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2004(571), January 2004.

[MS05] Ezra Miller and Bernd Sturmfels. *Combinatorial Commutative Algebra*. Graduate Texts in Mathematics. Springer-Verlag, New York, 2005.

[MT17] Angelos Mantzaflaris and Elias Tsigaridas. Resultants and Discriminants for Bivariate Tensor-product Polynomials. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, page 8, July 2017.

[NRS10] Jiawang Nie, Kristian Ranestad, and Bernd Sturmfels. The algebraic degree of semidefinite programming. *Mathematical Programming*, 122(2):379–405, 2010.

[OO13] Luke Oeding and Giorgio Ottaviani. Eigenvectors of tensors and algorithms for Waring decomposition. *Journal of Symbolic Computation*, 54:9–35, July 2013.

[OSS80] Christian Okonek, Michael Schneider, and Heinz Spindler. *Vector Bundles on Complex Projective Spaces: With an Appendix by S. I. Gelfand*. Modern Birkhäuser Classics. Birkhäuser Basel, 1980.

[Pan01] Victor Y. Pan. *Structured Matrices and Polynomials: Unified Superfast Algorithms*. Birkhäuser Basel, 2001.

[Pan02] Victor Y. Pan. Univariate polynomials: Nearly optimal algorithms for numerical factorization and root-finding. *Journal of Symbolic Computation*, 33(5):701 – 733, 2002.

[Per07] Daniel Perrin. *Algebraic Geometry: An Introduction*. Springer Science & Business Media, December 2007.

[PMH18] Bor Plestenjak, Andrej Muhič, and Pavel Holoborodko. Multipareig 2.5.0.0, 2018. `https://www.mathworks.com/matlabcentral/fileexchange/47844-multipareig`.

[PS96] Paul Pedersen and Bernd Sturmfels. Mixed monomial bases. In Laureano González-Vega and Tomás Recio, editors, *Algorithms in Algebraic Geometry and Applications*, Progress in Mathematics, pages 307–316. Birkhäuser Basel, 1996.

[PS05] Lior Pachter and Bernd Sturmfels. *Algebraic Statistics for Computational Biology*, volume 13. Cambridge University Press, 2005.

[PT17a] Victor Y. Pan and Elias Tsigaridas. Accelerated approximation of the complex roots and factors of a univariate polynomial. *Theoretical Computer Science*, 681:138–145, 2017.

[PT17b] Victor Y. Pan and Elias Tsigaridas. Nearly optimal computations with structured matrices. *Theoretical Computer Science*, 681:117–137, 2017.

[Rém01] Gaël Rémond. Élimination multihomogène. In Yuri Valentinovich Nesterenko and Patrice Philippon, editors, *Introduction to algebraic independence theory*, chapter 5, pages 53–81. Springer Science & Business Media, 2001.

[Rez92] Bruce Reznick. *Sums of even powers of real linear forms*, volume 463. American Mathematical Society, 1992.

[Rez96] Bruce Reznick. Homogeneous Polynomial Solutions to Constant Coefficient PDE's. *Advances in Mathematics*, 117(2):179–192, February 1996.

[Rez13a] Bruce Reznick. On the Length of Binary Forms. In Krishnaswami Alladi, Manjul Bhargava, David Savitt, and Pham Huu Tiep, editors, *Quadratic and Higher Degree Forms*, Developments in Mathematics, pages 207–232. Springer New York, New York, NY, 2013.

[Rez13b] Bruce Reznick. Some new canonical forms for polynomials. *Pacific Journal of Mathematics*, 266(1):185–220, September 2013.

[Rit32] Joseph Fels Ritt. *Differential equations from the algebraic standpoint*, volume 14. American Mathematical Society, 1932.

[RLY18] Jose Israel Rodriguez, Lek-Heng Lim, and Yiling You. Fiber product homotopy method for multiparameter eigenvalue problems. *arXiv:1806.10578 [math]*, June 2018.

[Roj99] Joseph Maurice Rojas. Solving Degenerate Sparse Polynomial Systems Faster. *Journal of Symbolic Computation*, 28(1):155–186, July 1999.

[RT17] Bruce Reznick and Neriman Tokcan. Binary forms with three different relative ranks. *Proceedings of the American Mathematical Society*, 145(12):5169–5177, 2017.

[SDC07] Hal Schenck, Alicia Dickenstein, and David Cox. A case study in bigraded commutative algebra. In Irena Peeva, editor, *Syzygies and Hilbert functions*, pages 74–118. Chapman and Hall/CRC, 2007.

[SEDS17] Mohab Safey El Din and Éric Schost. Bit complexity for multi-homogeneous polynomial system solving - application to polynomial minimization. *Journal of Symbolic Computation*, 2017.

[SFR14] Sylvain Soliman, François Fages, and Ovidiu Radulescu. A constraint solving approach to model reduction by tropical equilibration. *Algorithms for molecular biology : AMB*, 9(1):24–24, 2014.

[Sha13] Igor R. Shafarevich. *Basic algebraic geometry 1, Varieties in projective space*. Springer Berlin, Berlin, 3. ed edition, 2013. Translated by Miles Reid.

[Som99] Martín Sombra. A Sparse Effective Nullstellensatz. *Advances in Applied Mathematics*, 22(2):271–295, February 1999.

[Sop13] Ivan Soprunov. Toric complete intersection codes. *Journal of Symbolic Computation*, 50(Supplement C):374–385, March 2013.

[Spa12] Pierre-Jean Spaenlehauer. Solving multi-homogeneous and determinantal systems. Algorithms - Complexity - Applications. PhD Thesis, 2012.

[SS01] Gunter Scheja and Uwe Storch. *Regular sequences and resultants*. AK Peters/CRC Press, 2001.

[Sto00] Arne Storjohann. *Algorithms for matrix canonical forms*. Doctoral Thesis, ETH Zurich, 2000.

[Stu93] Bernd Sturmfels. Sparse elimination theory. In David Eisenbud and Lorenzo Robbiano, editors, *Proceedings Computational Algebraic Geometry and Commutative Algebra*, pages 264–298. Cambridge University Press, 1993.

[Stu94] Bernd Sturmfels. On the Newton Polytope of the Resultant. *Journal of Algebraic Combinatorics*, 3(2):207–236, April 1994.

[Stu96] Bernd Sturmfels. *Gröbner Bases and Convex Polytopes*. American Mathematical Society, 1996.

[Stu02] Bernd Sturmfels. *Solving Systems of Polynomial Equations*. American Mathematical Society, 2002.

[Sva14] Jules Svartz. *Solving zero-dimensional structured polynomial systems*. phdthesis, Université Pierre et Marie Curie - Paris VI, October 2014.

[SVT06] Jessica Sidman and Adam Van Tuyl. Multigraded regularity: syzygies and fat points. *Contributions to Algebra and Geometry*, 47(1):1–22, 2006.

[SW05] Andrew Sommese and Charles Wampler. *The numerical solution of systems of polynomials: Arising in engineering and science*. World Scientific, January 2005.

[Syl04a] James Joseph Sylvester. An essay on canonical forms, supplement to a sketch of a memoir on elimination, transformation and canonical forms. In *The collected papers of James Joseph Sylvester*, volume 1, pages 203–216. Cambridge University Press, 1904.

[Syl04b] James Joseph Sylvester. On a remarkable discovery in the theory of canonical forms and of hyperdeterminants. In *The collected papers of James Joseph Sylvester*, volume 1, pages 265–283. Cambridge University Press, 1904.

[SZ94] Bernd Sturmfels and Andrei Zelevinsky. Multigraded Resultants of Sylvester Type. *Journal of Algebra*, 163(1):115–127, January 1994.

[Sza10] Agnes Szanto. Multivariate subresultants using Jouanolou matrices. *Journal of Pure and Applied Algebra*, 214(8):1347–1369, August 2010.

[TMVB18] Simon Telen, Bernard Mourrain, and Marc Van Barel. Solving Polynomial Systems via a Stabilized Representation of Quotient Algebras. *SIAM Journal on Matrix Analysis and Applications*, 39(3):1421–1447, October 2018.

[Tra89] Carlo Traverso. Gröbner Trace Algorithms. In *Proceedings of the International Symposium ISSAC'88 on Symbolic and Algebraic Computation*, ISAAC '88, pages 125–138, London, UK, UK, 1989. Springer-Verlag.

[Tru98] Ngô Trung. The Castelnuovo regularity of the Rees algebra and the associated graded ring. *Transactions of the American Mathematical Society*, 350(7):2813–2832, 1998.

[vdW78] Bartel Leendert van der Waerden. On varieties in multiple-projective spaces. *Indagationes Mathematicae (Proceedings)*, 81(1):303–312, January 1978.

[vdW91] Bartel Leendert van der Waerden. *Algebra: Volume II*. Springer-Verlag, New York, 1991.

[Vil18] Gilles Villard. On Computing the Resultant of Generic Bivariate Polynomials. In *Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*, ISSAC '18, pages 391–398, New York, NY, USA, 2018. ACM.

[Vol88] Hans Volkmer. *Multiparameter Eigenvalue Problems and Expansion Theorems*. Lecture Notes in Mathematics. Springer-Verlag, Berlin Heidelberg, 1988.

[VVC94] Jan Verschelde, Pierre Verlinden, and Ronald Cools. Homotopies exploiting Newton polytopes for solving sparse polynomial systems. *SIAM Journal on Numerical Analysis*, 31(3):915–930, June 1994.

[Wey94] Jerzy Weyman. Calculating discriminants by higher direct images. *Transactions of the American Mathematical Society*, 343(1):367–389, January 1994.

[Wey03] Jerzy Weyman. *Cohomology of Vector Bundles and Syzygies*. Cambridge University Press, June 2003.

[WS11] Charles W. Wampler and Andrew J. Sommese. Numerical algebraic geometry and algebraic kinematics. *Acta Numerica*, 20:469–567, May 2011.

[WZ94] Jerzy Weyman and Andrei Zelevinsky. Determinantal formulas for multigraded resultants. *Journal of Algebraic Geometry*, 3(4):569–598, 1994.

Bender, Matías Rafael – Thèse de doctorat – 2019

Résumé :

La résolution de systèmes polynomiaux est l'un des problèmes les plus anciens et importants en mathématiques informatiques et a des applications dans plusieurs domaines des sciences et de l'ingénierie. C'est un problème intrinsèquement difficile avec une complexité au moins exponentielle du nombre de variables. Cependant, dans la plupart des cas, les systèmes polynomiaux issus d'applications ont une structure quelconque. Dans cette thèse, nous nous concentrons sur l'exploitation de la structure liée à la faible densité des supports des polynômes; c'est-à-dire que nous exploitons le fait que les polynômes n'ont que quelques monômes à coefficients non nuls. Notre objectif est de résoudre les systèmes plus rapidement que les estimations les plus défavorables, qui supposent que tous les termes sont présents. Nous disons qu'un système creux est non mixte si tous ses polynômes ont le même polytope de Newton, et mixte autrement. La plupart des travaux sur la résolution de systèmes creux concernent le cas non mixte, à l'exception des résultants creux et des méthodes d'homotopie. Nous développons des algorithmes pour des systèmes mixtes. Nous utilisons les résultantes creux et les bases de Groebner. Nous travaillons sur chaque théorie indépendamment, mais nous les combinons également: nous tirons parti des propriétés algébriques des systèmes associés à une résultante non nulle pour améliorer la complexité du calcul de leurs bases de Groebner; par exemple, nous exploitons l'exactitude du complexe de Koszul pour déduire un critère d'arrêt précoce et éviter tout les réductions à zéro.

De plus, nous développons des algorithmes quasi-optimaux pour décomposer des formes binaires.

Mots clés : Résolution de Systèmes Polynomiaux; Systèmes Polynomiaux Creux; Résultant; Base de Gröbner; Décomposition du Tenseur; Théorie de l'Élimination Creux; Systèmes Multi-homogènes;

Abstract:

Solving polynomial systems is one of the oldest and most important problems in computational mathematics and has many applications in several domains of science and engineering. It is an intrinsically hard problem with complexity at least single exponential in the number of variables. However, in most of the cases, the polynomial systems coming from applications have some kind of structure. In this thesis we focus on exploiting the structure related to the sparsity of the supports of the polynomials; that is, we exploit the fact that the polynomials only have a few monomials with non-zero coefficients. Our objective is to solve the systems faster than the worst case estimates that assume that all the terms are present. We say that a sparse system is unmixed if all its polynomials have the same Newton polytope, and mixed otherwise. Most of the work on solving sparse systems concern the unmixed case, with the exceptions of mixed sparse resultants and homotopy methods. In this thesis, we develop algorithms for mixed systems. We use two prominent tools in nonlinear algebra: sparse resultants and Groebner bases. We work on each theory independently, but we also combine them to introduce new algorithms: we take advantage of the algebraic properties of the systems associated to a non-vanishing resultant to improve the complexity of computing their Groebner bases; for example, we exploit the exactness of some strands of the associated Koszul complex to deduce an early stopping criterion for our Groebner bases algorithms and to avoid every redundant computation (reductions to zero).

In addition, we introduce quasi-optimal algorithms to decompose binary forms.

Keywords: Gröbner Basis; Mixed Sparse System; Multihomogeneous System; Resultant; Solving Polynomial Systems; Sparse Elimination Theory; Tensor decomposition.