ABSTRACT

HOUGH, ZACHARY CLARK. μ-bases and algebraic moving frames: theory and computation. (Under the direction of Hoon Hong and Irina A. Kogan.)

We examine both the theory and computation of μ-bases and algebraic moving frames for a vector of univariate polynomials, along with several related observations. First, we present a new algorithm for computing a μ-basis of the syzygy module of n polynomials in one variable over an arbitrary field $\mathbb{K}$. The algorithm is conceptually different from the previously-developed algorithms by Cox, Sederberg, Chen, Zheng, and Wang for $n = 3$, and by Song and Goldman for an arbitrary n. The algorithm involves computing a "partial" reduced row-echelon form of a $(2d+1) \times n(d+1)$ matrix over $\mathbb{K}$, where d is the maximum degree of the input polynomials. The proof of the algorithm is based on standard linear algebra and is completely self-contained. The proof includes a proof of the existence of the μ-basis and as a consequence provides an alternative proof of the freeness of the syzygy module. The theoretical (worst case asymptotic) computational complexity of the algorithm is $O(d^2n + d^3 + n^2)$. We have implemented this algorithm (HHK) and the one developed by Song and Goldman (SG). Experiments on random inputs indicate that SG is faster than HHK when d is sufficiently large for a fixed n, and that HHK is faster than SG when n is sufficiently large for a fixed d. We also provide a generalization of the HHK algorithm to compute minimal bases for the kernels of $m \times n$ polynomial matrices.

We also characterize a relationship between μ-bases and Gröbner bases for the syzygy module of a vector of n univariate polynomials. Roughly put, we show that "every μ-basis is a minimal TOP Gröbner basis" and that "every minimal TOP Gröbner basis is a μ-basis." Precisely stated, we prove that, for $U \subset \mathbb{K}[s]^n$, the following two statements are equivalent:

(A) U is a μ-basis of syz($\mathbf{a}$)

(B) U is a minimal TOP$_B$-Gröbner basis of syz($\mathbf{a}$) for some ordered basis B of $\mathbb{K}^n$

where TOP$_B$ stands for the TOP ordering among the monomials defined by B.

Furthermore, we give an example showing that *not* every μ-basis is a TOP$_E$-Gröbner basis, where E stands for the standard basis of $\mathbb{K}^n$. We prove that the μ-basis produced by the HHK algorithm is the reduced TOP$_E$-Gröbner basis. We also give an example showing that not every minimal POT$_B$-Gröbner basis is a μ-basis.

We then turn our attention to algebraic moving frames. A *moving frame* at a rational curve is a basis of vectors moving along the curve. When the rational curve is given parametrically by a row vector $\mathbf{a}$ of univariate polynomials, a moving frame with important algebraic properties can be defined by the columns of an invertible polynomial matrix P, such that $\mathbf{a}P = [\gcd(\mathbf{a}), 0, \ldots, 0]$. We call such a matrix an algebraic moving frame. A *degree-optimal moving frame* has column-wise minimal degree, where the degree of a column is defined to be the maximum of the degrees of its components. Algebraic moving frames are closely related to the univariate versions of the celebrated Quillen-Suslin problem, effective Nullstellensatz problem, and syzygy module problem. The focus of these problems, however, is not degree optimality. By contrast, we develop the theory of and an efficient algorithm for constructing a degree-optimal moving frame. We also establish several new theoretical results concerning the degrees of an optimal moving frame and its components.

We present a new degree-optimal moving frame (OMF) algorithm for n relatively prime polynomials (i.e. $\gcd(\mathbf{a}) = 1$). We present a modification for the case when $\gcd(\mathbf{a}) \neq 1$. In addition, we show that any deterministic algorithm for computing a degree-optimal algebraic moving frame can be augmented so that it assigns a degree-optimal moving frame in a $GL_n(\mathbb{K})$-equivariant manner. Equivariance is a crucial property of classical geometric moving frames. We then compare our algorithm with other possible approaches, based on already available algorithms, and show that it is more efficient. We examine other algorithms for computing algebraic moving frames that are not-necessarily-optimal. We also present two new algorithms for computing moving frames based on Gröbner basis computations. One, using POT ordering, is not optimal, while the other, using TOP ordering, is degree-optimal. We also generalize the OMF algorithm to handle matrix inputs.

μ-bases and algebraic moving frames: theory and computation

by
Zachary Clark Hough

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Mathematics

Raleigh, North Carolina

2018

APPROVED BY:

Hoon Hong Co-chair of Advisory Committee	Irina A. Kogan Co-chair of Advisory Committee
Agnes Szanto	Bojko Bakalov

Russell Philbrick

DEDICATION

To my parents.

BIOGRAPHY

Still need to write.

ACKNOWLEDGEMENTS

Still need to write.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER

1

INTRODUCTION

Consider a vector $\mathbf{a}[s]=[a_1(s),a_2(s),\ldots,a_n(s)]$ of univariate polynomials over an arbitrary field $\mathbb{K}$. Such a vector will be our primary object of interest. Let n be the length of **a**, and let d be the degree of **a**, by which we mean the maximum of the degrees of the component polynomials a_i. The main problems we consider with regards to **a** are: computing a μ-basis for the syzygy module of **a**, denoted syz(**a**); and computing a degree-optimal algebraic moving frame at **a**.

It is well-known that the syzygy module of **a**, consisting of linear relations over $\mathbb{K}[s]$ among $a_1(s),\ldots,a_n(s)$:

$$\mathrm{syz}(\mathbf{a})=\{\mathbf{h}\in\mathbb{K}[s]^n\,|\,a_1\,h_1+\cdots+a_n\,h_n=0\}$$

is free.[1] This means that the syzygy module has a basis, and, in fact, infinitely many bases. Moreover, if one views **a** as a parametric curve, then one can take a basis of syz(**a**) and use it as a set of moving lines whose intersections trace out the curve. A μ-basis is a basis with particularly nice properties, which make the problem of finding a μ-basis an important one to study. Namely, a μ-basis is a minimal-degree basis of syzygies. Additional nice properties that μ-bases provide include point-wise linear independence (i.e. linear independence at all s_0 in the algebraic closure $\overline{\mathbb{K}}$) and independence of leading vectors, among others. Furthermore, μ-bases have nice applications in geometric modelling.

[1] Freeness of the syzygy module in the one-variable case can be deduced from the Hilbert Syzygy Theorem [Hil90]. In the multivariable case, the syzygy module of a polynomial vector is not always free (see, for instance, [Cox98a])

In addition to giving one a means to represent a curve as described above, μ-bases also allow one to classify curves (see [CI15]). Specifically, although a μ-basis is not unique, the degrees of the μ-basis are unique, so one can partition the collection of μ-bases by degrees, which then induces a partition on the collection of curves. A μ-basis also can be used to find the implicit equation of a curve, to determine if and where two curves intersect, and to determine if a point lies on a curve, among other applications. They do this more computationally efficiently than other bases of syzygies due to their minimal-degree property.

These applications and properties motivate the development of efficient algorithms for computing μ-bases. The concept of a μ-basis first appeared in [Cox98b], along with an algorithm for computing a μ-basis when the length of the input vector is three. Subsequent algorithms, also for the case when the input length is three, appeared in [ZS01] and [CW02]. The first algorithm for input vectors of arbitrary length appeared in [SG09]. These algorithms all require that $\gcd(\mathbf{a}) = 1$, and they are described in greater detail in Chapter 2.

Our contribution to the μ-basis community is a new algorithm for computing μ-bases for input vectors of arbitrary length and arbitrary GCD. We now briefly describe our contribution. It is well-known that the syzygy module of $\mathbf{a}$, $\operatorname{syz}(\mathbf{a})$, is generated by the set $\operatorname{syz}_d(\mathbf{a})$ of syzygies of degree at most $d = \deg(\mathbf{a})$. The set $\operatorname{syz}_d(\mathbf{a})$ is obviously a $\mathbb{K}$-subspace of $\mathbb{K}[s]^n$. Using the standard monomial basis, it is easy to see that this subspace is isomorphic to the kernel of a certain linear map $A\colon \mathbb{K}^{n(d+1)} \longrightarrow \mathbb{K}^{2d+1}$ (explicitly given by (3.7)). Now we come to the *key* idea: one can *systematically* choose a suitable finite subset of the kernel of A so that the corresponding subset of $\operatorname{syz}_d(\mathbf{a})$ forms a μ-basis. We elaborate on how this is done. Recall that a column of a matrix is called *non-pivotal* if it is either the first column and zero, or it is a linear combination of the previous columns. Now we observe and prove a remarkable fact: the set of indices of non-pivotal columns of A splits into exactly $n-1$ sets of modulo-n-equivalent integers. By taking the smallest representative in each set, we obtain $n-1$ integers, which we call *basic non-pivotal* indices. The set of non-pivotal indices of A is equal to the set of non-pivotal indices of its reduced row-echelon form E. From each non-pivotal column of E, an element of $\ker(A)$ can easily be read off, that, in turn, gives rise to an element of $\operatorname{syz}(\mathbf{a})$, which we call *a row-echelon syzygy*. We prove that the row-echelon syzygies corresponding to the $n-1$ *basic non-pivotal* indices comprise a μ-basis. Thus, a μ-basis can be found by computing the reduced row-echelon form of a single $(2d+1)\times n(d+1)$ matrix A over $\mathbb{K}$. Actually, it is sufficient to compute only a “partial” reduced row-echelon form containing only the basic non-pivotal columns and the preceding pivotal columns.

Performance-wise, our algorithm compares favorably to existing approaches. We show that the new algorithm has theoretical complexity $O(d^2n + d^3 + n^2)$, assuming that the arithmetic takes constant time (which is the case when the field $\mathbb{K}$ is finite). We have implemented our algorithm

(HHK), as well as Song and Goldman's [SG09] algorithm (SG) in Maple [Ber15]. Experiments on random inputs indicate that SG is faster than HHK when d is sufficiently large for a fixed n and that HHK is faster than SG when n is sufficiently large for a fixed d.

Let us now turn our attention to algebraic moving frames. As mentioned above, a nonzero row vector $\mathbf{a} \in \mathbb{K}[s]^n$ defines a parametric curve in $\mathbb{K}^n$. The columns of a matrix $P \in GL_n(\mathbb{K}[s])$ assign a basis of vectors in $\mathbb{K}^n$ at each point of the curve. Here, $GL_n(\mathbb{K}[s])$ denotes the set of invertible $n \times n$ matrices over $\mathbb{K}[s]$, or equivalently, the set of matrices whose columns are *point-wise* linearly independent over the algebraic closure $\overline{\mathbb{K}}$. In other words, the columns of the matrix P can be viewed as a coordinate system, or a frame, that moves along the curve. To be of interest, however, such assignment should not be arbitrary, but instead be related to the curve in a meaningful way. From now on, we require that $\mathbf{a}P = [\gcd(\mathbf{a}), 0, \ldots, 0]$, where $\gcd(\mathbf{a})$ is the monic greatest common divisor of the components of $\mathbf{a}$. We will call a matrix P with the above property an *algebraic moving frame at* $\mathbf{a}$. We observe that for any nonzero monic polynomial $\lambda(s)$, a moving frame at $\mathbf{a}$ is also a moving frame at $\lambda\mathbf{a}$. Therefore, we can obtain an equivalent construction in the projective space $\mathbb{PK}^{n-1}$ by considering only polynomial vectors $\mathbf{a}$ such that $\gcd(\mathbf{a}) = 1$. Then P can be thought of as an element of $PGL_n(\mathbb{K}[s]) = GL_n(\mathbb{K}[s])/cI$, where $c \neq 0 \in \mathbb{K}$ and I is an identity matrix. A canonical map of $\mathbf{a}$ to any of the affine subsets $\mathbb{K}^{n-1} \subset \mathbb{PK}^{n-1}$ produces a rational curve in $\mathbb{K}^n$, and P assigns a projective moving frame at $\mathbf{a}$. We are particularly interested in *degree-optimal* algebraic moving frames – frames that column-wise have minimal degrees, where the degree of a column is defined to be the maximum of the degrees of its components (see Definitions 6 and 62).

The problem of finding a degree-optimal algebraic moving frame is worthwhile to study for various reasons. First of all, it is clear that the first column of an algebraic moving frame P satisfies $\mathbf{a}P_{*1} = \gcd(\mathbf{a})$. That is, the first column is a Bézout vector of $\mathbf{a}$, a vector comprised of the coefficients appearing in the output of the extended Euclidean algorithm. Thus, the first column of a degree-optimal moving frame is a minimal-degree Bézout vector. Our literature search did not yield any efficient algorithm for computing a minimal-degree Bézout vector. Of course, one can compute such a vector by a brute-force method, namely by searching for a Bézout vector of a fixed degree, starting from degree zero, increasing the degree by one, and terminating the search once a Bézout vector is found, but this procedure is very inefficient. As such, computing a degree-optimal moving frame would allow for a more efficient computation of a minimal-degree Bézout vector. Secondly, it is obvious that the last $n-1$ columns of an algebraic moving frame P are syzygies of $\mathbf{a}$. In Proposition 67, we prove that these last $n-1$ columns of P comprise a point-wise linearly independent basis of the syzygy module of $\mathbf{a}$. Thus, the last $n-1$ columns of a degree-optimal moving frame form a basis of the syzygy module of $\mathbf{a}$ of *optimal degree*, i.e. a μ-basis. As mentioned above, μ-bases have a long history of applications in geometric modeling, originating with works by Sederberg and Chen

[SC95], Cox, Sederberg and Chen [Cox98b]. Further development of this topic appeared in [Che05; SG09; JG09; TW14]. Hence, degree-optimal moving frames inherit these important connections to geometric modeling. Furthermore, degree-optimal moving frames have additional application aspects as well. A very important area of applications where utilization of degree-optimal moving frames is beneficial is control theory. In particular, the use of degree-optimal frames can lower differential degrees of "flat outputs" (see, for instance, Polderman and Willems [PW98], Martin, Murray and Rouchon [Mar01], Fabiańska and Quadrat [FQ07], Antritter and Levine [AL10], Imae, Akasawa, and Kobayashi [Ima15]). Another interesting application of algebraic frames can be found in the paper [Elk12] by Elkadi, Galligo and Ba, devoted to the following problem: given a vector of polynomials with gcd 1, find small degree perturbations so that the perturbed polynomials have a large-degree gcd. As discussed in Example 3 of [Elk12], the perturbations produced by the algorithm presented in this paper do not always have minimal degrees. It would be worthwhile to study if the usage of degree-optimal moving frames can decrease the degrees of the perturbations.

The applications available to degree-optimal moving frames motivate the development of efficient algorithms for their construction. Algebraic moving frames appeared in a number of important proofs and constructions under a variety of names. For example, in the constructive proofs of the celebrated Quillen-Suslin theorem [FG90], [LS92], [Can93], [PW95], [LY05], [FQ07], given a polynomial *unimodular* $m \times n$ matrix $\mathbf{A}$, one constructs a unimodular matrix P such that $\mathbf{A}P = [I_m, \mathbf{0}]$, where I_m is an $m \times m$ identity matrix. In the univariate case with $m = 1$, the matrix P is an algebraic moving frame. However, the above works were not concerned with the problem of finding P of optimal degree for every input $\mathbf{A}$. We describe these approaches in greater detail in Chapter 2.

Our contribution to this problem is the theory of and a new algorithm for computing a degree-optimal algebraic moving frame. The advantage of the theory developed here is that it describes how to simultaneously construct a minimal-degree Bézout vector and a μ-basis. Theorem 88 is crucial for this construction, because it shows how a minimal-degree Bézout vector can be read off a Sylvester-type matrix associated with $\mathbf{a}$, the same matrix that is used in Section 3.1 for computing a μ-basis. This theorem leads to an algorithm for computing a degree-optimal moving frame (see Section 5.1). We now briefly elaborate on this algorithm, which consists of the following three steps: (1) build a Sylvester-type $(2d+1)\times(nd+n)$ matrix A, associated with $\mathbf{a}$, where d is the maximal degree of the components of the vector $\mathbf{a}$, and append an additional column to A; (2) run a single partial row-echelon reduction of the resulting $(2d+1)\times(nd+n+1)$ matrix; (3) read off an optimal moving frame from appropriate columns of the partial reduced row-echelon form. We implemented the algorithm in the computer algebra system Maple. The codes and examples are available on the web: `http://www.math.ncsu.edu/~zchough/frame.html`. As mentioned, the advantage of the algorithm is that it simultaneously constructs a minimal-degree Bézout vector and a μ-basis.

Performance-wise, our algorithm compares very favorably to existing approaches. As mentioned above, a degree-optimal moving frame consists of a minimal-degree Bézout vector and a μ-basis. Hence, one may attempt to construct an optimal moving frame by putting together a minimal-degree Bézout vector and a μ-basis. Indeed, as previously discussed, algorithms for computing μ-bases are well-developed. The most straightforward (and computationally inefficient) approach consists of computing the reduced Gröbner basis of the syzygy module with respect to a term-over-position monomial ordering (see Section 3.3). More efficient algorithms have been developed by Cox, Sederberg, and Chen[Cox98b], Zheng and Sederberg [ZS01], Chen and Wang [CW02] for the $n=3$ case, and by Song and Goldman [SG09] and Hong, Hough and Kogan [Hon17] for arbitrary n. The problem of computing a μ-basis also can be viewed as a particular case of the problem of computing optimal-degree kernels of $m\times n$ polynomial matrices of rank m (see for instance Beelen [Bee87], Antoniou, Vardulakis, and Vologiannidis [Ant05], Zhou, Labahn, and Storjohann [Zho12] and references therein). However, these approaches are insufficient due to the lack of an efficient method for computing minimal-degree Bézout vectors.

Alternatively, one can first construct a non-optimal moving frame by algorithms using, for instance, a generalized version of Euclid's extended gcd algorithm, as described by Polderman and Willems in [PW98], or various algorithms presented in the literature devoted to the constructive Quillen-Suslin theorem and the related problem of unimodular completion: Fitchas and Galligo [FG90], Logar and Sturmfels [LS92], Caniglia, Cortiñas, Danón, Heintz, Krick, and Solernó [Can93], Park and Woodburn [PW95], Lombardi and Yengui [LY05], Fabiańska and Quadrat [FQ07], Zhou-Labahn [ZL14]. Then a degree-reduction procedure can be performed, for instance, by computing the Popov normal form of the last $n-1$ columns of a non-optimal moving frame, as discussed in [Bec06], and then reducing the degree of its first column. We discuss this approach in Section 5.4, and demonstrate that it is less efficient than the direct algorithm based on the theory from Sections 4.1 and 4.2 and presented in Section 5.1.

In addition to developing the theory behind an algorithm for computing an optimal moving frame, we prove new results about the degrees of optimal moving frames and its building blocks. These degrees play an important role in the classification of rational curves, because although a degree-optimal moving frame is not unique, its columns have canonical degrees. The list of degrees of the last $n-1$ columns (μ-basis columns) is called the μ-type of an input polynomial vector, and μ-strata analysis was performed in D'Andrea [D'A04], Cox and Iarrobino [CI15]. In Theorem 75, we show that the degree of the first column (Bézout vector) is bounded by the maximal degree of the other columns, while Proposition 76 shows that this is the only restriction that the μ-type imposes on the degree of a minimal Bézout vector. Thus, one can refine the μ-strata analysis to the

(β, μ)-strata analysis, where β denotes the degree of a minimal-degree Bézout vector. This work can have potential applications to rational curve classification problems. In Proposition 91 and Theorem 95, we establish sharp lower and upper bounds for the degree of an optimal moving frame and show that for a generic vector $\mathbf{a}$, the degree of an optimal moving frame equals to the sharp lower bound.

We now summarize the contents of the remaining chapters. In Chapter 2, we conduct a more thorough review of previous approaches for computing μ-bases and algebraic moving frames. In Chapter 3, we examine the theory and computation of μ-bases. After providing a background on μ-bases, we present a new algorithm for computing a μ-basis for a vector of n univariate polynomials. We also provide a generalization of the algorithm to compute minimal bases for the kernels of $m \times n$ polynomial matrices. Additionally, we characterize a relationship between μ-bases and Gröbner bases for the syzygy module of a vector of n univariate polynomials. In Chapter 4, we examine the theory of algebraic moving frames. After providing a background on moving frames, we develop the framework for an algorithm for computing a degree-optimal moving frame at a vector of n univariate polynomials. We also examine the degrees of optimal moving frames. In Chapter 5, we examine the computation of algebraic moving frames. We present a new algorithm for computing degree-optimal moving frames, based on the theory developed in Chapter 4. We present a modification for the case when $\gcd(\mathbf{a}) \neq 1$, and we present an augmentation that allows for the computation of equivariant degree-optimal moving frames as well. We examine other possible approaches, and we compare our algorithm with these approaches. We also generalize our algorithm to handle unimodular matrix inputs.

CHAPTER

2

REVIEW

In this chapter, we provide a more thorough review of the previous work on μ-bases and algebraic moving frames. In particular, we examine some of the previous algorithms for their construction, along with other related observations.

2.1 μ-bases

As mentioned in the Chapter 1, one of the main problems we consider is the problem of computing a μ-basis for a vector **a** of univariate polynomials. The concept of a μ-basis first appeared in [Cox98b], motivated by the search for new, more efficient methods for solving implicitization problems for rational curves, and as a further development of the method of moving lines (and, more generally, moving curves) proposed in [SC95]. Since then, a large body of literature on the applications of μ-bases to various problems involving vectors of univariate polynomials has appeared, such as [Che05; SG09; JG09; TW14].[1] The variety of possible applications motivates the development of algorithms for computing μ-bases. Although a proof of the existence of a μ-basis for arbitrary n appeared already in [Cox98b], the algorithms were first developed for the $n = 3$ case only [Cox98b;

[1] A notion of a μ-basis for vectors of polynomials in two variables also has been developed and applied to the study of rational surfaces in three-dimensional projective space (see, for instance, [Che05; Shi12]). We focus solely on the one-variable case.

ZS01; CW02]. The first algorithm for arbitrary n appeared in [SG09], as a generalization of [CW02]. We now review some of these approaches.

2.1.1 Original definition

The original definition of a μ-basis appeared on p 824 of a paper by Cox, Sederberg, and Chen [Cox98b]. Given a vector $\mathbf{a} \in \mathbb{K}[s]^n$ with $\gcd(\mathbf{a}) = 1$ and $\deg(\mathbf{a}) = d$, a μ-basis of $\mathbf{a}$ is a generating set $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ of syz($\mathbf{a}$) satisfying $\deg(\mathbf{u}_1) + \cdots + \deg(\mathbf{u}_{n-1}) = d$. In the paper, the degrees of $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ were denoted $\mu_1, \dots, \mu_{n-1}$, and the term μ-basis was coined. Also in the paper, the authors prove that a μ-basis always exists for vectors of arbitrary length. The proof of the existence theorem (Theorem 1 on p. 824 of [Cox98b]) appeals to the celebrated Hilbert Syzygy Theorem [Hil90] and utilizes Hilbert polynomials, which first appeared in the same paper [Hil90] under the name of characteristic functions.

The proof of existence for vectors of arbitrary n is non-constructive. However, the authors of [Cox98b] do implicitly suggest an algorithm for the $n = 3$ case. Later, it was explicitly described in the Introduction of [ZS01]. The algorithm relies on the fact that, in the $n = 3$ case, there are only two elements in a μ-basis, and their degrees (denoted as μ_1 and μ_2) can be determined *prior* to computing the basis (see Corollary 2 on p. 811 of [Cox98b] and p. 621 of [ZS01]). Once the degrees are determined, two syzygies are constructed from null vectors of two linear maps $A_1 : \mathbb{K}^{3(\mu_1+1)} \longrightarrow \mathbb{K}^{\mu_1+d+1}$ and $A_2 : \mathbb{K}^{3(\mu_2+1)} \longrightarrow \mathbb{K}^{\mu_2+d+1}$. Special care is taken to ensure that these syzygies are linearly independent over $\mathbb{K}[s]$. These two syzygies comprise a μ-basis. It is not clear, however, how this method can be generalized to arbitrary n. First, as far as we are aware, there is not yet an efficient way to determine the degrees of μ-basis members *a priori*. Second, there is not yet an efficient way for choosing appropriate null vectors so that the resulting syzygies are linearly independent.

2.1.2 Subsequent developments

The next approach we review is by Zheng and Sederberg [ZS01], who gave a different (more efficient) algorithm for the $n = 3$ case, based on Buchberger-type reduction. The algorithm makes use of the fact that, when $\gcd(\mathbf{a}) = 1$, the three obvious syzygies $[a_2, -a_1, 0]$, $[a_3, 0, -a_1]$, and $[0, a_3, -a_2]$ generate syz($\mathbf{a}$). Then Buchberger-type reduction is used to reduce the degree of one of the syzygies at a time. This is done until one of the syzygies reduces to 0, after which the remaining two syzygies comprise a μ-basis. Here, we provide the details of the algorithm. Note, the notation LT refers to the leading term of a polynomial vector. A term of a vector in $\mathbb{K}[s]^n$ is $c\,s^k \mathbf{e}_i$, where c is a constant and $\mathbf{e}_i$ is the i-th standard basis vector. The leading term of a vector is the term of highest degree and highest position, while the constant in the leading term is called the leading coefficient and denoted LC.

Input: $\mathbf{a} \in \mathbb{K}[s]^3$ with $\gcd(\mathbf{a}) = 1$
Output: A μ-basis of $\mathbf{a}$

1. $\mathbf{u}_1 = [a_2, -a_1, 0]^T$, $\mathbf{u}_2 = [a_3, 0, -a_1]^T$, $\mathbf{u}_3 = [0, a_3, -a_2]^T$

2. Choose $\mathbf{u}_i$, $\mathbf{u}_j$ so that $LT(\mathbf{u}_i)$ and $LT(\mathbf{u}_j)$ contain the same basis vector, and $\deg(\mathbf{u}_i) \geq \deg(\mathbf{u}_j)$.

3. Replace $\mathbf{u}_i$ with

$$\mathbf{u}_i \longleftarrow \frac{LCM(LC(\mathbf{u}_i), LC(\mathbf{u}_j))}{LC(\mathbf{u}_i)} \mathbf{u}_i - \frac{LCM(LC(\mathbf{u}_i), LC(\mathbf{u}_j))}{LC(\mathbf{u}_j)} s^{\deg(\mathbf{u}_i) - \deg(\mathbf{u}_j)} \mathbf{u}_j$$

   where LCM represents the least common multiple.

4. If $\mathbf{u}_i = 0$, output remaining nonzero vectors. Else go to step 2.

Example 1. *We demonstrate the update procedure on* $\mathbf{a} = \left[1 + s^2, 2 + s^2, 3 + s^2\right]$.
Start with

$$\mathbf{u}_1 = \begin{bmatrix} 2 + s^2 \\ -1 - s^2 \\ 0 \end{bmatrix}, \ \mathbf{u}_2 = \begin{bmatrix} 3 + s^2 \\ 0 \\ -1 - s^2 \end{bmatrix}, \ \mathbf{u}_3 = \begin{bmatrix} 0 \\ 3 + s^2 \\ -2 - s^2 \end{bmatrix}.$$

1. *Since* $LT(\mathbf{u}_1) = s^2\mathbf{e}_1$ *and* $LT(\mathbf{u}_2) = s^2\mathbf{e}_1$, *we update*

$$\mathbf{u}_1 = \mathbf{u}_1 - \mathbf{u}_2 = \begin{bmatrix} -1 \\ -1 - s^2 \\ 1 + s^2 \end{bmatrix}.$$

2. *Since* $LT(\mathbf{u}_1) = -s^2\mathbf{e}_2$ *and* $LT(\mathbf{u}_3) = s^2\mathbf{e}_2$, *we update*

$$\mathbf{u}_1 = \mathbf{u}_1 + \mathbf{u}_3 = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}.$$

3. *Since* $LT(\mathbf{u}_1) = -\mathbf{e}_1$ *and* $LT(\mathbf{u}_2) = s^2\mathbf{e}_1$, *we update*

$$\mathbf{u}_2 = \mathbf{u}_2 + s^2\mathbf{u}_1 = \begin{bmatrix} 3 \\ 2s^2 \\ -1 - 2s^2 \end{bmatrix}.$$

4. *Since $LT(\mathbf{u}_2) = 2s^2\mathbf{e}_2$ and $LT(\mathbf{u}_3) = s^2\mathbf{e}_2$, we update*

$$\mathbf{u}_2 = \mathbf{u}_2 - 2\mathbf{u}_3 = \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix}.$$

5. *Since $LT(\mathbf{u}_1) = -\mathbf{e}_1$ and $LT(\mathbf{u}_2) = 3\mathbf{e}_1$, we update*

$$\mathbf{u}_2 = \mathbf{u}_2 + 3\mathbf{u}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We then output the μ-basis

$$\left\{ \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3+s^2 \\ -2-s^2 \end{bmatrix} \right\}.$$

A more efficient modification of this procedure was proposed by Chen and Wang [CW02], where instead of reducing based on leading terms, the reduction is done using relationships among leading vectors. Note, the degree of a polynomial vector **h** is the maximum of the degrees of its components, and the leading vector is the vector of deg(**h**) coefficients. Here, we provide the details of the algorithm.

Input: $\mathbf{a} \in \mathbb{K}[s]^3$ with gcd($\mathbf{a}$) = 1
Output: A μ-basis of **a**

1. $\mathbf{u}_1 = [a_2, -a_1, 0]^T$, $\mathbf{u}_2 = [a_3, 0, -a_1]^T$, $\mathbf{u}_3 = [0, a_3, -a_2]^T$

2. Set $m_i = LV(\mathbf{u}_i)$ and $d_i = \deg(\mathbf{u}_i)$ for $i = 1, 2, 3$.

3. Sort d_i so that $d_1 \geq d_2 \geq d_3$ and re-index $\mathbf{u}_i$, m_i.

4. Find real numbers $\alpha_1, \alpha_2, \alpha_3$ such that $\alpha_1 m_1 + \alpha_2 m_2 + \alpha_3 m_3 = 0$.

5. If $\alpha_1 \neq 0$, update $\mathbf{u}_1$ by

$$\mathbf{u}_1 = \alpha_1\mathbf{u}_1 + \alpha_2 s^{d_1-d_2}\mathbf{u}_2 + \alpha_3 s^{d_1-d_3}\mathbf{u}_3$$

   and update m_1 and d_1 accordingly. If $\alpha_1 = 0$, update $\mathbf{u}_2$ by

$$\mathbf{u}_2 = \alpha_2\mathbf{u}_2 + \alpha_3 s^{d_2-d_3}\mathbf{u}_3$$

and update m_2 and d_2 accordingly.

6. If one of the vectors is zero, then output the remaining nonzero vectors and stop; otherwise, go to Step 3.

Example 2. *We demonstrate the modified update procedure on* $\mathbf{a} = \left[1+s^2, 2+s^2, 3+s^2\right]$.
Start with

$$\mathbf{u}_1 = \begin{bmatrix} 2+s^2 \\ -1-s^2 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 3+s^2 \\ 0 \\ -1-s^2 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 0 \\ 3+s^2 \\ -2-s^2 \end{bmatrix}.$$

1. *Since* $m_1 = LV(\mathbf{u}_1) = [1,-1,0]^T$, $m_2 = LV(\mathbf{u}_2) = [1,0,-1]^T$, *and* $m_3 = LV(\mathbf{u}_3) = [0,1,-1]^T$, *we have* $\alpha_1 = 1$, $\alpha_2 = -1$, *and* $\alpha_3 = 1$, *so we update*

$$\mathbf{u}_1 = \mathbf{u}_1 - \mathbf{u}_2 + \mathbf{u}_3 = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$$

and reindex

$$\mathbf{u}_1 = \begin{bmatrix} 3+s^2 \\ 0 \\ -1-s^2 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 0 \\ 3+s^2 \\ -2-s^2 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}.$$

2. *Since* $m_1 = [1,0,-1]^T$, $m_2 = [0,1,-1]^T$, *and* $m_3 = [-1,2,-1]^T$, *we have* $\alpha_1 = 1$, $\alpha_2 = -2$, *and* $\alpha_3 = 1$, *so we update*

$$\mathbf{u}_1 = \mathbf{u}_1 - 2\mathbf{u}_2 + s^2\mathbf{u}_3 = \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix}$$

and reindex

$$\mathbf{u}_1 = \begin{bmatrix} 0 \\ 3+s^2 \\ -2-s^2 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}.$$

3. *Since* $m_1 = [0,1,-1]^T$, $m_2 = [3,-6,3]^T$, *and* $m_3 = [-1,2,-1]^T$, *we have* $\alpha_1 = 0$, $\alpha_2 = 1$, *and* $\alpha_3 = 3$, *so we update*

$$\mathbf{u}_2 = \mathbf{u}_2 + 3\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We then output the μ-basis

$$\left\{ \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3+s^2 \\ -2-s^2 \end{bmatrix} \right\}.$$

Observe that the modified procedure using leading vectors works in fewer steps than the original procedure using leading terms. The algorithm in [CW02] was subsequently generalized to arbitrary n by Song and Goldman [SG09]. The general algorithm starts by observing that the set of the obvious syzygies $\{[\quad -a_j \quad a_i \quad] \,|\, 1 \le i < j \le n\}$ generates syz($\mathbf{a}$), provided gcd($\mathbf{a}$) = 1. Then Buchberger-type reduction is used to reduce the degree of one of the syzygies at a time. It is proved that when such reduction becomes impossible, one is left with exactly $n-1$ non-zero syzygies that comprise a μ-basis. If gcd($\mathbf{a}$) is non-trivial, then the output is a μ-basis multiplied by gcd($\mathbf{a}$). The full details of the algorithm are below.

Input: $\mathbf{a} \in \mathbb{K}[s]^n$ with gcd($\mathbf{a}$) = 1
Output: A μ-basis of $\mathbf{a}$

1. Create the $r = C_2^n$ “obvious" syzygies and label them $\mathbf{u}_1, \ldots, \mathbf{u}_r$.

2. Set $m_i = LV(\mathbf{u}_i)$ and $d_i = \deg(\mathbf{u}_i)$ for $i = 1, \ldots, r$.

3. Sort d_i so that $d_1 \ge d_2 \ge \ldots \ge d_r$ and re-index $\mathbf{u}_i$, m_i.

4. Find real numbers $\alpha_1, \alpha_2, \ldots, \alpha_r$ such that $\alpha_1 m_1 + \alpha_2 m_2 + \cdots + \alpha_r m_r = 0$.

5. Choose the lowest integer j such that $\alpha_j \neq 0$, and update $\mathbf{u}_j$ by setting

$$\mathbf{u}_j = \alpha_j \mathbf{u}_j + \alpha_{j+1} s^{d_j - d_{j+1}} \mathbf{u}_{j+1} + \cdots + \alpha_r s^{d_j - d_r} \mathbf{u}_r.$$

   If $\mathbf{u}_j \equiv 0$, discard $\mathbf{u}_j$ and set $r = r-1$; otherwise set $m_j = LV(\mathbf{u}_j)$ and $d_j = \deg(\mathbf{u}_j)$.

6. If $r = n-1$, then output the $n-1$ non-zero vector polynomials $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ and stop; otherwise, go to Step 3.

2.2 Algebraic moving frames

The other main problem we consider is that of computing a degree-optimal algebraic moving frame at a vector $\mathbf{a}$ of univariate polynomials. An algebraic moving frame at $\mathbf{a} \in \mathbb{K}[s]^n$ is a matrix

$P \in GL_n(\mathbb{K}[s])$ such that $\mathbf{a}P = [\gcd(\mathbf{a}), 0, \ldots, 0]$. A degree-optimal moving frame has minimal column-wise degree. The problem of constructing an algebraic moving frame is a particular case of the well-known problem of providing a constructive proof of the Quillen-Suslin theorem [FG90], [LS92], [Can93], [PW95], [LY05], [FQ07]. In those papers, the multivariate problem is reduced inductively to the univariate case, and then an algorithm for the univariate case is proposed. Those univariate algorithms produce algebraic moving frames. As far as we are aware, the produced moving frames are *usually not* degree-optimal. However, the algorithms are very efficient. We now review some of these approaches.

2.2.1 Fitchas-Galligo algorithm

We start by discussing an algorithm that appeared in [FG90] by Fitchas and Galligo. Before presenting the algorithm, however, we need the following lemma.

Lemma 3. *Let* $n \geq 3$. *Let* $\mathbf{a} = [a_1, \ldots, a_n] \in \mathbb{K}[s]^n$ *be such that* $\gcd(a_1, \ldots, a_n) = 1$. *Then there exist* $k_3, \ldots, k_n \in \mathbb{K}$ *such that* $\gcd(a_1 + k_3 a_3 + \cdots + k_n a_n\,,\ a_2) = 1$.

Proof. Let d be the degree of a_2. Let $\beta_1, \ldots, \beta_d$ be the roots of a_2 in the algebraic closure of $\mathbb{K}$. Consider the following set

$$C = \mathbb{K}^{n-2} \setminus (S_1 \cup \cdots \cup S_d)$$

where

$$S_i = \left\{ (k_3, \ldots, k_n) \in \mathbb{K}^{n-2} \mid a_1\left(\beta_i\right) + k_3\, a_3\left(\beta_i\right) + \cdots + k_n\, a_n\left(\beta_i\right) = 0 \right\}$$

The lemma is immediate from the following claims.

1. $\dim C = n - 2 \geq 1$.

   Proof: Note that S_i is a $\mathbb{K}$-affine space. Since $\gcd(a_1, \ldots, a_n) = 1$, for every $i \in \{1, \ldots, d\}$, at least one of the following is non-zero: $a_1\left(\beta_i\right), a_3\left(\beta_i\right), \ldots, a_n\left(\beta\right)$. Thus dim $S_i \leq n - 3$. In turn dim $C = n - 2 \geq 1$.

2. $\forall (k_3, \ldots, k_n) \in C$ $\gcd(a_1 + k_3 a_3 + \cdots + k_n a_n, a_2) = 1$.

   Proof: Let $(k_3, \ldots, k_n) \in C$. Then for every $i \in \{1, \ldots, d\}$, we have

   $$a_1\left(\beta_i\right) + k_3 a_3\left(\beta_i\right) + \cdots + k_n a_n\left(\beta_i\right) \neq 0$$

   Thus

   $$\gcd(a_1 + k_3 a_3 + \cdots + k_n a_n,\ a_2) = 1.$$

□

We remark that finding the constants k_i as described in the lemma can be completed with a random search. Moreover, for random inputs, $\gcd(a_1, a_2) = 1$ and one can take each $k_i = 0$.

Now, [FG90] describes the following general algorithm for the univariate case:

1. Find $M \in \mathbb{K}[s]^{n\times n}$ with $|M| = 1$ such that $\mathbf{a}M = \mathbf{a}(0)$.
2. Find $T \in \mathbb{K}^{n\times n}$ with $|T| = 1$ such that $\mathbf{a}(0)T = [1,0,\ldots,0]$.
3. $P = MT$.

To find matrix M, they do the following. Assume a_1 and a_2 are relatively prime (otherwise, find constants as described in Lemma 3). Then there exist $f_1, f_2 \in \mathbb{K}[s]$ such that $a_1 f_1 + a_2 f_2 = 1$.
Define

$$A = \begin{bmatrix} f_1(s) & -a_2(s) & f_1(s)[a_3(0)-a_3(s)] & \cdots & f_1(s)[a_n(0)-a_n(s)] \\ f_2(s) & a_1(s) & f_2(s)[a_3(0)-a_3(s)] & \cdots & f_2(s)[a_n(0)-a_n(s)] \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}.$$

Then $\mathbf{a}A = [1, 0, a_3(0), \ldots, a_n(0)]$ and $|A| = 1$. Define

$$B = \begin{bmatrix} a_1(0) & a_2(0) & & & \\ -f_2(0) & f_1(0) & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$

Then $[1, 0, a_3(0), \ldots, a_n(0)]B = [a_1(0), a_2(0), a_3(0), \ldots, a_n(0)]$ and $|B| = 1$.
Let $M = AB$. Then $\mathbf{a}M = \mathbf{a}(0)$ and $|M| = 1$.

The authors of [FG90] do not explicitly describe how to find matrix T with $|T| = 1$ such that $\mathbf{a}(0)T = [1,0,\ldots,0]$. However, this is relatively straightforward, and we present one such method here.

Lemma 4. *Let i be any such that $\mathbf{a}(0)_i \neq 0$. Let j be any such that $1 \le j \le n$ and $j \neq i$. Let*

$$\hat{e}_k = \begin{cases} e_k & \text{if} \quad k \neq j \\ \frac{(-1)^{1+i}}{\mathbf{a}(0)_i} e_k & \text{if} \quad k = j \end{cases}$$

where e_k is the k-th standard unit (row) vector. Finally let

$$T = \begin{bmatrix} \mathbf{a}(0) \\ \hat{e}_1 \\ \vdots \\ \hat{e}_{i-1} \\ \hat{e}_{i+1} \\ \vdots \\ \hat{e}_n \end{bmatrix}^{-1}$$

Then $\mathbf{a}(0)\,T = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$ *and* $\left|T\right| = 1$.

Proof. For the first claim, we observe that

$$\begin{aligned}\mathbf{a}(0)\,T &= \left(\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \mathbf{a}(0) \\ \hat{e}_1 \\ \vdots \\ \hat{e}_{i-1} \\ \hat{e}_{i+1} \\ \vdots \\ \hat{e}_n \end{bmatrix} \right) T \\ &= \left(\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} T^{-1} \right) T = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} T^{-1} T = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}\end{aligned}$$

To show the second claim, consider

$$\left|T^{-1}\right| = \left| \begin{bmatrix} \mathbf{a}(0) \\ \hat{e}_1 \\ \vdots \\ \hat{e}_{i-1} \\ \hat{e}_{i+1} \\ \vdots \\ \hat{e}_n \end{bmatrix} \right| = \frac{(-1)^{1+i}}{\mathbf{a}(0)_i} \left| \begin{bmatrix} \mathbf{a}(0) \\ e_1 \\ \vdots \\ e_{i-1} \\ e_{i+1} \\ \vdots \\ e_n \end{bmatrix} \right|.$$

By subtracting from the first row appropriate multiples of the other rows, and then using basic properties of the determinant, we get

$$\left|T^{-1}\right| = \frac{(-1)^{1+i}}{\mathbf{a}(0)_i} \left| \begin{bmatrix} \mathbf{a}(0)_i\, e_i \\ e_1 \\ \vdots \\ e_{i-1} \\ e_{i+1} \\ \vdots \\ e_n \end{bmatrix} \right| = \left| \begin{bmatrix} e_1 \\ \vdots \\ e_{i-1} \\ e_i \\ e_{i+1} \\ \vdots \\ e_n \end{bmatrix} \right| = 1$$

Therefore $\left|T\right| = 1$. □

We now present the Fitchas-Galligo algorithm.

Input: $\mathbf{a} \neq 0 \in \mathbb{K}[s]^n$ with $\gcd(\mathbf{a}) = 1$
Output: A moving frame at **a**

1. Find constants $k_3, \ldots, k_n$ such that $\gcd(a_1 + k_3 a_3 + \cdots + k_n a_n, a_2) = 1$.

2. Find $f_1, f_2 \in \mathbb{K}[s]$ such that $(a_1 + k_3 a_3 + \cdots + k_n a_n) f_1 + a_2 f_2 = 1$. This can be done by using the Extended Euclidean Algorithm.

3. $M \longleftarrow KAB$, where

$$K = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ k_3 & & 1 & & \\ \vdots & & & \ddots & \\ k_n & & & & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} f_1 & -a_2 & f_1[a_3(0)-a_3] & \cdots & f_1[a_n(0)-a_n] \\ f_2 & a_1' & f_2[a_3(0)-a_3] & \cdots & f_2[a_n(0)-a_n] \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} a_1'(0) & a_2(0) & & & \\ -f_2(0) & f_1(0) & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix},$$

and $a_1' = a_1 + k_3 a_3 + \cdots + k_n a_n$.

4. Find matrix T such that $\mathbf{a}'(0)T = [1, 0, \ldots, 0]$ (where $\mathbf{a}' = [a_1', a_2, \ldots, a_n]$).

   (a) $i \longleftarrow$ such that $\mathbf{a}'(0)_i \neq 0$

   (b) $j \longleftarrow$ such that $1 \leq j \leq n$ and $j \neq i$

   (c) $\hat{e}_k = \begin{cases} e_k & \text{if } k \neq j \\ (-1)^{1+i} / \mathbf{a}'(0)_i \, e_k & \text{if } k = j \end{cases}$ where e_k is the k-th standard unit (row) vector

   (d) $T \longleftarrow \begin{bmatrix} \mathbf{a}'(0) \\ \hat{e}_1 \\ \vdots \\ \hat{e}_{i-1} \\ \hat{e}_{i+1} \\ \vdots \\ \hat{e}_n \end{bmatrix}^{-1}$

5. $P \longleftarrow MT$

The proof of the algorithm is immediate from the fact that $\mathbf{a}P = \mathbf{a}MT = \mathbf{a}'(0)T = [1, 0, \ldots, 0]$ and $|P| = |MT| = |M||T| = 1$.

2.2.2 Algorithm based on Euclidean division

Another simple and elegant algorithm for constructing not-necessarily-optimal moving frames, based on a generalized version of Euclid's extended gcd algorithm, was first mentioned in [LS92] and [PW95]. The authors did not explicitly describe the algorithm, though. However, it can be extracted from Theorem B.1.16 of the book "Introduction to the Mathematical Theory of Systems and Control" by Polderman and Willems [PW98], and it later appeared in [Elk12] as well. We present it here, and we denote it MF_GE (for Moving Frame by Generalized Euclid's algorithm).

```
Input:   $\mathbf{a} \in \mathbb{K}[s]^n$, $\mathbf{a} \neq 0$
Output: $P$, a moving frame at $\mathbf{a}$

  1. Let $k$ be such that $\mathbf{a} = \begin{bmatrix} a_1 & \cdots & a_k & 0 & \cdots & 0 \end{bmatrix}$ where $a_k \neq 0$.

  2. If $k = 1$ then set
         $P = \begin{bmatrix} \frac{1}{\mathrm{lc}(a_1)} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$
     and return $P$. (Here, $\mathrm{lc}(a_1)$ denotes the leading coefficient of $a_1$.)

  3. (Find $q_2, \ldots, q_k, r \in \mathbb{K}[s]^n$ such that $a_1 = q_2 a_2 + \ldots + q_k a_k + r$.)

     (a) $r \leftarrow a_1$

     (b) For $i = 2, \ldots, k$ do
             $q_i \leftarrow \mathrm{quo}(r, a_i)$
             $r \leftarrow \mathrm{rem}(r, a_i)$

  4. $\mathbf{a}' \leftarrow \begin{bmatrix} a_2 & \cdots & a_k & r & 0 & \cdots & 0 \end{bmatrix}$.
```

5. $T \leftarrow \begin{bmatrix} & & & 1 & & & \\ 1 & & & -q_2 & & & \\ & \ddots & & \vdots & & & \\ & & 1 & -q_k & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \in \mathbb{K}[s]^{n\times n},$
   where the q's are placed in the k-th column

6. $P' \leftarrow \text{MF\_GE}(\mathbf{a}')$.

7. $P \leftarrow TP'$.

8. Return P.

The proof of the algorithm is immediate from the fact that, at each step of the algorithm, $\mathbf{a}T = \mathbf{a}'$ and $|T| = 1$.

2.2.3 Fabianska-Quadrat algorithm

Another very efficient algorithm for constructing not-necessarily-optimal algebraic moving frames appeared in [FQ07]. We present it here.

Input: $\mathbf{a} \neq 0 \in \mathbb{K}[s]^n$ with $\gcd(\mathbf{a}) = 1$
Output: A moving frame at $\mathbf{a}$

1. Find constants $k_3, \ldots, k_n$ such that $\gcd(a_1 + k_3 a_3 + \cdots + k_n a_n, a_2) = 1$.

2. Find $f_1, f_2 \in \mathbb{K}[s]$ such that $(a_1 + k_3 a_3 + \cdots + k_n a_n) f_1 + a_2 f_2 = 1$. This can be done by using the Extended Euclidean Algorithm.

3. $P \longleftarrow \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ k_3 & & 1 & & \\ \vdots & & & \ddots & \\ k_n & & & & 1 \end{bmatrix} \begin{bmatrix} f_1 & -a_2 & & & \\ f_2 & a_1' & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -a_3 & \cdots & -a_n \\ 0 & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix},$
   where $a_1' = a_1 + k_3 a_3 + \cdots + k_n a_n$.

The proof of the algorithm is immediate from the fact that $\mathbf{a}P = [1, 0, \ldots, 0]$ and $|P| = 1$.

CHAPTER

3

μ-BASES

This chapter examines both the theory and computation of μ-bases. In Section 3.1, we present a new and alternative algorithm for computing a μ-basis for vectors of arbitrary length n. The proof of the algorithm does not rely on previously established theorems about the freeness of the syzygy module or the existence of a μ-basis, and, therefore, as a by-product, provides an alternative, self-contained, constructive proof of these facts. Before presenting the algorithm, we give a rigorous definition of a μ-basis, describe its characteristic properties, and formulate the problem of computing a μ-basis. We prove several lemmas about the vector space of syzygies of degree at most d, and the role they play in generating the syzygy module. We define the notion of *row-echelon syzygies* and explain how they can be computed. We then present our main theoretical result, Theorem 31, which explicitly identifies a subset of *row-echelon syzygies* that comprise a μ-basis. We present an algorithm for computing a μ-basis, we analyze the theoretical (worst case asymptotic) computational complexity of this algorithm, we discuss implementation and experiments, and we compare the performance of the algorithm presented here with the one described in [SG09].

In Section 3.2, we consider a natural generalization of the μ-basis problem. Namely, that of considering kernels, or nullspaces, of $m \times n$ polynomial matrices of rank m. An optimal degree basis of the nullspace is called a minimal basis. We show that the μ-basis algorithm presented in this chapter can be generalized to compute minimal bases of the kernels of $m \times n$ polynomial matrices.

In Section 3.3, we characterize an interesting relationship between μ-bases and Gröbner bases of the syzygy module. We define monomial orderings on $\mathbb{K}[s]^n$ and highlight two in particular, TOP and POT. We provide a background on Gröbner bases for a submodule in $\mathbb{K}[s]^n$. We then show that every μ-basis is a minimal TOP Gröbner basis and that every minimal TOP Gröbner basis is a μ-basis. We also show that the μ-basis algorithm presented in this chapter produces a μ-basis that is the reduced TOP Gröbner basis of the syzygy module relative to the standard basis of $\mathbb{K}^n$. We also give an example showing that not every minimal POT Gröbner basis is a μ-basis.

3.1 An algorithm for computing μ-bases

In this section, we provide a background on μ-bases, develop the theory of a new μ-basis algorithm, present a new algorithm for computing a μ-basis, analyze the theoretical complexity of the algorithm, and compare the algorithm with the one described in [SG09].

3.1.1 Definitions and problem statement

As mentioned previously, $\mathbb{K}$ denotes a field and $\mathbb{K}[s]$ denotes a ring of polynomials in one indeterminate s. The symbol n will be reserved for the length of the polynomial vector $\mathbf{a}$, whose syzygy module we are considering, and from now on we assume $n > 1$, because for the $n = 1$ case the problem is trivial. The symbol d is reserved for the degree of $\mathbf{a}$. We also will assume that $\mathbf{a}$ is a *non-zero vector*. All vectors are implicitly assumed to be *column vectors*, unless specifically stated otherwise. Superscript ^{T} denotes transposition.

Definition 5 (Syzygy)**.** *Let* $\mathbf{a} = [a_1, \ldots, a_n] \in \mathbb{K}[s]^n$ *be a row n-vector of polynomials. The syzygy set of* $\mathbf{a}$ *is*

$$\operatorname{syz}(\mathbf{a}) = \{\mathbf{h} \in \mathbb{K}[s]^n \mid \mathbf{a}\mathbf{h} = 0\}.$$

We emphasize that $\mathbf{h}$ is by default a column vector and $\mathbf{a}$ is explicitly defined to be a row vector, so that the product $\mathbf{a}\mathbf{h}$ is well-defined. It is easy to check that syz($\mathbf{a}$) is a $\mathbb{K}[s]$-module. To define a μ-basis, we need the following terminology:

Definition 6 (Leading vector)**.** *For* $\mathbf{h} \in \mathbb{K}[s]^n$ *we define the degree and the leading vector of* $\mathbf{h}$ *as follows:*

- $\deg(\mathbf{h}) = \max_{i=1,\ldots,n} \deg(h_i)$.
- $LV(\mathbf{h}) = [\operatorname{coeff}(h_1, t), \ldots, \operatorname{coeff}(h_n, t)]^T \in \mathbb{K}^n$*, where* $t = \deg(\mathbf{h})$ *and* $\operatorname{coeff}(h_i, t)$ *denotes the coefficient of s^t in h_i.*

Example 7. *Let* $\mathbf{h} = \begin{bmatrix} 1-2s-2s^2-s^3 \\ 2+2s+s^2+s^3 \\ -3 \end{bmatrix}$. *Then* $\deg(\mathbf{h}) = 3$ *and* $LV(\mathbf{h}) = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

Before giving the definition of a μ-basis, we state a proposition that asserts the equivalence of several statements, each of which can be taken as a definition of a μ-basis.

Proposition 8. *For a subset* $u = \{\mathbf{u}_1, \dots, \mathbf{u}_{n-1}\} \subset \operatorname{syz}(\mathbf{a})$, *ordered so that* $\deg(\mathbf{u}_1) \le \cdots \le \deg(\mathbf{u}_{n-1})$, *the following properties are equivalent:*

1. *[independence of the leading vectors] The set u generates* $\operatorname{syz}(\mathbf{a})$, *and the leading vectors* $LV(\mathbf{u}_1), \dots, LV(\mathbf{u}_{n-1})$ *are independent over* $\mathbb{K}$.
2. *[minimality of the degrees] The set u generates* $\operatorname{syz}(\mathbf{a})$, *and if* $\mathbf{h}_1, \dots, \mathbf{h}_{n-1}$ *is any generating set of* $\operatorname{syz}(\mathbf{a})$, *such that* $\deg(\mathbf{h}_1) \le \cdots \le \deg(\mathbf{h}_{n-1})$, *then* $\deg(\mathbf{u}_i) \le \deg(\mathbf{h}_i)$ *for* $i = 1, \dots, n-1$.
3. *[sum of the degrees] The set u generates* $\operatorname{syz}(\mathbf{a})$, *and* $\deg(\mathbf{u}_1) + \cdots + \deg(\mathbf{u}_{n-1}) = \deg(\mathbf{a}) - \deg(\gcd(\mathbf{a}))$.
4. *[reduced representation] For every* $\mathbf{h} \in \operatorname{syz}(\mathbf{a})$, *there exist* $g_1, \dots, g_{n-1} \in \mathbb{K}[s]$ *such that* $\deg(g_i) \le \deg(\mathbf{h}) - \deg(\mathbf{u}_i)$ *and*
$$\mathbf{h} = \sum_{i=1}^{n-1} g_i\, \mathbf{u}_i. \tag{3.1}$$
5. *[outer product] There exists a non-zero constant* $\alpha \in \mathbb{K}$ *such that the outer product of* $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ *is equal to* $\alpha\, \mathbf{a}/\gcd(\mathbf{a})$.

Here and below $\gcd(\mathbf{a})$ denotes the greatest common monic devisor of the polynomials $a_1, \dots, a_n$. The above proposition is a slight rephrasing of Theorem 2 in [SG09]. The only notable difference is that we drop the assumption that $\gcd(\mathbf{a}) = 1$ and modify Statements 3 and 5 accordingly. After making an observation that $\operatorname{syz}(\mathbf{a}) = \operatorname{syz}\big(\mathbf{a}/\gcd(\mathbf{a})\big)$, one can easily check that a proof of Proposition 8 can follow the same lines as the proof of Theorem 2 in [SG09]. *We do not use Proposition 8 to derive and justify our algorithm for computing a μ-basis, and therefore we are not including its proof.* However, we do prove some of these properties in Section 4.2.4. We include this proposition to underscore several important properties of a μ-basis and to facilitate comparison with the previous work on the subject.

Following [SG09], we base our definition of a μ-basis on Statement 1 of Proposition 8. We are making this choice, because in the process of proving the existence of a μ-basis, we explicitly construct a set of $n-1$ syzygies for which Statement 1 can be easily verified, while verification of the other statements of Proposition 8 is not immediate. The original definition of a μ-basis (p. 824 of

[Cox98b]) is based on the sum of the degrees property (Statement 2 of Proposition 8). In Section 3.1.9, we discuss the advantages of the original definition.

Definition 9 (μ-basis)**.** *For a non-zero row vector* $\mathbf{a} \in \mathbb{K}[s]^n$*, a subset* $u \subset \mathbb{K}[s]^n$ *of polynomial vectors is called a* μ*-basis of* $\mathbf{a}$*, or, equivalently, a* μ*-basis of* $\operatorname{syz}(\mathbf{a})$*, if the following three properties hold:*

1. u *has exactly* $n-1$ *elements;*
2. $LV(\mathbf{u}_1), \ldots, LV(\mathbf{u}_{n-1})$ *are independent over* $\mathbb{K}$*;*
3. u *is a basis of* $\operatorname{syz}(\mathbf{a})$*, the syzygy module of* $\mathbf{a}$*.*

As we show in Lemma 30 below, the $\mathbb{K}$-linear independence of leading vectors of any set of polynomial vectors immediately implies the $\mathbb{K}[s]$-linear independence of the polynomial vectors themselves. Therefore, a set u satisfying Statement 1 of Proposition 8 is a basis of syz($\mathbf{a}$). Thus, the apparently stronger Definition 9 is, in fact, equivalent to Statement 1 of Proposition 8.

In the next two sections, through a series of lemmas culminating in Theorem 31, we give a self-contained constructive proof of the existence of a μ-basis. This, in turn, leads to an algorithm, presented in Section 3.1.4, for solving the following problem:

Problem:

Input: $\mathbf{a} \neq 0 \in \mathbb{K}[s]^n$, row vector, where $n > 1$ and $\mathbb{K}$ is a computable field.[1]

Output: $M \in \mathbb{K}[s]^{n\times(n-1)}$, such that the columns of M form a μ-basis of $\mathbf{a}$.

Example 10 (Running example)**.** *We will be using the following simple example throughout this section to illustrate the theoretical ideas/findings and the resulting algorithm.*

Input: $\mathbf{a} = \begin{bmatrix} 1+s^2+s^4 & 1+s^3+s^4 & 1+s^4 \end{bmatrix} \in \mathbb{Q}[s]^3$

Output: $M = \begin{bmatrix} -s & 1-2s-2s^2-s^3 \\ 1 & 2+2s+s^2+s^3 \\ -1+s & -3 \end{bmatrix}$

In contrast to the algorithm developed by Song and Goldman in [SG09], the algorithm presented in this section produces a μ-basis even when the input vector $\mathbf{a}$ has a non-trivial greatest common divisor (see Section 3.1.9 for more details).

[1] A field is *computable* if there are algorithms for carrying out the arithmetic $(+,-,\times,/)$ operations among the field elements.

It is worthwhile emphasizing that not every basis of syz($\mathbf{a}$) is a μ-basis. Indeed, let $\mathbf{u}_1$ and $\mathbf{u}_2$ be the columns of matrix M in Example 10. Then $\mathbf{u}_1 + \mathbf{u}_2$ and $\mathbf{u}_2$ is a basis of syz($\mathbf{a}$), but not a μ-basis, because $LV(\mathbf{u}_1 + \mathbf{u}_2) = LV(\mathbf{u}_2)$. A μ-basis is not canonical: for instance, $\mathbf{u}_1$ and $\mathbf{u}_1 + \mathbf{u}_2$ will provide another μ-basis for syz($\mathbf{a}$) in Example 10. However, Statement 2 of Proposition 8 implies that the degrees of the members of a μ-basis are canonical. In [Cox98b], these degrees were denoted by $\mu_1, \ldots, \mu_{n-1}$ and the term "μ-basis" was coined. A more in-depth comparison with previous works on μ-bases and discussion of some related problems can be found in Section 3.1.9.

3.1.2 Syzygies of bounded degree.

From now on, let $\langle \square \rangle_{\mathbb{K}[s]}$ stand for the $\mathbb{K}[s]$-module generated by $\square$. It is known that syz($\mathbf{a}$) is generated by polynomial vectors of degree at most $d = \deg(\mathbf{a})$. To keep our presentation self-contained, we provide a proof of this fact (adapted from Lemma 2 of [SG09]).

Lemma 11. *Let* $\mathbf{a} \in \mathbb{K}[s]^n$ *be of degree* d. *Then* syz($\mathbf{a}$) *is generated by polynomial vectors of degree at most* d.

Proof. Let $\tilde{\mathbf{a}} = \mathbf{a}/gcd(\mathbf{a}) = [\tilde{a}_1, \ldots, \tilde{a}_n]$. For all $i < j$, let

$$\mathbf{u}_{ij} = [\quad -\tilde{a}_j \quad \tilde{a}_i \quad]^T,$$

with $-\tilde{a}_j$ in i-th position, $\tilde{a}_i$ in j-th position, and all the other elements equal to zero. We claim that the $\mathbf{u}_{ij}$'s are the desired polynomial vectors. First note that

$$\max_{1\le i<j\le n} \deg(\mathbf{u}_{ij}) = \max_{1\le i\le n} \tilde{a}_i \le \deg \mathbf{a} = d.$$

It remains to show that $\text{syz}(\mathbf{a}) = \left\langle \mathbf{u}_{ij} \,|\, 1 \le i < j \le n \right\rangle_{\mathbb{K}[s]}$. Obviously we have

$$\text{syz}(\mathbf{a}) = \text{syz}(\tilde{\mathbf{a}}) \tag{3.2}$$

Since $\mathbf{u}_{ij}$ belongs to syz($\tilde{\mathbf{a}}$), we have

$$\text{syz}(\tilde{\mathbf{a}}) \supset \left\langle \mathbf{u}_{ij} \,|\, 1 \le i < j \le n \right\rangle_{\mathbb{K}[s]}. \tag{3.3}$$

Since $\gcd(\tilde{\mathbf{a}}) = 1$, there exists a polynomial vector $\mathbf{f} = [f_1, \ldots, f_n]^T$ such that

$$\tilde{a}_1 f_1 + \cdots + \tilde{a}_n f_n = 1.$$

For any $\mathbf{h} = [h_1, \ldots, h_n]^T \in \operatorname{syz}(\tilde{\mathbf{a}})$, by definition

$$\tilde{a}_1 h_1 + \cdots + \tilde{a}_n h_n = 0.$$

Therefore, for each h_i,

$$\begin{aligned} h_i &= (\tilde{a}_1 f_1 + \cdots + \tilde{a}_n f_n) h_i \\ &= \tilde{a}_1 f_1 h_i + \cdots + \tilde{a}_{i-1} f_{i-1} h_i + \quad \tilde{a}_i f_i h_i \quad + \tilde{a}_{i+1} f_{i+1} h_i + \cdots + \tilde{a}_n f_n h_i \\ &= \tilde{a}_1 f_1 h_i + \cdots + \tilde{a}_{i-1} f_{i-1}\, h_i - f_i \sum_{k \neq i, k=1}^{n} \tilde{a}_k h_k + \tilde{a}_{i+1} f_{i+1} h_i + \cdots + \tilde{a}_n f_n h_i \\ &= \tilde{a}_1 (f_1 h_i - f_i h_1) + \cdots + \tilde{a}_n (f_n h_i - f_i h_n) = \sum_{k \neq i, k=1}^{n} [k, i]\, \tilde{a}_k, \end{aligned}$$

where we denote $f_k\, h_i - f_i h_k$ by $[k, i]$. Since $[k, i] = -[i, k]$, it follows that

$$\mathbf{h} = [h_1, \ldots, h_n]^T = \sum_{1 \le i < j \le n} [i, j][\quad -\tilde{a}_j \quad \tilde{a}_i \quad]^T.$$

That is,

$$h = \sum_{1 \le i < j \le n} (f_i h_j - f_j h_i) \mathbf{u}_{ij}.$$

Therefore

$$\operatorname{syz}(\tilde{\mathbf{a}}) \subset \left\langle \mathbf{u}_{ij} \mid 1 \le i < j \le n \right\rangle_{\mathbb{K}[s]}. \tag{3.4}$$

Putting (3.2), (3.3) and (3.4) together, we have

$$\operatorname{syz}(\mathbf{a}) = \left\langle \mathbf{u}_{ij} \mid 1 \le i < j \le n \right\rangle_{\mathbb{K}[s]}.$$

□

Let $\mathbb{K}[s]_d$ denote the set of polynomials of degree at most d, let $\mathbb{K}[s]_d^n$ denote the set of polynomial vectors of degree at most d, and let

$$\operatorname{syz}_d(\mathbf{a}) = \{\mathbf{h} \in \mathbb{K}[s]_d^n \mid \mathbf{a}\mathbf{h} = 0\}$$

be the set of all syzygies of degree at most d.

It is obvious that $\mathbb{K}[s]_d$ is a $(d+1)$-dimensional vector space over $\mathbb{K}$. Therefore, the set $\mathbb{K}[s]_d^n$ is an $n(d+1)$-dimensional vector space over $\mathbb{K}$. It is straightforward to check that $\operatorname{syz}_d(\mathbf{a})$ is a vector subspace of $\mathbb{K}[s]_d^n$ over $\mathbb{K}$ and, therefore, is finite-dimensional. The following lemma states that a $\mathbb{K}$-basis of the vector space $\mathrm{syz}_d(\mathbf{a})$ generates the $\mathbb{K}[s]$-module $\mathrm{syz}(\mathbf{a})$. The proof of this lemma follows from Lemma 11 in a few trivial steps and is included for the sake of completeness.

Lemma 12. *Let* $\mathbf{a} \in \mathbb{K}[s]^n$ *be of degree* d *and* $\mathbf{h}_1, \dots \mathbf{h}_l$ *be a basis of the* $\mathbb{K}$*-vector space* $\mathrm{syz}_d(\mathbf{a})$. *Then* $\mathrm{syz}(\mathbf{a}) = \langle \mathbf{h}_1, \dots, \mathbf{h}_l \rangle_{\mathbb{K}[s]}$.

Proof. From Lemma 11, it follows that there exist $\mathbf{u}_1, \dots, \mathbf{u}_r \in \mathrm{syz}_d(\mathbf{a})$ that generate the $\mathbb{K}[s]$-module $\mathrm{syz}(\mathbf{a})$. Therefore, for any $\mathbf{f} \in \mathrm{syz}(\mathbf{a})$, there exist $g_1, \dots, g_r \in \mathbb{K}[s]$, such that

$$\mathbf{f} = \sum_{i=1}^{r} g_i \, \mathbf{u}_i. \tag{3.5}$$

Since $\mathbf{h}_1, \dots \mathbf{h}_l$ is a basis of the $\mathbb{K}$-vector space $\mathrm{syz}_d(\mathbf{a})$, there exist $\alpha_{ij} \in \mathbb{K}$ such that

$$\mathbf{u}_i = \sum_{j=1}^{l} \alpha_{ij} \, \mathbf{h}_j. \tag{3.6}$$

Combining (3.5) and (3.6) we get:

$$\mathbf{f} = \sum_{i=1}^{r} g_i \sum_{j=1}^{l} \alpha_{ij} \, \mathbf{h}_j = \sum_{j=1}^{l} \left(\sum_{i=1}^{r} \alpha_{ij} g_j \right) \mathbf{h}_j.$$

□

The next step is to show that the vector space $\mathrm{syz}_d(\mathbf{a})$ is isomorphic to the kernel of a linear map $A \colon \mathbb{K}^{n(d+1)} \to \mathbb{K}^{2d+1}$ defined as follows: for $\mathbf{a} = \sum_{0 \le j \le d} c_j s^j \in \mathbb{K}_d^n[s]$, where $c_j = [c_{1j}, \dots, c_{nj}] \in \mathbb{K}^n$ are *row* vectors, define

$$A = \begin{bmatrix} c_0 & & \\ \vdots & \ddots & \\ c_d & \vdots & c_0 \\ & \ddots & \vdots \\ & & c_d \end{bmatrix} \in \mathbb{K}^{(2d+1) \times n(d+1)}, \tag{3.7}$$

with the blank spaces filled by zeros.

For this purpose, we define an explicit isomorphism between vector spaces $\mathbb{K}[s]_t^m$ and $\mathbb{K}^{m(t+1)}$, where t and m are arbitrary natural numbers. Any polynomial m-vector $\mathbf{h}$ of degree at most t can be written as $\mathbf{h} = w_0 + s\, w_1 + \cdots + s^t\, w_t$ where $w_i = [w_{1i}, \ldots, w_{mi}]^T \in \mathbb{K}^m$. It is clear that the map

$$\sharp_t^m : \mathbb{K}[s]_t^m \to \mathbb{K}^{m(t+1)}$$

$$\mathbf{h} \to \mathbf{h}^{\sharp_t^m} = \begin{bmatrix} w_0 \\ \vdots \\ w_t \end{bmatrix} \tag{3.8}$$

is linear. It is easy to check that the inverse of this map

$$\flat_t^m : \mathbb{K}^{m(t+1)} \to \mathbb{K}[s]_t^m$$

is given by a linear map:

$$v \to v^{\flat_t^m} = S_t^m\, v \tag{3.9}$$

where

$$S_t^m = \begin{bmatrix} I_m & s\,I_m & \cdots s^t\, I_m \end{bmatrix} \in \mathbb{K}[s]^{m \times m(t+1)}.$$

Here I_m denotes the $m \times m$ identity matrix. For the sake of notational simplicity, we will often write $\sharp$, $\flat$ and S instead of $\sharp_t^m$, $\flat_t^m$ and S_t^m when the values of m and t are clear from the context.

Example 13. For

$$\mathbf{h} = \begin{bmatrix} 1-2s-2s^2-s^3 \\ 2+2s+s^2+s^3 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} + s \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix} + s^2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + s^3 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix},$$

we have

$$\mathbf{h}^\sharp = [1, 2, -3, -2, 2, 0, -2, 1, 0, -1, 1, 0]^T.$$

Note that

$$\mathbf{h} = (\mathbf{h}^\sharp)^\flat = S\,\mathbf{h}^\sharp = \begin{bmatrix} I_3 & s I_3 & s^2 I_3 & s^3 I_3 \end{bmatrix} \mathbf{h}^\sharp.$$

With respect to the isomorphisms $\sharp$ and $\flat$, the $\mathbb{K}$-linear map $\mathbf{a}\colon \mathbb{K}[s]_d^n \to \mathbb{K}[s]_{2d}$ corresponds to the $\mathbb{K}$ linear map $A\colon \mathbb{K}^{n(d+1)} \to \mathbb{K}^{2d+1}$ in the following sense:

Lemma 14. *Let* $\mathbf{a} = \sum_{0 \le j \le d} c_j s^j \in \mathbb{K}_d^n[s]$ *and* $A \in \mathbb{K}^{(2d+1)\times n(d+1)}$ *defined as in* (3.7)*. Then for any* $v \in \mathbb{K}^{n(d+1)}$*:*

$$\mathbf{a} v^\flat = (Av)^\flat. \tag{3.10}$$

Proof. A vector $v \in \mathbb{K}^{n(d+1)}$ can be split into $(d+1)$ blocks

$$\begin{bmatrix} w_0 \\ \vdots \\ w_d \end{bmatrix},$$

where $w_i \in \mathbb{K}^n$ are column vectors. For $j < 0$ and $j > d$, let us define $c_j = 0 \in \mathbb{K}^n$. Then Av is a $(2d+1)$-vector with $(k+1)$-th entry

$$(Av)_{k+1} = c_k\, w_0 + c_{k-1}\, w_1 + \cdots + c_{k-d}\, w_d = \sum_{0 \le i \le d} c_{k-i}\, w_i,$$

where $k = 0, \ldots, 2d$. Then

$$\begin{aligned} \mathbf{a} v^\flat &= \mathbf{a} S_d^n\, v = \left(\sum_{0 \le j \le d} c_j s^j \right) \left(\sum_{0 \le i \le d} w_i\, s^i \right) = \sum_{0 \le i,j \le d} c_j\, w_i s^{i+j} \\ &= \sum_{0 \le k \le 2d} s^k \left(\sum_{0 \le i \le d} c_{k-i}\, w_i \right) = \sum_{0 \le k \le 2d} s^k\, (Av)_{k+1} = S_{2d}^1\, (Av) = (Av)^\flat. \end{aligned}$$

□

Example 15. For the row vector **a** in the running example (Example 10), we have $n = 3$, $d = 4$,

$$c_0 = [1,1,1],\ c_1 = [0,0,0],\ c_2 = [1,0,0],\ c_3 = [0,1,0],\ c_4 = [1,1,1]$$

and

$$A = \begin{bmatrix} 1 & 1 & 1 & & & & & & & & & & & & \\ 0 & 0 & 0 & 1 & 1 & 1 & & & & & & & & & \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & & & & \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\ & & & & & & & & & & & & 1 & 1 & 1 \end{bmatrix}.$$

Let $v = [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15]^T$. Then

$$Av = [6,15,25,39,60,33,48,47,42]^T$$

and so

$$(Av)^\flat = S^1_{2d}(Av) = S^1_8(Av) = 6 + 15s + 25s^2 + 39s^3 + 60s^4 + 33s^5 + 48s^6 + 47s^7 + 42s^8.$$

On the other hand, since

$$v^\flat = S^n_d v = S^3_4 v = \begin{bmatrix} 1+4s+7s^2+10s^3+13s^4 \\ 2+5s+8s^2+11s^3+14s^4 \\ 3+6s+9s^2+12s^3+15s^4 \end{bmatrix},$$

we have

$$\begin{aligned} \mathbf{a}v^\flat &= \begin{bmatrix} 1+s^2+s^4 & 1+s^3+s^4 & 1+s^4 \end{bmatrix} \begin{bmatrix} 1+4s+7s^2+10s^3+13s^4 \\ 2+5s+8s^2+11s^3+14s^4 \\ 3+6s+9s^2+12s^3+15s^4 \end{bmatrix} \\ &= 42s^8 + 47s^7 + 48s^6 + 33s^5 + 60s^4 + 39s^3 + 25s^2 + 15s + 6. \end{aligned}$$

We observe that

$$\mathbf{a}v^\flat = (Av)^\flat.$$

Lemma 16. *$v \in \ker(A)$ if and only if $v^\flat \in \operatorname{syz}_d(\mathbf{a})$.*

Proof. Follows immediately from (3.10). □

We conclude this section by describing an explicit generating set for the syzygy module.

Lemma 17. *Let $b_1, \ldots b_l$ comprise a basis of $\ker(A)$, then*

$$\operatorname{syz}(\mathbf{a}) = \left\langle b^\flat_1, \ldots, b^\flat_l \right\rangle_{\mathbb{K}[s]}.$$

Proof. Lemma 16 shows that the isomorphism (3.9) between vector spaces $\mathbb{K}^{n(d+1)}$ and $\mathbb{K}(s)^n_d$ induces an isomorphism between their respective subspaces $\ker(A)$ and $\operatorname{syz}_d(\mathbf{a})$. Therefore, $b^\flat_1, \ldots, b^\flat_l$ is a basis of $\operatorname{syz}_d(\mathbf{a})$. The conclusion then follows from Lemma 12. □

3.1.3 “Row-echelon” generators and μ-bases.

In the previous section, we proved that any basis of ker(A) gives rise to a generating set of syz($\mathbf{a}$). In this section, we show that a particular basis of ker(A), which can be “read off” the reduced row-echelon form of A, contains $n-1$ vectors that give rise to a μ-basis of syz($\mathbf{a}$). In this and the following sections, quo(i, j) denotes the quotient and rem(i, j) denotes the remainder generated by dividing of an integer i by an integer j.

We start with the following important definition:

Definition 18. *A column of any matrix N is called pivotal if it is either the first column and is non-zero or it is linearly independent of all previous columns. The rest of the columns of N are called non-pivotal. The index of a pivotal (non-pivotal) column is called a pivotal (non-pivotal) index.*

From this definition, using induction, it follows that every non-pivotal column can be written as a linear combination of the preceding *pivotal columns*.

We denote the set of pivotal indices of A as p and the set of its non-pivotal indices as q. In the following two lemmas, we show how the specific structure of the matrix A is reflected in the structure of the set of non-pivotal indices q.

Lemma 19 (periodicity)**.** *If $j \in q$ then $j + kn \in q$ for $0 \le k \le \left\lfloor \frac{n(d+1)-j}{n} \right\rfloor$. Moreover,*

$$A_{*j} = \sum_{r<j} \alpha_r A_{*r} \quad \Longrightarrow \quad A_{*j+kn} = \sum_{r<j} \alpha_r A_{*r+kn}, \tag{3.11}$$

where A_{*j} denotes the j-th column of A.

Proof. To prove the statement, we need to examine the structure of the $(2d+1) \times n(d+1)$ matrix A:

$$\begin{bmatrix}
c_{01} & \cdots & c_{0n} & & & & & & & \\
\vdots & \cdots & \vdots & c_{01} & \cdots & c_{0n} & & & & \\
\vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \ddots & & & \\
c_{d1} & \cdots & c_{dn} & \vdots & \cdots & \vdots & \ddots & c_{01} & \cdots & c_{0n} \\
 & & & c_{d1} & \cdots & c_{dn} & \ddots & \vdots & \cdots & \vdots \\
 & & & & & & \ddots & \vdots & \cdots & \vdots \\
 & & & & & & & c_{d1} & \cdots & c_{dn}
\end{bmatrix}. \tag{3.12}$$

The j-th column of A has the first quo$(j-1, n)$ and the last $(d - \text{quo}(j-1, n))$ entries zero. For

$1 \leq j \leq nd$ the $(n+j)$-th column is obtained by shifting all entries of the j-th column down by 1 and then putting an extra zero on the top. We consider two cases:

1. Integer $j = 1$ is in q if and only if the first column of A is zero. From the structure of A it follows that any column indexed by $1 + kn$ is zero and therefore, $(1 + kn) \in q$ for $\left\lfloor \frac{n(d+1)-1}{n} \right\rfloor = d \geq k \geq 0$.

2. Let us embed A in an infinite matrix indexed by integers. By inspection of the structure of A given by (3.12), we see immediately

$$A_{i,r+kn} = A_{i-k,r}. \tag{3.13}$$

   Then, for a non pivotal index $j > 1$ and $0 \leq k \leq \left\lfloor \frac{n(d+1)-j}{n} \right\rfloor$ we have:

$$\begin{aligned}
& A_{*j} = \sum_{r<j} \alpha_r A_{*r} \\
\iff & \forall_{i \in \mathbb{Z}} \; A_{i,j} = \sum_{r=1}^{j-1} \alpha_r A_{i,r} \\
\iff & \forall_{i \in \mathbb{Z}} \; A_{i-k,j} = \sum_{r=1}^{j-1} \alpha_r A_{i-k,r} \qquad \text{(by reindexing the row)} \\
\iff & \forall_{i \in \mathbb{Z}} \; A_{i,j+kn} = \sum_{r=1}^{j-1} \alpha_r A_{i,r+kn} \qquad \text{(from (3.13))} \\
\implies & \; A_{*j+kn} = \sum_{r<j} \alpha_r A_{*r+kn},
\end{aligned}$$

   Therefore $(j + kn) \in q$ for $\left\lfloor \frac{n(d+1)-j}{n} \right\rfloor \geq k \geq 0$ and equation (3.11) holds.

□

Definition 20. Let q be the set of non-pivotal indices. Let $q/(n)$ denote the set of equivalence classes of q modulo n. Then the set $\tilde{q} = \{\min \varrho \,|\, \varrho \in q/(n)\}$ will be called the set of *basic non-pivotal indices*.

Example 21. For the matrix A in Example 15, we have $n = 3$ and $q = \{6, 9, 11, 12, 14, 15\}$. Then $q/(n) = \big\{\{6, 9, 12, 15\}, \{11, 14\}\}\big\}$ and $\tilde{q} = \{6, 11\}$.

Lemma 22. *There are exactly $n-1$ basic non-pivotal indices: $|\tilde{q}| = n-1$.*

Proof. We prove this lemma by showing that $|\tilde{q}| < n$ and $|\tilde{q}| > n-2$.

1. Since there are at most n equivalence classes in q modulo n, it follows from the definition of $\tilde{q}$ that $|\tilde{q}| \leq n$. Moreover, the $(2d+1)$-th row of the last block of n-columns of A is given by the row vector $c_d = (c_{1d}, \ldots, c_{nd}) = LV(a)$, which is non-zero. Thus, there exists $r \in \{1, \ldots, n\}$, such that $c_{rd} \neq 0$. Suppose r is minimal such that $c_{rd} \neq 0$. Then the $(nd+r)$-th column of A is independent from the first $nd + r - 1$ columns (since each of these columns has a zero in the $(2d+1)$-th position). Hence, there exists $r \in \{1, \ldots, n\}$ such that $nd + r$ is a pivotal index. From the periodicity Lemma 19, it follows that for every $k = 0, \ldots d$, index $r + kn$ is pivotal and therefore no integer from the class r modulo n belongs to $\tilde{q}$. Thus $|\tilde{q}| < n$.

2. Assume $|\tilde{q}| \leq n-2$. From the periodicity Lemma 19, it follows that the set of non-pivotal indices is the union of the sets $\{j + kn | j \in \tilde{q}, 0 \leq k \leq l_j\}$, where $l_j \leq d$ is some integer. Therefore

   $$|q| \leq |\tilde{q}|(d+1) \leq (n-2)(d+1) = nd + n - 2d - 2.$$

   On the other hand, $|q| = n(d+1) - |p|$. It is well-known (and easy to check) that $|p| = \text{rank}(A)$. Since rank(A) cannot exceed the number of rows of A, $|p| \leq 2d+1$. Therefore

   $$|q| \geq n(d+1) - (2d+1) = nd + n - 2d - 1.$$

   Contradiction. Hence $|\tilde{q}| > n-2$.

□

From the matrix A we will now construct a square $n(d+1) \times n(d+1)$ matrix V in the following way. For $i \in p$, the i-th column of V has 1 in the i-th row and 0's in all other rows. For $i \in q$ we define the i-th column from the linear relationship

$$A_{*i} = \sum_{\{j \in p \,|\, j < i\}} \alpha_j A_{*j} \tag{3.14}$$

as follows: for $j \in p$ such that $j < i$ we set $V_{ji} = \alpha_j$. All the remaining elements of the column V_{*i} are zero. For simplicity we will denote the i-th column of V as v_i. We note two important properties of V:

1. Matrix V has the same linear relationships among its columns as A.

2. Vectors $\{b_i = e_i - v_i \,|\, i \in q\}$, where by e_i we denote a column vector that has 1 in the i-th position and 0's in all others, comprise a basis of ker(A).

The corresponding syzygies $\{b_i^\flat \mid i \in q\}$ will be called *row-echelon syzygies* because the α's appearing in (3.14) can be read off the reduced row-echelon form E of A. (We remind the reader that the $(2d+1) \times n(d+1)$ matrix E has the following property: for all $i \in q$, the non-zero entries of the i-th column consist of $\{\alpha_j \mid j \in p,\ j < i\}$ and α_j is located in the row that corresponds to the place of j in the *increasingly ordered* list p.) The reduced row-echelon form can be computed using Gauss-Jordan elimination or some other methods.

Example 23. For the matrix A in Example 15, we have $n = 3$, $d = 4$, and

$$V = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & -2 & -1 & 0 & -1 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 3 & 1 & 0 & 0 & -2 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 2 & 1 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & -2 & -1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 2 & 1 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & -1 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.$$

The non-pivotal indices are $q = \{6, 9, 11, 12, 14, 15\}$. We have

$$\begin{aligned}
b_6 &= e_6 - v_6 = [0,1,-1,-1,0,1,0,0,0,0,0,0,0,0,0]^T \\
b_9 &= e_9 - v_9 = [0,1,-1,-1,1,0,-1,0,1,0,0,0,0,0,0]^T \\
b_{11} &= e_{11} - v_{11} = [1,2,-3,-2,2,0,-2,1,0,-1,1,0,0,0,0]^T \\
b_{12} &= e_{12} - v_{12} = [0,1,-1,-1,1,0,-1,1,0,-1,0,1,0,0,0]^T \\
b_{14} &= e_{14} - v_{14} = [-1,1,0,0,0,0,0,1,0,-1,0,0,-1,1,0]^T \\
b_{15} &= e_{15} - v_{15} = [-1,-1,2,1,-1,0,1,0,0,0,0,0,-1,0,1]^T
\end{aligned}$$

and the corresponding row-echelon syzygies are

$$b_6^\flat = \begin{bmatrix} -s \\ 1 \\ -1+s \end{bmatrix} \qquad b_9^\flat = \begin{bmatrix} -s-s^2 \\ 1+s \\ -1+s^2 \end{bmatrix}$$

$$b_{11}^\flat = \begin{bmatrix} 1-2s-2s^2-s^3 \\ 2+2s+s^2+s^3 \\ -3 \end{bmatrix} \qquad b_{12}^\flat = \begin{bmatrix} -s-s^2-s^3 \\ 1+s+s^2 \\ -1+s^3 \end{bmatrix}$$

$$b_{14}^\flat = \begin{bmatrix} -1-s^3-s^4 \\ 1+s^2+s^4 \\ 0 \end{bmatrix} \qquad b_{15}^\flat = \begin{bmatrix} -1+s+s^2-s^4 \\ -1-s \\ 2+s^4 \end{bmatrix}.$$

The following lemma shows a crucial relationship between the row-echelon syzygies. Note that, in this lemma, we use i to denote a non-pivotal index and ι to denote a *basic* non-pivotal index.

Lemma 24. *Let v_r, $r \in \{1,\dots,n(d+1)\}$ denote columns of the matrix V. For $i \in q$, let*

$$b_i = e_i - v_i \tag{3.15}$$

Then for any $\iota \in \tilde{q}$ and any integer k such that $0 \le k \le \left\lfloor \frac{n(d+1)-\iota}{n} \right\rfloor$

$$b^\flat_{\iota+kn} = s^k\, b^\flat_\iota + \sum_{\{j\in p \,|\, j<\iota,\, j+kn\in q\}} \alpha_j\, b^\flat_{j+kn}, \tag{3.16}$$

where constants α_j appear in the expression of the ι-th column of A as a linear combination of the previous pivotal columns:

$$A_{*\iota} = \sum_{\{j\in p \,|\, j<\iota\}} \alpha_j A_{*j}.$$

Proof. We start by stating identities, which we use in the proof. By definition of V, we have for any $j \in p$:

$$v_j = e_j \tag{3.17}$$

and for any $\iota \in \tilde{q}$:

$$v_\iota = \sum_{\{j\in p \,|\, j<\iota\}} \alpha_j v_j = \sum_{\{j\in p \,|\, j<\iota\}} \alpha_j e_j. \tag{3.18}$$

Since V has the same linear relationships among its columns as A, it inherits periodicity property

(3.11). Therefore, for any $\iota \in \tilde{q}$ and any integer k such that $0 \le k \le \left\lfloor \frac{n(d+1)-\iota}{n} \right\rfloor$:

$$v_{\iota+kn} = \sum_{\{j \in p \,|\, j<\iota\}} \alpha_j \, v_{j+kn}. \tag{3.19}$$

We also will use an obvious relationship for any $r \in \{1, \dots, n(d+1)\}$ and $0 \le k \le \left\lfloor \frac{n(d+1)-r}{n} \right\rfloor$:

$$e^\flat_{r+kn} = s^k e^\flat_r \tag{3.20}$$

and the fact that the set $\{1, \dots, n(d+1)\}$ is a disjoint union of the sets p and q. Then

$$\begin{aligned}
b^\flat_{\iota+kn} &= (e_{\iota+kn} - v_{\iota+kn})^\flat = s^k e^\flat_\iota - \sum_{\{j \in p \,|\, j<\iota\}} \alpha_j \, v^\flat_{j+kn} && \text{by (3.15), (3.20) and (3.19)}\\
&= s^k e^\flat_\iota - \sum_{\{j \in p \,|\, j<\iota,\, j+kn \in p\}} \alpha_j \, v^\flat_{j+kn} - \sum_{\{j \in p \,|\, j<\iota,\, j+kn \in q\}} \alpha_j \, v^\flat_{j+kn} && \text{(disjoint union)}\\
&= s^k e^\flat_\iota - \sum_{\{j \in p \,|\, j<\iota,\, j+kn \in p\}} \alpha_j \, e^\flat_{j+kn} - \sum_{\{j \in p \,|\, j<\iota,\, j+kn \in q\}} \alpha_j \, v^\flat_{j+kn} && \text{by (3.17)}\\
&= s^k e^\flat_\iota - \sum_{\{j \in p \,|\, j<\iota\}} \alpha_j e^\flat_{j+kn} + \sum_{\{j \in p \,|\, j<\iota,\, j+kn \in q\}} \alpha_j \left(e^\flat_{j+kn} - v^\flat_{j+kn} \right) && \text{(disjoint union)}\\
&= s^k e^\flat_\iota - \sum_{\{j \in p \,|\, j<\iota\}} s^k \alpha_j \, e^\flat_j + \sum_{\{j \in p \,|\, j<\iota,\, j+kn \in q\}} \alpha_j \, b^\flat_{j+kn} && \text{by (3.20) and (3.15)}\\
&= s^k \left(e_\iota - \sum_{\{j \in p \,|\, j<\iota\}} \alpha_j \, e_j \right)^\flat + \sum_{\{j \in p \,|\, j<\iota,\, j+kn \in q\}} \alpha_j \, b^\flat_{j+kn}\\
&= s^k b^\flat_\iota + \sum_{\{j \in p \,|\, j<\iota,\, j+kn \in q\}} \alpha_j \, b^\flat_{j+kn}. && \text{(3.18) and (3.15)}
\end{aligned}$$

□

Example 25. Continuing with Example 23, where $q = \{6, 9, 11, 12, 14, 15\}$ and $\tilde{q} = \{6, 11\}$ and $p = \{1, 2, 3, 4, 5, 7, 8, 10, 13\}$, we have:

$$\begin{array}{rcl}
b^\flat_9 &=& s\, b^\flat_6 + 1\, b^\flat_6,\\
b^\flat_{12} &=& s^2\, b^\flat_6 + 1\, b^\flat_9 + 0\, b^\flat_{11},\\
b^\flat_{14} &=& s\, b^\flat_{11} + 3\, b^\flat_6 + (-1)\, b^\flat_{11}\\
b^\flat_{15} &=& s^3\, b^\flat_6 + (-1)\, b^\flat_{11} + 1\, b^\flat_{12} + 0\, b^\flat_{14}.
\end{array} \tag{3.21}$$

In the next lemma, we show that the subset of row-echelon syzygies indexed by the $n-1$ basic non-pivotal indices is sufficient to generate syz($\mathbf{a}$).

Lemma 26. *Let $\tilde{q}$ denote the set of basic non-pivotal indices of A. Then*

$$\mathrm{syz}(\mathbf{a}) = \left\langle b_r^\flat \mid r \in \tilde{q} \right\rangle_{\mathbb{K}[s]}.$$

Proof. Since $\{b_i \mid i \in q\}$ comprise a basis of $\ker(A)$, we know from Lemma 17 that $\mathrm{syz}(\mathbf{a}) = \left\langle b_i^\flat \mid i \in q \right\rangle_{\mathbb{K}[s]}$. Equation (3.16) implies that for all $i \in q$, there exist constant β's such that

$$b_i^\flat = s^k\, b_\iota^\flat + \sum_{\{r \in q \mid r < i\}} \beta_r\, b_r^\flat, \tag{3.22}$$

where $\iota \in \tilde{q}$ is equal to i modulo n. It follows that inductively we can express $b_i^\flat$ as a $\mathbb{K}[s]$-linear combination of $\{b_r | r \in \tilde{q}\}$ and the conclusion of the lemma follows. □

Example 27. Continuing with Example 23, we have from (3.21):

$$\begin{aligned}
b_9^\flat &= (s+1)\, b_6^\flat, \\
b_{12}^\flat &= (s^2+s+1)\, b_6^\flat + 0\, b_{11}^\flat, \\
b_{14}^\flat &= 3\, b_6^\flat + (s-1)\, b_{11}^\flat, \\
b_{15}^\flat &= (s^3+s^2+s+1)\, b_6^\flat + (-1)\, b_{11}^\flat.
\end{aligned}$$

We next establish linear independence of the corresponding leading vectors:

Lemma 28. *The leading vectors $LV(b_r^\flat)$, $r \in \tilde{q}$ are linearly independent over $\mathbb{K}$.*

Proof. The leading vector $LV(b_r^\flat)$ is equal to the last non-zero n-block in the $n(d+1)$-vector b_r. By construction, the last non-zero element of b_r is equal to 1 and occurs in the r-th position. Then $LV(b_r^\flat)$ has 1 in $\bar{r} = (r \mod n)$ (the reminder of division of r by n) position. All elements of $LV(b_r^\flat)$ positioned after $\bar{r}$ are zero. Since all integers in $\tilde{q}$ are distinct (modulo n), $LV(b_r^\flat)$, $r \in \tilde{q}$ are linearly independent over $\mathbb{K}$.

□

Example 29. The basic non-pivotal columns of the matrix V in Example (23) are columns 6 and 11.

We previously computed

$$b_6 = e_6 - v_6 = [0,1,-1,-1,0,1,0,0,0,0,0,0,0,0,0]^T$$
$$b_{11} = e_{11} - v_{11} = [1,2,-3,-2,2,0,-2,1,0,-1,1,0,0,0,0]^T.$$

The last non-zero n-blocks of b_6 and b_{11} are $[-1,0,1]$ and $[-1,1,0]$, respectively. These blocks coincide with $LV(b_6^\flat)$ and $LV(b_{11}^\flat)$ computed in Example 23. We observe that these vectors are linearly independent, as expected.

Lemma 30. *Let polynomial vectors* $\mathbf{h}_1,\dots,\mathbf{h}_l \in \mathbb{K}[s]^n$ *be such that* $LV(\mathbf{h}_1),\dots,LV(\mathbf{h}_l)$ *are independent over* $\mathbb{K}$. *Then* $\mathbf{h}_1,\dots,\mathbf{h}_l$ *are independent over* $\mathbb{K}[s]$.

Proof. Assume that $\mathbf{h}_1,\dots,\mathbf{h}_l$ are linearly *dependent* over $\mathbb{K}[s]$, i.e. there exist polynomials $g_1,\dots,g_l \in \mathbb{K}[s]$, not all zero, such that

$$\sum_{i=1}^{l} g_i\,\mathbf{h}_i = 0. \tag{3.23}$$

Let $m = \max_{i=1,\dots,l}\big(\deg(g_i)+\deg(\mathbf{h}_i)\big)$ and let $\mathscr{I}$ be the set of indices on which this maximum is achieved. Then (3.23) implies

$$\sum_{i\in\mathscr{I}} LC(g_i)\,LV(\mathbf{h}_i) = 0,$$

where $LC(g_i)$ is the leading coefficient of g_i and is non-zero for $i \in \mathscr{I}$. This identity contradicts our assumption that $LV(\mathbf{h}_1),\dots,LV(\mathbf{h}_l)$ are linearly independent over $\mathbb{K}$. □

Theorem 31 (Main). *The set* $u = \{b_r^\flat \mid r \in \tilde{q}\}$ *is a* μ*-basis of* $\mathbf{a}$.

Proof. We will check that u satisfies the three conditions of a μ-basis in Definition 9.

1. From Lemma 22, there are exactly $n-1$ elements in $\tilde{q}$. Since $b_{r_1} \neq b_{r_2}$ for $r_1 \neq r_2 \in \tilde{q}$ and since $\flat$ is an isomorphism, the set u contains exactly $n-1$ elements.

2. From Lemma 28, we know that the leading vectors $LV(b_r^\flat)$, $r \in \tilde{q}$ are linearly independent over $\mathbb{K}$.

3. Lemma 26 asserts that the set u generates syz(a). By combining Lemmas 28 and 30, we see that the elements of this set are independent over $\mathbb{K}[s]$. Therefore u is a basis of syz($\mathbf{a}$).

□

Remark 32. *We note that by construction the last non-zero entry of vector b_r is in the r-th position, and therefore*

$$\deg(b_r^\flat) = \lceil r/n \rceil - 1.$$

Thus we can determine the degrees of the μ-basis elements prior to computing the μ-basis from the set of basic non-pivotal indices.

Example 33. For the row vector **a** given in the running example (Example 10), we determined that $\tilde{q} = \{6, 11\}$. Therefore, prior to computing a μ-basis, we can determine the degrees of its members: $\mu_1 = \lceil 6/3 \rceil - 1 = 1$ and $\mu_2 = \lceil 11/3 \rceil - 1 = 3$. We now can apply Theorem 31 and the computation we performed in Example 23 to write down a μ-basis:

$$b_6^\flat = \begin{bmatrix} -s \\ 1 \\ -1+s \end{bmatrix} \text{ and } b_{11}^\flat = \begin{bmatrix} 1-2s-2s^2-s^3 \\ 2+2s+s^2+s^3 \\ -3 \end{bmatrix}.$$

We observe that our degree prediction is correct.

3.1.4 Algorithm

We now describe an algorithm for computing μ-bases of univariate polynomials. We assume that the reader is familiar with Gauss-Jordan elimination (for computing reduced row-echelon forms and in turn null vectors), which can be found in any standard linear algebra textbook. The theory developed in the previous sections can be recast into the following computational steps:

1. *Construct a matrix $A \in \mathbb{K}^{(2d+1)\times n(d+1)}$ whose null space corresponds to* $\operatorname{syz}_d(\mathbf{a})$.

2. *Compute the reduced row-echelon form E of A.*

3. *Construct a matrix $M \in \mathbb{K}[s]^{n\times(n-1)}$ whose columns form a μ-basis of* **a**, as follows:

    (a) Construct the matrix $B \in \mathbb{K}^{n(d+1)\times(n-1)}$ whose columns are the null vectors of E corresponding to its basic non-pivot columns:

    - $B_{\tilde{q}_j, j} = 1$
    - $B_{p_r, j} = -E_{r, \tilde{q}_j}$ for all r
    - All other entries are zero

    where p is the list of the pivotal indices and $\tilde{q}$ is the list of the basic non-pivotal indices of E.

(b) Translate the columns of B into polynomials.

However, steps 2 and 3 do some wasteful operations and they can be improved, as follows:

- Note that step 2 constructs the entire reduced row-echelon form of A, even though we only need $n-1$ null vectors corresponding to its basic non-pivot columns. Hence, it is natural to optimize this step so that only the $n-1$ null vectors are constructed: instead of using Gauss-Jordan elimination to compute the entire reduced row-echelon form, one performs operations column by column only on the pivot columns and basic non-pivot columns. One aborts the elimination process as soon as $n-1$ basic non-pivot columns are found, resulting in a partial reduced row-echelon form of A.
- Note that step 3 constructs the entire matrix B even though many entries are zero. Hence, it is natural to optimize this step so that we bypass constructing the matrix B, but instead construct the matrix M directly from the matrix E. This is possible because the matrix E contains all the information about the matrix B.

Below, we describe the resulting algorithm in more detail and illustrate its operation on our running example (Example 10).

μ-Basis Algorithm

Input: $\mathbf{a} \neq 0 \in \mathbb{K}[s]^n$, row vector, where $n > 1$ and $\mathbb{K}$ is a computable field

Output: $M \in \mathbb{K}[s]^{n\times(n-1)}$ such that its columns form a μ-basis of $\mathbf{a}$

1. *Construct a matrix* $A \in \mathbb{K}^{(2d+1)\times n(d+1)}$ *whose null space corresponds to* $\operatorname{syz}_d(\mathbf{a})$.

   (a) $d \longleftarrow \deg(\mathbf{a})$

   (b) Identify the row vectors $c_0, \ldots, c_d \in \mathbb{K}^n$ such that $\mathbf{a} = c_0 + c_1 s + \cdots + c_d s^d$.

   (c) $$A \longleftarrow \begin{bmatrix} c_0 & & \\ \vdots & \ddots & \\ c_d & \vdots & c_0 \\ & \ddots & \vdots \\ & & c_d \end{bmatrix}$$

2. *Construct the "partial" reduced row-echelon form* E *of* A.

   This can be done by using Gauss-Jordan elimination (forward elimination, backward elimination, and normalization), with the following optimizations:

- Stop the forward elimination as soon as $n-1$ basic non-pivot columns are detected.
- Skip over periodic non-pivot columns.
- Carry out the row operations only on the required columns.

3. *Construct a matrix* $M \in \mathbb{K}[s]^{n\times(n-1)}$ *whose columns form a* μ*-basis of* $\mathbf{a}$.

   Let p be the list of the pivotal indices and let $\tilde{q}$ be the list of the basic non-pivotal indices of E.

```
(a) Initialize an $n \times n-1$ matrix $M$ with 0 in every entry.
(b) For $j = 1,\dots,n-1$
        $r \leftarrow \operatorname{rem}\left(\tilde{q}_j - 1, n\right) + 1$
        $k \leftarrow \operatorname{quo}\left(\tilde{q}_j - 1, n\right)$
        $M_{r,j} \leftarrow M_{r,j} + s^k$
(c) For $i = 1,\dots,|p|$
        $r \leftarrow \operatorname{rem}\left(p_i - 1, n\right) + 1$
        $k \leftarrow \operatorname{quo}\left(p_i - 1, n\right)$
        For $j = 1,\dots,n-1$
            $M_{r,j} \leftarrow M_{r,j} - E_{i,\tilde{q}_j} s^k$
```

Theorem 34. *Let M be the output of the μ-Basis Algorithm on the input $a \in \mathbb{K}[s]^n$. Then the columns of M form a μ-basis for* $\mathbf{a}$.

Proof. In step 1, we construct the matrix A whose null space corresponds to $\operatorname{syz}_d(\mathbf{a})$ as has been shown in Lemma 16. In step 2, we perform partial Gauss-Jordan operations on A to identify the $n-1$ basic non-pivot columns of its reduced row-echelon form E. In Lemma 22, we showed that there are exactly $n-1$ such columns. In step 3, we convert the basic non-pivot columns of E into polynomial vectors, using the $\flat$-isomorphism described in Section 3.1.2, and return these polynomial vectors as columns of the matrix M. From Theorem 31 it follows that the columns of M indeed form a μ-basis of $\mathbf{a}$, because they satisfy the generating, leading vector, and linear independence conditions of Definition 9 of a μ-basis. □

Example 35. *We trace the algorithm (with partial Gauss-Jordan) on the input vector from Example 10:*

$$\mathbf{a} = \begin{bmatrix} 1+s^2+s^4 & 1+s^3+s^4 & 1+s^4 \end{bmatrix} \in \mathbb{Q}[s]^3$$

1. *Construct a matrix* $A \in \mathbb{K}^{(2d+1)\times n(d+1)}$ *whose null space corresponds to* $\operatorname{syz}_d(\mathbf{a})$:

(a) $d \longleftarrow 4$

(b) $c_0, c_1, c_2, c_3, c_4 \longleftarrow [\,1\ 1\ 1\,], [\,0\ 0\ 0\,], [\,1\ 0\ 0\,], [\,0\ 1\ 0\,], [\,1\ 1\ 1\,]$

(c) $$A \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
0 & 0 & 0 & 1 & 1 & 1 & & & & & & & & & \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

A blank indicates that the entry is zero due to structural reasons.

2. *Construct the "partial" reduced row-echelon form E of A*:

For this step, we will maintain/update the following data structures.

- E: the matrix initialized with A and updated by the Gauss-Jordan process.
- p: the set of the pivotal indices found.
- $\tilde{q}$: the set of the basic non-pivotal indices found.
- O: the list of the row operations, represented as follows.

  (i, i') : swap $E_{i,j}$ with $E_{i',j}$

  (i, w, i'): $E_{i,j} \longleftarrow E_{i,j} + w \cdot E_{i',j}$

  where j is the current column index.

We will also indicate the update status of the columns of E using the following color codings.

gray	:	not yet updated
blue	:	pivot
red	:	basic non-pivot
brown	:	periodic non-pivot

Now we show the trace.

(a) Initialize.

$p \longleftarrow \{\,\}$

$\tilde{q} \longleftarrow \{\,\}$

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
0 & 0 & 0 & 1 & 1 & 1 & & & & & & & & & \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

$O \longleftarrow [\,]$

(b) $j \longleftarrow 1$

Carry out the row operations in O on column 1. (Nothing to do.)

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
0 & 0 & 0 & 1 & 1 & 1 & & & & & & & & & \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Identify column 1 as a pivot.

$p \longleftarrow \{1\}$

$\tilde{q} \longleftarrow \{\,\}$

Carry out the row operations $(3,-1,1),(5,-1,1)$ on column 1.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & 0 & 0 & 1 & 1 & 1 & & & & & & & & & \\
 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Append $(3,-1,1),(5,-1,1)$ to O.

$O \longleftarrow [(3,-1,1),(5,-1,1)]$

(c) $j \longleftarrow 2$

Carry out the row operations in O on column 2.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & 0 & 0 & 1 & 1 & 1 & & & & & & & & & \\
 & -1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Identify column 2 as a pivot.

$p \longleftarrow \{1,2\}$

$\tilde{q} \longleftarrow \{\}$

Carry out the row operations $(3,2),(4,1,2)$ on column 2.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & 0 & 1 & 1 & 1 & & & & & & & & & \\
 & & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Append $(3,2),(4,1,2)$ to O.

$O \longleftarrow [(3,-1,1),(5,-1,1),(3,2),(4,1,2)]$

(d) $j \longleftarrow 3$

Carry out the row operations in O on column 3.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 1 & 1 & 1 & & & & & & & & & \\
 & & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Identify column 3 as a pivot.

$p \longleftarrow \{1,2,3\}$

$\tilde{q} \longleftarrow \{\,\}$

Carry out the row operation $(4,3)$ on column 3.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 1 & 1 & 1 & & & & & & & & & \\
 & & -1 & 0 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Append $(4,3)$ to O.

$O \longleftarrow [(3,-1,1),(5,-1,1),(3,2),(4,1,2),(4,3)]$

(e) $j \longleftarrow 4$

Carry out the row operations in O on column 4.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 1 & 1 & & & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Identify column 4 as a pivot.

$p \longleftarrow \{1,2,3,4\}$

$\tilde{q} \longleftarrow \{\,\}$

Carry out the row operation $(6,-1,4)$ on column 4.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 1 & 1 & & & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Append $(6,-1,4)$ to O.

$O \longleftarrow [(3,-1,1),(5,-1,1),(3,2),(4,1,2),(4,3),(6,-1,4)]$

(f) $j \longleftarrow 5$

Carry out the row operations in O on column 5.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 1 & & & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Identify column 5 as a pivot.

$p \longleftarrow \{1,2,3,4,5\}$

$\tilde{q} \longleftarrow \{\,\}$

No row operations needed on column 5.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 1 & & & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Nothing to append to O.

$O \longleftarrow [(3,-1,1),(5,-1,1),(3,2),(4,1,2),(4,3),(6,-1,4)]$

(g) $j \longleftarrow 6$

Carry out the row operations in O on column 6.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & & & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Identify column 6 as a basic non-pivot: column 6 is non-pivotal because it does not have non-zero entries below the 5-th row and therefore it is a linear combination of the five previous pivotal columns: $E_{*6} = -E_{*2} + E_{*3} + E_{*4}$. Column 6 is basic because its index is minimal in its equivalence class $q/(3)$.

$p \longleftarrow \{1,2,3,4,5\}$

$\tilde{q} \longleftarrow \{6\}$

No row operations needed on column 6.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & & & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Nothing to append to O.

$O \longleftarrow [(3,-1,1),(5,-1,1),(3,2),(4,1,2),(4,3),(6,-1,4)]$

(h) $j \longleftarrow 7$

Carry out the row operations in O on column 7.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & 1 & & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Identify column 7 as a pivot.

$p \longleftarrow \{1,2,3,4,5,7\}$

$\tilde{q} \longleftarrow \{6\}$

Carry out the row operations $(7,6)$ on column 7.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & 1 & & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Append $(7,6)$ to O.

$O \longleftarrow [(3,-1,1),(5,-1,1),(3,2),(4,1,2),(4,3),(6,-1,4),(7,6)]$

(i) $j \longleftarrow 8$

Carry out the row operations in O on column 8.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Identify column 8 as a pivot.

$p \longleftarrow \{1,2,3,4,5,7,8\}$

$\tilde{q} \longleftarrow \{6\}$

No row operations needed on column 8.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Nothing to append to O.

$O \longleftarrow [(3,-1,1),(5,-1,1),(3,2),(4,1,2),(4,3),(6,-1,4),(7,6)]$

(j) $j \longleftarrow 9$

Identify column 9 as periodic non-pivot.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & & & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & & & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

(k) $j \longleftarrow 10$

Carry out the row operations in O on column 10.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & & 0 & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Identify column 10 as a pivot.

$p \longleftarrow \{1,2,3,4,5,7,8,10\}$

$\tilde{q} \longleftarrow \{6\}$

No row operations needed on column 10.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & & 0 & & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Nothing to append to O.

$O \longleftarrow [(3,-1,1),(5,-1,1),(3,2),(4,1,2),(4,3),(6,-1,4),(7,6)]$

(l) $j \longleftarrow 11$

Carry out the row operations in O on column 11.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & & 0 & 0 & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Identify column 11 as a basic non-pivot.

$p \longleftarrow \{1,2,3,4,5,7,8,10\}$

$\tilde{q} \longleftarrow \{6,11\}$

No row operations needed on column 11.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & 1 & 1 & & & & & & & & & & & & \\
 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & & 0 & 0 & & & & \\
 & & -1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & & & & \\
 & & & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & & & \\
 & & & & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
 & & & & & & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

Nothing to append to O.

$O \longleftarrow [(3,-1,1),(5,-1,1),(3,2),(4,1,2),(4,3),(6,-1,4),(7,6)]$

We have identified $n-1$ basic non-pivot columns, so we abort forward elimination.

(m) Perform backward elimination on the pivot columns and basic non-pivot columns.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & & & & & 0 & & & & & -1 & & & & \\
 & -1 & & & & 1 & & & & & 2 & & & & \\
 & & -1 & & & -1 & & & 1 & & -3 & & & & \\
 & & & 1 & & 1 & & & 0 & & 2 & 1 & & & \\
 & & & & 1 & 0 & & & 0 & & -2 & 0 & 1 & 1 & 1 \\
 & & & & & & 1 & & 0 & & 2 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & & -1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

(n) Perform normalization on the pivot columns and basic non-pivot columns.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc}
1 & & & & & 0 & & & & & -1 & & & & \\
 & 1 & & & & -1 & & & & & -2 & & & & \\
 & & 1 & & & 1 & & & 1 & & 3 & & & & \\
 & & & 1 & & 1 & & & 0 & & 2 & 1 & & & \\
 & & & & 1 & 0 & & & 0 & & -2 & 0 & 1 & 1 & 1 \\
 & & & & & & 1 & & 0 & & 2 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & & -1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

3. *Construct a matrix $M \in \mathbb{K}[s]^{n\times(n-1)}$ whose columns form a μ-basis of* $\mathbf{a}$:

(a) $M \longleftarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

(b) $M \longleftarrow \begin{bmatrix} 0 & 0 \\ 0 & s^3 \\ s & 0 \end{bmatrix}$

(c) $M \longleftarrow \begin{bmatrix} -s & 1-2s-2s^2-s^3 \\ 1 & 2+2s+s^2+s^3 \\ -1+s & -3 \end{bmatrix}$

3.1.5 Theoretical Complexity Analysis

In this subsection, we analyze the theoretical (asymptotic worst case) complexity of the μ-basis algorithm given in the previous subsection. We will do so under the assumption that the time for any arithmetic operation is constant.

Theorem 36. *The complexity of the μ-basis algorithm given in the previous section is*

$$O(d^2n + d^3 + n^2).$$

Proof. We will trace the theoretical complexity for each step of the algorithm.

1. (a) To determine d, we scan through each of the n polynomials in **a** to identify the highest degree term, which is always $\leq d$. Thus, the complexity for this step is $O(dn)$.

   (b) We identify $n(d+1)$ values to make up $c_0, \ldots, c_d$. Thus, the complexity for this step is $O(dn)$.

   (c) We construct a matrix with $(2d+1)n(d+1)$ entries. Thus, the complexity for this step is $O(d^2n)$.

2. With the partial Gauss-Jordan elimination, we perform row operations only on the (at most) $2d+1$ pivot columns of A and the $n-1$ basic non-pivot columns of A. Thus, we perform Gauss-Jordan elimination on a $(2d+1)\times(2d+n)$ matrix. In general, for a $k \times l$ matrix, Gauss-Jordan elimination has complexity $O(k^2l)$. Thus, the complexity for this step is $O(d^2(d+n))$.

3. (a) We fill 0 into the entries of an $n \times (n-1)$ matrix M. Thus, the complexity of this step is $O(n^2)$.

   (b) We update entries of the matrix $n-1$ times. Thus, the complexity of this step is $O(n)$.

   (c) We update entries of the matrix $|p| \times (n-1)$ times. Note that $|p| = \text{rank}(A) \leq 2d+1$. Thus the complexity of this step is $O(dn)$.

By summing up, we have

$$O\left(dn+dn+d^2n+d^2(d+n)+n^2+n+dn\right)=O\left(d^2n+d^3+n^2\right)$$

□

Remark 37. Note that the n^2 term in the above complexity is solely due to step 3(a), where the matrix M is initialized with zeros. If one uses a *sparse* representation of the matrix (storing only non-zero elements), then one can skip the initialization of the matrix M. As a result, the complexity can be improved to $O\left(d^2n+d^3\right)$.

Remark 38. [Comparison with Song-Goldman Algorithm] As far as we are aware, the theoretical complexity of the algorithm by Song and Goldman [SG09] has not yet been published. Here we roughly estimate the complexity of this algorithm to be $O(dn^5+d^2n^4)$. It will require a more rigorous analysis to prove/refute this apparent complexity, which is beyond the scope of this section. For the readers' convenience, we reproduce the algorithm published in [SG09] on pp. 220 – 221 in our notation, before analyzing its complexity.

Input: $\mathbf{a}\in\mathbb{K}[s]^n$ with $\gcd(\mathbf{a})=1$
Output: A μ-basis of $\mathbf{a}$

1. Create the $r=C_2^n$ "obvious" syzygies as described in Lemma 11 and label them $\mathbf{u}_1,\dots,\mathbf{u}_r$.
2. Set $m_i=LV(\mathbf{u}_i)$ and $d_i=\deg(\mathbf{u}_i)$ for $i=1,\dots,r$.
3. Sort d_i so that $d_1\ge d_2\ge\ldots\ge d_r$ and re-index $\mathbf{u}_i$, m_i.
4. Find real numbers $\alpha_1,\alpha_2,\dots,\alpha_r$ such that $\alpha_1m_1+\alpha_2m_2+\cdots+\alpha_rm_r=0$.
5. Choose the lowest integer j such that $\alpha_j\neq 0$, and update $\mathbf{u}_j$ by setting

   $$\mathbf{u}_j=\alpha_j\mathbf{u}_j+\alpha_{j+1}s^{d_j-d_{j+1}}\mathbf{u}_{j+1}+\cdots+\alpha_rs^{d_j-d_r}\mathbf{u}_r.$$

   If $\mathbf{u}_j\equiv 0$, discard $\mathbf{u}_j$ and set $r=r-1$; otherwise set $m_j=LV(\mathbf{u}_j)$ and $d_j=\deg(\mathbf{u}_j)$.
6. If $r=n-1$, then output the $n-1$ non-zero vector polynomials $\mathbf{u}_1,\dots,\mathbf{u}_{n-1}$ and stop; otherwise, go to Step 3.

Finding a null vector in step 4 by partial Gauss-Jordan elimination requires performing row operations on (at most) $n+1$ columns. Since each column contains n entries, we conclude that this step has complexity $O(n^3)$. Performing the “update" operation in step 5 of the algorithm has complexity $O(dn^2)$. Step 6 implies that, in the worst case, the algorithm repeats steps 4 and 5 at most $\left(d\,\frac{n(n-1)}{2} - d\right)$ times. The reason is as follows. Since the algorithm starts with the $C_2^n = \frac{n(n-1)}{2}$ obvious syzygies and each has degree $\leq d$, the (worst case) total degree of the syzygies at the beginning of the algorithm is $d\,\frac{n(n-1)}{2}$. The algorithm ends only when the total degree is d. If each repetition of steps 4 and 5 reduces the total degree by 1, then the steps are repeated $\left(d\,\frac{n(n-1)}{2} - d\right)$ times. Thus, the total computational complexity appears to be $O(dn^5 + d^2n^4)$.

3.1.6 Implementation

We implemented the μ-basis algorithm presented in this section and the one described in Song-Goldman [SG09]. For the sake of simplicity, from now on, we will call these two algorithms HHK and SG. We now discuss our implementation. We implemented both algorithms (HHK and SG) in the computer algebra system Maple [Ber15]. The codes and examples are available on the web:

`http://www.math.ncsu.edu/~zchough/mubasis.html`

We post two versions of the code:

`program_rf` : over rational number field $\mathbb{Q}$.

`program_ff` : over finite field $\mathbb{F}_p$ where p is an arbitrary prime number.

Now we explain how the two algorithms (HHK and SG) have been implemented.

- Although both algorithms could be written in a non-interpreted language such as the C-language, making the running time significantly shorter, we implemented both algorithms in Maple [Ber15] for the following reasons.

    1. Maple allows fast prototyping of the algorithms, making it easier to write and read the programs written in Maple.
    2. It is expected that potential applications of μ-bases will often be written in computer algebra systems such as Maple.

- Both algorithms contain a step in which null vectors are computed (step 2 of HHK and step 4 of SG). Although Maple has a built-in routine for computing a basis of the null space for the input matrix, we did not use this built-in routine because we do not need the entire null basis, but only a certain subset of basis vectors with desired properties, consisting of $n-1$ vectors for HHK and a single vector for SG. For this reason, we implemented partial Gauss-Jordan elimination.

- For the rational field implementation of the SG algorithm, we produced the null vector in step 4 with integer entries in order to avoid rational number arithmetic (which is expensive due to integer gcd computations) in the subsequent steps of the algorithm.
- Dense representations of matrices were used for both algorithms. As shown in Remark 37, it is easy to exploit sparse representations for HHK, but it was not clear how one could exploit sparse representations for SG. Thus, in order to ensure fair comparison, we used dense representations for both algorithms.

3.1.7 Experiments, timing, and fitting

We now describe the experimental performance of both algorithms. We explain the setup for our experiments so that the timings reported here can be reproduced independently.

- The programs were executed using Maple 2015 version running on Apple iMac (Intel i 7-2600, 3.4 GHz, 16GB).
- The inputs were randomly-generated: for various values of d and n, the coefficients were taken randomly from $\mathbb{F}_5$, with a uniform distribution.
- In order to get reliable timings, especially when the computing time is small relative to the clock resolution, we ran each program several times on the same input and computed the average of the computing times.
- The execution of a program on an input was cut off if its computing time exceeded 120 seconds.

Figure 3.1 shows the experimental timing for the HHK algorithm, while Figure 3.2 shows the experimental timing for the SG algorithm. An experimental timing corresponds to a point (d, n, t), where d is the degree, n is the length of the input polynomial vector, and t is the time in seconds it took for our codes to produce the output. The algorithms were run on randomly-generated examples with specified d and n, and they ran in time t. For each figure, the axes represent the range of values $d = 3, \ldots, 200, \quad n = 3, \ldots, 200$, and $t = 0, \ldots, 120$, where t is the timing in seconds. Each dot (d, n, t) represents an experimental timing.

For each algorithm, we fit a surface through the experimental data points. The background gray surfaces are fitted to the experimental data. Our fitting models are based on the theoretical complexities obtained in Section 3.1.5. The fitting was computed using least squares. For HHK, based on Theorem 36, we chose a model for the timing, $t = \alpha_1 d^2 n + \alpha_2 d^3 + \alpha_3 n^2$, where α's are unknown constants to be determined. After substituting the experimental values (d, n, t), we obtain

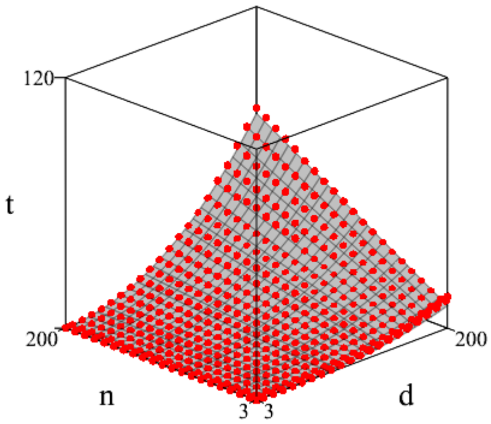

Figure 3.1 HHK algorithm timing

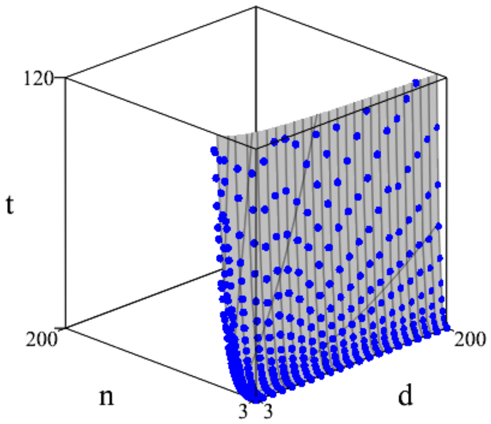

Figure 3.2 SG algorithm timing

an over-determined system of linear equations in the α's. We find α's that minimize the sum of squares of errors. For SG, we used the same procedure with the timing model $t = \beta_1 d n^5 + \beta_2 d^2 n^4$ based on Remark 38.

We generated the following functions:

$$t_{HHK} \approx 10^{-6} \cdot (7.4\, d^2 n + 1.2\, d^3 + 1.2\, n^2) \tag{3.24}$$

$$t_{SG} \approx 10^{-7} \cdot (2.6\, d n^5 + 0.6\, d^2 n^4) \tag{3.25}$$

For our experimental data, the residual standard deviation for the HHK-timing model (3.24) is 0.686 seconds, while the residual standard deviation for the SG-timing model (3.25) is 11.886 seconds.

We observe from Figures 3.1 and 3.2 that for a fixed d, the HHK algorithm's dependence on n is almost linear, while the SG algorithm's dependence on n is highly nonlinear. In fact, for the latter, the dependence is so steep that the algorithm was unable to terminate in under 120 seconds for most values of n, thus explaining the large amount of blank space in Figure 3.2. For a fixed n, the HHK algorithm's dependence on d is nonlinear, while the SG algorithm's dependence on d is almost linear.

3.1.8 Comparison

We now compare the performance of the two algorithms. Two pictures below represent performance comparisons.

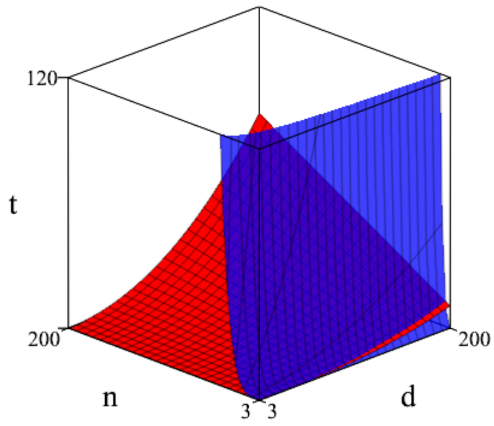

Figure 3.3 HHK (red) and SG (blue).

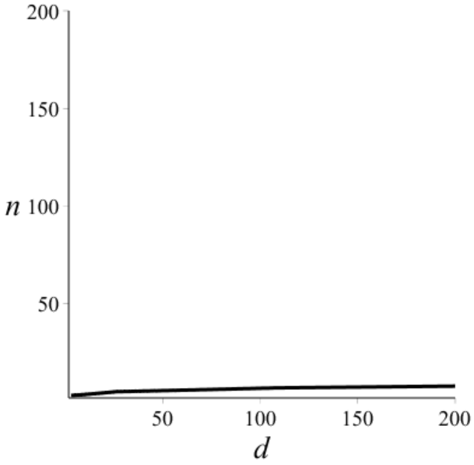

Figure 3.4 Tradeoff graph

- Figure 3.3 shows the fitted surfaces from Figures 3.1 and 3.2 on the same graph. The axes represent the range of values $n = 3, \ldots, 200$, $d = 3, \ldots, 200$, and $t = 0, \ldots, 120$, where t is the timing of the algorithms in seconds.
- Figure 3.4 shows a tradeoff graph for the two algorithms. The curve in the figure represents values of d and n for which the two algorithms run with the same timing. Below the curve, the SG algorithm runs faster, while above the curve, the HHK algorithm runs faster. The ratio of the dominant terms in the fitted formulae is $d : n^4$. This ratio manifests itself in the shape of the tradeoff curve presented in Figure 3.4.

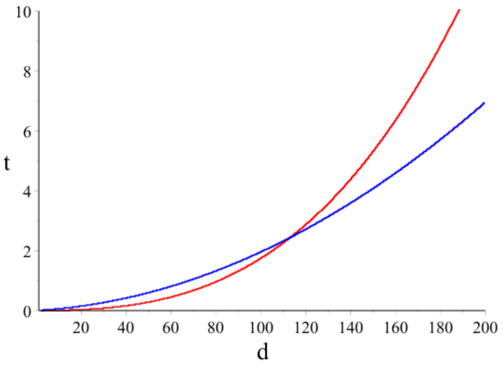

Figure 3.5 $n = 7$

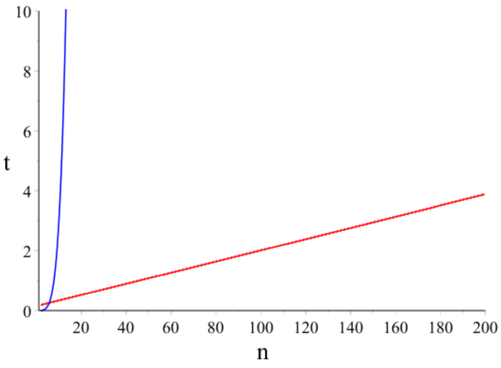

Figure 3.6 $d = 50$

From Figure 3.3, we observe that for a fixed d, as n increases the HHK algorithm vastly outperforms the SG algorithm. In contrast, for a fixed value of n, as d increases the SG algorithm outperforms the HHK algorithm. The order by which SG runs faster is less than the order by which HHK runs faster for fixed d and increasing n. We underscore this observation by displaying two-dimensional slices of Figure 3.3. Figure 3.5 represents the slice in the d-direction with $n = 7$, while Figure 3.6 represents the slice in the n-direction with $d = 50$. As before, HHK is represented by red and SG by blue.

3.1.9 Original definition, homogeneous version, and gcd

We now elaborate on a few additional topics related to μ-bases. Namely, the original definition of a μ-basis from [Cox98b], the homogeneous version of the μ-basis problem, and how μ-basis computations relate to gcd.

The original definition and proof of existence: The original definition of a μ-basis appeared on p 824 of a paper by Cox, Sederberg, and Chen [Cox98b] and is based on the "sum of the degrees" property (Statement 2 of Proposition 8). The definition also mentions an equivalent "reduced representation" (Statement 4 of Proposition 8). The proof of the existence theorem (Theorem 1 on p. 824 of [Cox98b]) appeals to the celebrated Hilbert Syzygy Theorem [Hil90] and utilizes Hilbert polynomials, which first appeared in the same paper [Hil90] under the name of characteristic functions. The definition of μ-basis in terms of the degrees, given in [Cox98b], is compatible with the tools that have been chosen to show its existence.

The homogeneous version of the problem: It is instructive to compare the inhomogeneous and homogenous versions of the problem. In fact, in order to invoke the Hilbert Syzygy Theorem in the proof of the existence of a μ-basis, Cox, Sederberg, and Chen restated the problem in the homogeneous setting (see pp. 824-825 of [Cox98b]).

Let $\hat{\mathbf{a}} = [a_1(\hat{x}, y), \dots, a_n(\hat{x}, y)]$ be a row vector of n homogeneous polynomials over a field $\mathbb{K}$, each of which has the same degree. As before, a syzygy of $\hat{\mathbf{a}}$ is a column vector $\mathbf{h} = [h_1(x, y), \dots, h_n(x, y)]^T$ of polynomials (not necessarily homogeneous), such that $\hat{\mathbf{a}}\,\mathbf{h} = 0$. The set syz($\hat{\mathbf{a}}$) is a module over $\mathbb{K}[x, y]$, and the Hilbert Syzygy Theorem implies that it is a free module of rank $n-1$ possessing a homogeneous basis. Let $n-1$ homogeneous polynomial vectors $\mathbf{u}_1(\hat{x}, y), \dots, \mathbf{u}_{\widehat{n-1}}(x, y)$ comprise an *arbitrary* homogeneous basis of syz($\hat{\mathbf{a}}$). Define dehomogenizations: $\mathbf{a}(s) = [a_1(s), \dots, a_n(s)]$, where $a_i(s) = \hat{a}_i(s, 1) \in \mathbb{K}[s]$, $i = 1, \dots, n$ and $\mathbf{u}_j(s) = \hat{\mathbf{u}}_j(s, 1) \in \mathbb{K}[s]^n$, $j = 1, \dots, n-1$. An argument, involving Hilbert polynomials on p. 825 of [Cox98b], shows that $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ is a *μ-basis* of syz($\mathbf{a}$).

Let us now start with a polynomial vector $\mathbf{a}(s) = [a_1(s), \dots, a_n(s)] \in \mathbb{K}[s]^n$ of degree d in the sense of Definition 6, and consider its homogenization $\hat{\mathbf{a}} = [\hat{a}_1(x, y), \dots, \hat{a}_n(x, y)]$, where $\hat{a}_i(x, y) =$

$y^d\, a_i\left(\frac{x}{y}\right)$, $i = 1,\dots,n$. It is not true that homogenezation of an arbitrary basis of syz($\mathbf{a}$) produces a basis of syz($\hat{\mathbf{a}}$). Indeed, let $\mathbf{u}_1$ and $\mathbf{u}_2$ be the columns of matrix M in Example 10. Then $\mathbf{u}_1 + \mathbf{u}_2$ and $\mathbf{u}_2$ is a basis of syz($\mathbf{a}$), with each vector having degree 3. Their homogenizations $\widehat{\mathbf{u}_1 + \mathbf{u}_2}$ and $\widehat{\mathbf{u}_2}$ are homogeneous polynomial vectors of degree 3, and, therefore, they can not possibly generate a homogeneous vector $\hat{\mathbf{u}}_1(x, y) = y\,\mathbf{u}_1\left(\frac{x}{y}\right) = [-x,\ y,\ x-y]^T$ of degree 1, which clearly belongs to syz($\hat{\mathbf{a}}$). A rather simple argument that utilizes the "reduced representation" property (Statement 4 of Proposition 8) can be used to show that for an arbitrary non-zero vector $\mathbf{a} \in \mathbb{K}[s]^n$, homogenization of any *μ-basis* of syz($\mathbf{a}$) produces a homogeneous basis of syz($\hat{\mathbf{a}}$).

The above discussion can be summarized in the following statement: the set of homogeneous bases of syz($\hat{\mathbf{a}}$) is in one-to-one correspondence with the set of μ-bases of syz($\mathbf{a}$), where $\mathbf{a} \in \mathbb{K}[s]^n$ is the dehomogenization of $\hat{\mathbf{a}} \in \mathbb{K}[x, y]^n$. Therefore, the algorithm developed in Section 3.1.4 can be used to compute homogeneous bases of syz($\hat{\mathbf{a}}$).

μ-basis algorithms and gcd *computation:* In contrast to the algorithm developed by Song and Goldman in [SG09], the algorithm presented in this section produces a μ-basis even when the input vector $\mathbf{a}$ has a non-trivial greatest common divisor. Moreover, once a μ-basis is computed, one can *immediately* find gcd($\mathbf{a}$) using Statement 5 of Proposition 8. Indeed, let $\mathbf{h}$ denote the outer product of a μ-basis $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$. If M is the matrix generated by the algorithm, then $h_i = (-1)^i\,|M_i|$, where M_i is an $(n-1)\times(n-1)$ submatrix of M constructed by removing the i-th row. By Statement 5 of Proposition 8, there exists a non-zero $\alpha \in \mathbb{K}$ such that

$$\mathbf{a} = \alpha\ \mathrm{gcd}(\mathbf{a})\,\mathbf{h}.$$

Let $i \in \{1,\dots,n\}$ be such that a_i is a non-zero polynomial. Then gcd($\mathbf{a}$) is computed by long division of a_i by h_i and then dividing the quotient by its leading coefficient to make it monic. In comparison, the algorithm developed in [SG09] produces a μ-basis of $\mathbf{a}$ multiplied by gcd($\mathbf{a}$). From the output of this algorithm and Statement 5 of Proposition 8, one finds gcd$(\mathbf{a})^{n-2}$. Song and Goldman discuss how to recover gcd($\mathbf{a}$) itself by *repeatedly* running their algorithm.

3.2 Matrix generalization: minimal bases

A natural generalization of the μ-basis problem is obtained by considering kernels, or nullspaces, of $m \times n$ polynomial matrices of rank m. A basis of the nullspace is called minimal if the "minimal degree" Statement 2 of Proposition 8 is satisfied (with $n-1$ replaced by $n-m$). One can easily adapt the argument in the proof of Theorem 2 in [SG09] to show that, in this more general setting, Statement 2 is equivalent to the "independence of the leading vectors" Statement 1 and to the "reduced representation" Statement 4 of Proposition 8. One can also show with an example that the "sum of the degrees" Statement 3 (with the degree of a polynomial matrix defined to be the maximum of the degrees of its entries) is no longer equivalent to Statements 1 and 4. There is a large body of work on computing minimal bases (see for instance [Bee87], [Ant05], [Zho12] and references therein). This research direction seems to be developing independently of the body of work devoted to μ-bases. The algorithm presented in Section 3.1.4 can be generalized to compute minimal bases of the kernels of $m \times n$ polynomial matrices. We now provide the details of this generalization.

Consider the matrix

$$\mathbf{A} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ \vdots & \vdots & \vdots \\ - & \mathbf{a}_m & - \end{bmatrix} \in \mathbb{K}[s]^{m\times n}$$

with rank m and degree at most d, where $m < n$. Here, we define the degree of a matrix to be the maximum of the degrees of its entries. Recall that in the case when $m = 1$, we look for syzygies of degree at most d by considering nullspace vectors of a matrix A in which the block of coefficients is repeated $d+1$ times. For arbitrary m, we look for kernel vectors of degree at most md (see [Zho12] and [SV05]). Thus, for each $\mathbf{a}_i$, we form its corresponding matrix A_i in which the block of coefficients for $\mathbf{a}_i$ is repeated $md+1$ times. That is, for each i, we have the matrix

$$A_i = \begin{bmatrix} c_0^i & & \\ \vdots & \ddots & \\ c_d^i & \vdots & c_0^i \\ & \ddots & \vdots \\ & & c_d^i \end{bmatrix} \in \mathbb{K}^{(md+d+1)\times n(md+1)}.$$

Then, to find the desired kernel vectors of $\mathbf{A}$, we look for vectors in $\mathbb{K}^{n(md+1)}$ that appear in the nullspace of each A_i, and then translate them into polynomials. In other words, we look for nullspace vectors of the stacked matrix

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} \in \mathbb{K}^{m(md+d+1)\times n(md+1)}.$$

All of the results in Section 3.1 can now be readily adapted to show that a minimal nullspace basis for $\mathbf{A}$ can be constructed by computing the nullspace vectors of A corresponding to its basic non-pivotal columns, and then translating these into polynomial vectors. The only difference is that there are now $n-m$ basic non-pivotal indices. We thus have the following algorithm.

Minimal Basis Algorithm

Input: $\mathbf{A} \neq 0 \in \mathbb{K}[s]^{m\times n}$ of rank m, where $m < n$ and $\mathbb{K}$ is a computable field

Output: $M \in \mathbb{K}[s]^{n\times(n-m)}$ such that its columns form a minimal nullspace basis of $\mathbf{A}$

1. *Construct stacked Sylvester-type matrix $A \in \mathbb{K}^{m(md+d+1)\times n(md+1)}$ whose null space corresponds to* $\ker_d(\mathbf{A})$*.*

   (a) $d \longleftarrow \deg(\mathbf{A})$

   (b) For each $i = 1, \ldots, m$, identify the row vectors $c_0^i, \ldots, c_d^i \in \mathbb{K}^n$ such that $\mathbf{a}_i = c_0^i + c_1^i s + \cdots + c_d^i s^d$.

   (c) $$A \longleftarrow \begin{bmatrix} c_0^1 & & \\ \vdots & \ddots & \\ c_d^1 & \vdots & c_0^1 \\ & \ddots & \vdots \\ & & c_d^1 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ c_0^m & & \\ \vdots & \ddots & \\ c_d^m & \vdots & c_0^m \\ & \ddots & \vdots \\ & & c_d^m \end{bmatrix}$$

2. *Construct the "partial" reduced row-echelon form E of A.*

   This can be done by using Gauss-Jordan elimination (forward elimination, backward elimination, and normalization), with the following optimizations:

   - Stop the forward elimination as soon as $n-m$ basic non-pivot columns are detected.
   - Skip over periodic non-pivot columns.
   - Carry out the row operations only on the required columns.

3. *Construct a matrix $M \in \mathbb{K}[s]^{n\times(n-m)}$ whose columns form a minimal nullspace basis of* $\mathbf{A}$.

   Let p be the list of the pivotal indices and let $\tilde{q}$ be the list of the basic non-pivotal indices of E.

```
(a) Initialize an $n \times n - m$ matrix $M$ with 0 in every entry.
(b) For $j = 1, \ldots, n - m$
        $r \leftarrow \operatorname{rem}(\tilde{q}_j - 1, n) + 1$
        $k \leftarrow \operatorname{quo}(\tilde{q}_j - 1, n)$
        $M_{r,j} \leftarrow M_{r,j} + s^k$
(c) For $i = 1, \ldots, |p|$
        $r \leftarrow \operatorname{rem}(p_i - 1, n) + 1$
        $k \leftarrow \operatorname{quo}(p_i - 1, n)$
        For $j = 1, \ldots, n - m$
            $M_{r,j} \leftarrow M_{r,j} - E_{i,\tilde{q}_j} s^k$
```

Example 39. *We trace the algorithm on the input matrix*

$$\mathbf{A} = \begin{bmatrix} s^3+s+1 & s^3+s^2+1 & s^3+1 & s^3+s^2+s+1 \\ -2\,s^3+s+1 & -2\,s^3+s^2+1 & -s^3+1 & s^3+1 \end{bmatrix} \in \mathbb{Q}^{2\times 4}$$

1. *Construct stacked Sylvester-type matrix $A \in \mathbb{K}^{m(md+d+1)\times n(md+1)}$ whose null space corresponds to* $\ker_d(\mathbf{A})$.

   (a) $d \longleftarrow 3$

   (b) $c_0^1, c_1^1, c_2^1, c_3^1 \longleftarrow \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$
   $c_0^2, c_1^2, c_2^2, c_3^2 \longleftarrow \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} -2 & -2 & -1 & 1 \end{bmatrix}$

   (c) $A \longleftarrow$

$$\left[\begin{array}{cccc|cccc|cccc|cccc|cccc|cccc|cccc}
1 & 1 & 1 & 1 & & & & & & & & & & & & & & & & & & & & & & & & \\
1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & & & & & & & & & & & & & & & & & & & & \\
0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & & & & & & & & & & & & & & & & \\
1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & & & & & & & & & & & & \\
 & & & & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & & & & & & & & \\
 & & & & & & & & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & & & & \\
 & & & & & & & & & & & & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
 & & & & & & & & & & & & & & & & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\
 & & & & & & & & & & & & & & & & & & & & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
 & & & & & & & & & & & & & & & & & & & & & & & & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & & & & & & & & & & & & & & & & & & & & & & & & \\
1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & & & & & & & & & & & & & & & & & & & & \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & & & & & & & & & & & & & & & & \\
-2 & -2 & -1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & & & & & & & & & & & & \\
 & & & & -2 & -2 & -1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & & & & & & & & \\
 & & & & & & & & -2 & -2 & -1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & & & & \\
 & & & & & & & & & & & & -2 & -2 & -1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
 & & & & & & & & & & & & & & & & -2 & -2 & -1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
 & & & & & & & & & & & & & & & & & & & & -2 & -2 & -1 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & & & & & & & & & & & & & & & & & & & -2 & -2 & -1 & 1
\end{array}\right]$$

2. *Construct the “partial" reduced row-echelon form E of A*

   $E \longleftarrow$

$$\left[\begin{array}{cccc|cccc|cccc|cccc|cccc|cccc|cccc}
\color{blue}1 & & & & & & & & & \color{red}-1 & & & & & & \color{red}-2 & & & & & & & & & & & & \\
 & \color{blue}1 & & & & & & & & \color{red}1 & & & & & & \color{red}1 & & & & & & & & & & & & \\
 & & \color{blue}1 & & & & & & & \color{red}0 & & & & & & \color{red}1 & & & & & & & & & & & & \\
 & & & \color{blue}1 & & & & & & \color{red}0 & & & & \color{brown}1 & & \color{red}0 & & & & & & & & & & & & \\
 & & & & \color{blue}1 & & & & & \color{red}-1 & & & & \color{brown}0 & & \color{red}-2 & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 & & & & & & & & \\
 & & & & & \color{blue}1 & & & & \color{red}-1 & & & & \color{brown}1 & & \color{red}-2 & \color{gray}1 & \color{gray}0 & \color{gray}0 & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 & & & & \\
 & & & & & & \color{blue}1 & & & \color{red}3 & & & & \color{brown}1 & & \color{red}6 & \color{gray}0 & \color{gray}1 & \color{gray}0 & \color{gray}1 & \color{gray}1 & \color{gray}0 & \color{gray}0 & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 \\
 & & & & & & & \color{blue}1 & & \color{red}0 & & & & & & \color{red}0 & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}0 & \color{gray}1 & \color{gray}0 & \color{gray}1 & \color{gray}1 & \color{gray}0 & \color{gray}0 & \color{gray}1 \\
 & & & & & & & & \color{blue}1 & \color{red}1 & & & & & & \color{red}1 & & & & & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}0 & \color{gray}1 & \color{gray}0 & \color{gray}1 \\
 & & & & & & & & & & \color{blue}1 & & & & & \color{red}-1 & & & & & & & & & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 \\
 & & & & & & & & & & & \color{blue}1 & & & & \color{red}1 & & & & & & & & & & & & \\
 & & & & & & & & & & & & \color{blue}1 & & & \color{red}-2 & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & \color{blue}1 & \color{red}3 & & & & & & & & & & & & \\
 & & & & & & & & & & & & & \color{brown}1 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & \color{brown}0 & & & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 & & & & & & & & \\
 & & & & & & & & & & & & & \color{brown}1 & & & \color{gray}1 & \color{gray}0 & \color{gray}0 & \color{gray}0 & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 & & & & \\
 & & & & & & & & & & & & & \color{brown}-2 & & & \color{gray}0 & \color{gray}1 & \color{gray}0 & \color{gray}0 & \color{gray}1 & \color{gray}0 & \color{gray}0 & \color{gray}0 & \color{gray}1 & \color{gray}1 & \color{gray}1 & \color{gray}1 \\
 & & & & & & & & & & & & & & & & \color{gray}-2 & \color{gray}-2 & \color{gray}-1 & \color{gray}1 & \color{gray}0 & \color{gray}1 & \color{gray}0 & \color{gray}0 & \color{gray}1 & \color{gray}0 & \color{gray}0 & \color{gray}0 \\
 & & & & & & & & & & & & & & & & & & & & \color{gray}-2 & \color{gray}-2 & \color{gray}-1 & \color{gray}1 & \color{gray}0 & \color{gray}1 & \color{gray}0 & \color{gray}0 \\
 & & & & & & & & & & & & & & & & & & & & & & & & \color{gray}-2 & \color{gray}-2 & \color{gray}-1 & \color{gray}1
\end{array}\right]$$

Here, blue denotes pivotal columns, red denotes basic non-pivotal columns, brown denotes periodic non-pivotal columns, and gray denotes unused columns.

3. *Construct matrix $M \in \mathbb{K}[s]^{n\times(n-m)}$ whose columns form a minimal nullspace basis of* $\mathbf{A}$.

   (a) $M \longleftarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

   (b) $M \longleftarrow \begin{bmatrix} 0 & 0 \\ s^2 & 0 \\ 0 & 0 \\ 0 & s^3 \end{bmatrix}$

   (c) $M \longleftarrow \begin{bmatrix} 1+s-s^2 & 2+2s-s^2+2s^3 \\ -1+s+s^2 & -1+2s \\ -3s & -1-6s+s^2-3s^3 \\ 0 & -s^2+s^3 \end{bmatrix}$

3.3 Relationship between μ-bases and Gröbner bases

As in previous sections, $\mathbb{K}$ is a field and $\mathbb{K}[s]$ is the ring of univariate polynomials over $\mathbb{K}$. This section discusses an interesting relationship between μ-bases and Gröbner bases of the syzygy module of a polynomial vector in $\mathbb{K}[s]^n$. For details on Gröbner bases see [Buc65; Cox98a] and for μ-bases see [Cox98b; SG09; Hon17]. We, however, point out that we allow *non-standard choices of monomials (see Definition 40)*. This will be crucial for characterizing the relationship between μ-bases and Gröbner bases.

3.3.1 Definitions

To define Gröbner bases of a submodule of $\mathbb{K}[s]^n$, one needs the notions of monomials and monomial ordering.

Definition 40 (Monomials). *Let $\mathbf{b}_1,\dots,\mathbf{b}_n$ be a basis of $\mathbb{K}^n$. Then the set of monomials consists of elements $s^k\mathbf{b}_j$ of $\mathbb{K}[s]^n$, where $k \geq 0$ and $1 \leq j \leq n$ are integers.*

Usually the standard basis $\mathbf{e}_1,\dots,\mathbf{e}_n$ of $\mathbb{K}^n$ is chosen, where $\mathbf{e}_i$ has 1 in the i-th position and zero everywhere else. However, *in this section, other bases of $\mathbb{K}^n$ also will be used*. Once we have fixed a basis of $\mathbb{K}^n$ (and in turn, the set of monomials), then we can impose an ordering on the monomials.

Definition 41 (Monomial ordering). *A monomial ordering is a total ordering $<$ on the set of monomials, such that for all pairs of monomials $\mathbf{m}_1$ and $\mathbf{m}_2$ and $k > 0$ the following holds*

$$\mathbf{m}_1 < \mathbf{m}_2 \Longrightarrow \mathbf{m}_1 < s^k\mathbf{m}_1 < s^k\mathbf{m}_2 \tag{3.26}$$

For $n > 1$, there is a variety of monomial orderings. To define an ordering, we start by ordering the chosen basis vectors $\mathbf{b}_1,\dots,\mathbf{b}_n$ of $\mathbb{K}^n$, by imposing $\mathbf{b}_{\pi_1} < \cdots < \mathbf{b}_{\pi_n}$ for a permutation π of $(1,\dots,n)$. Let B denote such an ordered basis. Then two of the most popular monomial orderings are *term over position ordering* defined as:

$$\mathrm{TOP}_B: \quad s^k\mathbf{b}_i < s^l\mathbf{b}_j \text{ if and only if } k < l \text{ or } (k = l \text{ and } \mathbf{b}_i < \mathbf{b}_j)$$

and *position over term ordering* defined as:

$$\mathrm{POT}_B: \quad s^k\mathbf{b}_i < s^l\mathbf{b}_j \text{ if and only if } \mathbf{b}_i < \mathbf{b}_j \text{ or } (i = j \text{ and } k < l).$$

Once we have fixed a monomial ordering, we can define the leading term/coefficient of any $\mathbf{a} \in \mathbb{K}[s]^n$ as follows.

Definition 42 (Leading term/coefficient)**.** *Let* $\mathbf{a} \in \mathbb{K}[s]^n$ *and* $<$ *be a monomial ordering on* $\mathbb{K}[s]^n$*. Then we can write* $\mathbf{a}$ *as a linear combination of monomials*

$$\mathbf{a} = \sum_{i=i}^{k} c_i \mathbf{m}_i$$

with $\mathbf{m}_1 < \mathbf{m}_2 < \mathbf{m}_3 < \cdots < \mathbf{m}_k$*. Then the leading term of* $\mathbf{a}$*, denoted* $LT(\mathbf{a})$*, is* $c_k \mathbf{m}_k$*. The coefficient* c_k *is called the leading coefficient of* $\mathbf{a}$*, denoted* $LC(\mathbf{a})$*.*

We are now ready to define a Gröbner basis of a submodule of $\mathbb{K}[s]^n$.

Definition 43 (Gröbner basis)**.** *Let* H *be a submodule of* $\mathbb{K}[s]^n$ *and* $<$ *a monomial ordering. Denote by* $\langle LT(H)\rangle$ *the submodule generated by the leading terms of all* $\mathbf{h} \in H$ *with respect to* $<$*. Then a finite set* $U = \{\mathbf{u}_1, \dots, \mathbf{u}_t\} \subset H$ *is a* $<$*-Gröbner basis of* H *if* $\langle LT(H)\rangle = \langle LT(\mathbf{u}_1), \dots, LT(\mathbf{u}_t)\rangle$*.*

Various criteria exist for testing whether a set U is a Gröbner basis. We would like to highlight one particular criterion. First, we need the following terminology.

Definition 44. *Fix a basis* $\mathbf{b}_1, \dots, \mathbf{b}_n$ *for* $\mathbb{K}^n$*.*

- *A monomial* $s^k \mathbf{b}_i$ *divides another monomial* $s^l \mathbf{b}_j$ *if and only if* $i = j$ *and* $k \leq l$*, in which case the quotient is* s^{l-k}*.*
- *The least common multiple of a pair of monomials* $s^k \mathbf{b}_i$ *and* $s^l \mathbf{b}_j$*, which we shall denote* $\mathrm{lcm}(s^k \mathbf{b}_i, s^l \mathbf{b}_j)$*, is defined by*

$$\mathrm{lcm}(s^k \mathbf{b}_i, s^l \mathbf{b}_j) = \begin{cases} s^{\max\{k,l\}} \mathbf{b}_i & \text{if} \quad i = j \\ \mathbf{0} & \text{if} \quad i \neq j \end{cases}$$

- *The* S*-vector for a pair of vectors* $\mathbf{u}_i$ *and* $\mathbf{u}_j$ *in* $\mathbb{K}[s]^n$*, which we shall denote* $S(\mathbf{u}_i, \mathbf{u}_j)$*, is defined by*

$$S(\mathbf{u}_i, \mathbf{u}_j) = \frac{\mathrm{lcm}(LT(\mathbf{u}_i), LT(\mathbf{u}_j))}{LT(\mathbf{u}_i)} \mathbf{u}_i - \frac{\mathrm{lcm}(LT(\mathbf{u}_i), LT(\mathbf{u}_j))}{LT(\mathbf{u}_j)} \mathbf{u}_j.$$

Definition 45. *Let* $F = (\mathbf{f}_1, \dots, \mathbf{f}_k)$ *be an ordered list, where* $\mathbf{f}_k \in \mathbb{K}[s]^n$*, and* $<$ *be a monomial ordering on* $\mathbb{K}[s]^n$ *Then the normal form of* $\mathbf{f} \in \mathbb{K}$ *with respect to* F *and* $<$*, denoted* $\overline{\mathbf{f}}^F$*, is the remainder of division of the vector* $\mathbf{f}$ *by the list of vectors in* F *(in the specified order).*

Lemma 46. *A set* $U = \{\mathbf{u}_1, \dots, \mathbf{u}_t\} \subset \mathbb{K}[s]^n$ *is a Gröbner basis of the submodule it generates if and only if* $\overline{S(\mathbf{u}_i, \mathbf{u}_j)}^{(\mathbf{u}_1, \dots, \mathbf{u}_t)} = 0$ *for all* $i \neq j$*.*

Lastly, we define the notions of minimal and reduced Gröbner bases.

Definition 47 (minimal Gröbner basis)**.** *A Gröbner basis U is minimal if:*

1. $LC(\mathbf{u}) = 1$ *for all* $\mathbf{u} \in U$.

2. $LT(\mathbf{u})$ *is not divisible by* $LT(\mathbf{u}')$ *for all* $\mathbf{u} \in U$ *and* $\mathbf{u}' \in U \backslash \{\mathbf{u}\}$.

Definition 48 (reduced Gröbner basis)**.** *A Gröbner basis U is reduced if:*

1. U *is minimal*

2. *No monomial in* $\mathbf{u}$ *is divisible by* $LT(\mathbf{u}')$ *for all* $\mathbf{u} \in U$ *and* $\mathbf{u}' \in U \backslash \{\mathbf{u}\}$.

For a given monomial order, every submodule has a unique reduced Gröbner basis.

3.3.2 Main result

We now state the main result on a relationship between μ-bases and Gröbner bases.

Theorem 49 (Main)**.** *Let* $\mathbf{a} \in \mathbb{K}[s]^n \backslash \{\mathbf{0}\}$ *and* $U \subset \mathbb{K}[s]^n$. *Then the following two statements are equivalent.*

(A) U *is a* μ*-basis of* $\operatorname{syz}(\mathbf{a})$.

(B) U *is a minimal* TOP_B*-Gröbner basis of* $\operatorname{syz}(\mathbf{a})$ *for some ordered basis* B *of* $\mathbb{K}^n$.

Each direction of the equivalence in the main theorem can be "informally" stated as follows.

- $(A) \Longrightarrow (B)$: "Every μ-basis is a minimal TOP Gröbner basis."
- $(B) \Longrightarrow (A)$: "Every minimal TOP Gröbner basis is a μ-basis."

In the following, we will prove the main theorem. We will prove the two directions one by one (Theorems 50 and 57). Furthermore, we will provide several related results/observations.

3.3.3 Every μ-basis is a minimal TOP Gröbner basis

Theorem 50. *Let* $\mathbf{a}\in\mathbb{K}[s]^n\setminus\{\mathbf{0}\}$ *and* $U\subset\mathbb{K}[s]^n$*. Let*

(A) U *is a* μ*-basis of* $\mathrm{syz}(\mathbf{a})$.

(B) U *is a minimal* TOP_B*-Gröbner basis of* $\mathrm{syz}(\mathbf{a})$ *for some ordered basis* B *of* $\mathbb{K}^n$*.*

Then $(A)\Longrightarrow(B)$.

Proof. Let $\mathbf{a}\in\mathbb{K}[s]^n\setminus\{\mathbf{0}\}$ and $U\subset\mathbb{K}[s]^n$. Assume (A), that is, U is a μ-basis of syz(a). We need to show (B), that is, U is a minimal TOP_B-Gröbner basis of $\mathrm{syz}(\mathbf{a})$ for some ordered basis B of $\mathbb{K}^n$.

Let $U=\{\mathbf{u}_1,\dots,\mathbf{u}_{n-1}\}$. An ordered basis B of $\mathbb{K}^n$ will be chosen in the following way. By definition of a μ-basis, the leading vectors $LV(\mathbf{u}_1),\dots,LV(\mathbf{u}_{n-1})$ are linearly independent over $\mathbb{K}$. We can, therefore, extend them to an ordered basis B of $\mathbb{K}^n$ by choosing a vector $\mathbf{b}_n\in\mathbb{K}^n$ that is not in the span of $\{LV(\mathbf{u}_1),\dots,LV(\mathbf{u}_{n-1})\}$, resulting in

$$B=(LV(\mathbf{u}_1),\dots,LV(\mathbf{u}_{n-1}),\mathbf{b}_n)$$

We can choose an arbitrary order for vectors in B and show that U is a minimal TOP_B-Gröbner basis of $\mathrm{syz}(\mathbf{a})$. By definition of μ-basis, $\mathbf{u}_1,\dots,\mathbf{u}_{n-1}$ generate $\mathrm{syz}(\mathbf{a})$. With respect to TOP_B, we have $LT(\mathbf{u}_i)=s^{\mu_i}LV(\mathbf{u}_i)$, where $\mu_i=\deg(\mathbf{u}_i)$, and, in particular, $LC(\mathbf{u}_i)=1$. Since $LV(\mathbf{u}_i)\neq LV(\mathbf{u}_j)$ for every pair $i\neq j$, we have $\mathrm{lcm}(LT(\mathbf{u}_i),LT(\mathbf{u}_j))=0$ for all $i\neq j$, and, therefore, $S(\mathbf{u}_i,\mathbf{u}_j)=0$ for all $i\neq j$. Then from Lemma 46 and Definition 47, it follows that U is a minimal TOP_B-Gröbner basis of $\mathrm{syz}(\mathbf{a})$. □

Example 51. *For* $\mathbf{a}=[1+s,1+s^2+s^3,1+s^4]\in\mathbb{Q}[s]^3$*, we computed a* μ*-basis:*

$$\mathbf{u}_1=\begin{bmatrix}-3s^2\\ s^2+s+2\\ -s-2\end{bmatrix},\quad \mathbf{u}_2=\begin{bmatrix}s^2-s+2\\ s^2-1\\ -s-1\end{bmatrix}.$$

by an implementation of the Song-Goldman algorithm [SG09]. We need to show that $\{\mathbf{u}_1,\mathbf{u}_2\}$ *is a minimal* TOP_B*-Gröbner basis of* $\mathrm{syz}(\mathbf{a})$ *for some ordered basis* B *of* $\mathbb{Q}^3$*.*

For this, we will chose an ordered basis of $\mathbb{Q}^3$*. Consider*

$$LV(\mathbf{u}_1)=\begin{bmatrix}-3\\ 1\\ 0\end{bmatrix}\ \textit{and}\ LV(\mathbf{u}_2)=\begin{bmatrix}1\\ 1\\ 0\end{bmatrix}.$$

We can extend these two vectors to a basis for $\mathbb{Q}^3$ by choosing

$$\mathbf{b}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

resulting in the following ordered basis of $\mathbb{Q}^3$:

$$B = \left(\begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right).$$

With respect to TOP_B*, we have $LT(\mathbf{u}_1) = s^2 LV(\mathbf{u}_1)$ and $LT(\mathbf{u}_2) = s^2 LV(\mathbf{u}_2)$. Since the basis vectors in these two monomials are distinct,* $\mathrm{lcm}(LT(\mathbf{u}_1), LT(\mathbf{u}_2)) = 0$*. Clearly, $\{\mathbf{u}_1, \mathbf{u}_2\}$ satisfies the conditions of a minimal* TOP_B*-Gröbner basis.*

Using the same example, however, we can show that *not* every μ-basis is a TOP_E-Gröbner basis, where E stands for the standard basis of $\mathbb{K}^n$ (or its permutations).

Example 52. *As in Example 51, we consider an input vector $\mathbf{a} = [1+s, 1+s^2+s^3, 1+s^4]$ and a μ-basis computed by an implementation of the Song-Goldman algorithm [SG09]:*

$$\mathbf{u}_1 = \begin{bmatrix} -3s^2 \\ s^2+s+2 \\ -s-2 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} s^2-s+2 \\ s^2-1 \\ -s-1 \end{bmatrix}.$$

We will show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is not a TOP Gröbner basis for all six possible choices of ordering of the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$:

1. *For* $\mathrm{TOP}_{\mathbf{e}_1<\mathbf{e}_2<\mathbf{e}_3}$*, $LT(\mathbf{u}_1) = [0, s^2, 0]^T$ and $LT(\mathbf{u}_2) = [0, s^2, 0]^T$. The S-vector of $\mathbf{u}_1$ and $\mathbf{u}_2$ is*

$$S(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{u}_1 - \mathbf{u}_2 = \begin{bmatrix} -4s^2+s-2 \\ s+3 \\ -1 \end{bmatrix}.$$

   With respect to the given ordering, the normal form of $S(\mathbf{u}_1, \mathbf{u}_2)$, relative to $\{\mathbf{u}_1, \mathbf{u}_2\}$ equals to $S(\mathbf{u}_1, \mathbf{u}_2)$, and, therefore, we conclude, using Lemma 46, that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is not a Gröbner basis.

2. *For* $\mathrm{TOP}_{\mathbf{e}_1<\mathbf{e}_3<\mathbf{e}_2}$ *and* $\mathrm{TOP}_{\mathbf{e}_3<\mathbf{e}_1<\mathbf{e}_2}$*, we again have $LT(\mathbf{u}_1) = [0, s^2, 0]^T$ and $LT(\mathbf{u}_2) = [0, s^2, 0]^T$*

and the same argument as in Case 1 shows that $\{\mathbf{u}_1, \mathbf{u}_2\}$ *is not a Gröbner basis with respect to either of these two orderings.*

3. *For* $\text{TOP}_{\mathbf{e}_2<\mathbf{e}_1<\mathbf{e}_3}$*, we have* $LT(\mathbf{u}_1) = [-3s^2, 0, 0]^T$ *and* $LT(u_2) = [s^2, 0, 0]^T$*. The S-vector of* $\mathbf{u}_1$ *and* $\mathbf{u}_2$ *is*
$$S(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{u}_1 + 3\,\mathbf{u}_2 = \begin{bmatrix} -3s+6 \\ 4s^2+s-1 \\ -4s-5 \end{bmatrix}$$
With respect to the given ordering, the normal form of $S(\mathbf{u}_1, \mathbf{u}_2)$*, relative to* $\{\mathbf{u}_1, \mathbf{u}_2\}$ *equals to* $S(\mathbf{u}_1, \mathbf{u}_2)$*, and, therefore, we conclude, using Lemma 46, that* $\{\mathbf{u}_1, \mathbf{u}_2\}$ *is not a Gröbner basis.*

4. *For* $\text{TOP}_{\mathbf{e}_2<\mathbf{e}_3<\mathbf{e}_1}$ *and* $\text{TOP}_{\mathbf{e}_3<\mathbf{e}_2<\mathbf{e}_1}$*, we again have* $LT(\mathbf{u}_1) = [-3s^2, 0, 0]^T$ *and* $LT(\mathbf{u}_2) = [s^2, 0, 0]^T$ *and the same argument as in Case 3, shows that* $\{\mathbf{u}_1, \mathbf{u}_2\}$ *is not a Gröbner basis with respect to either of these two orderings.*

In contrast, we can show that the μ-basis produced by the HHK algorithm, given in Section 3.1, is the reduced $\text{TOP}_{\mathbf{e}_1<\ldots<\mathbf{e}_n}$-Gröbner basis of syz($\mathbf{a}$).

Proposition 53. *For every* $\mathbf{a} \in \mathbb{K}[s]^n \setminus \{\mathbf{0}\}$*, the* μ*-basis produced by the HHK algorithm is the reduced* $\text{TOP}_{\mathbf{e}_1<\cdots<\mathbf{e}_n}$*-Gröbner basis of* syz($\mathbf{a}$).

Proof. Let $d = \deg(\mathbf{a})$. It is known that $\deg(\mathbf{u}_i) \le d$ for $i = 1, \ldots, n-1$. Relative to $\text{TOP}_{\mathbf{e}_1<\ldots<\mathbf{e}_n}$ ordering, the monomials up to degree d are ordered as follows:

$$\mathbf{e}_1 < \cdots < \mathbf{e}_n < s\,\mathbf{e}_1 < \cdots < s\,\mathbf{e}_n < \cdots < s^d\,\mathbf{e}_1 \cdots < s^d\,\mathbf{e}_n. \tag{3.27}$$

Theorem 27 of [Hon17] and the results preceding this theorem assert that, for each $i = 1, \ldots n-1$, the coefficients of monomials (3.27) of the vector $\mathbf{u}_i$ appear, in the same order, as the entries of the vector $b_{\tilde{q}_i}$ described on page 855 of [Hon17]. Indices $\tilde{q}_i$ are such that $1 \le \tilde{q}_1 < \cdots < \tilde{q}_{n-1} \le (d+1)n$. These indices are all distinct modulo n. For each i, the vector $b_{\tilde{q}_i}$ has 1 in the $\tilde{q}_i$-th component, and all components with larger indices are zero. Moreover, for each $j > i$, vector $b_{\tilde{q}_j}$ has zeros in all components with indices k, such that $k \ge \tilde{q}_i$ and $k = \tilde{q}_i \mod n$.

Let $\mu_i = \lceil r/n \rceil - 1$, where $\lceil\ \rceil$ denotes ceiling function, and let the integer r_i, between 1 and n, be such that $r_i = \tilde{q}_i \mod n$. Then $LT(\mathbf{u}_i) = s^{\mu_i}\,\mathbf{e}_{r_i}$. The monomials reducible by $LT(\mathbf{u}_i)$ must be equal to $s^t\,\mathbf{e}_{r_i}$, where $t \ge \mu_i$. However, for any $j \ne i$, it follows from the structure of the coefficient vectors $b_{\tilde{q}_j}$ that all such monomials in the polynomial vector $\mathbf{u}_j$ have zero coefficients. Thus U is the reduced $\text{TOP}_{\mathbf{e}_1<\ldots<\mathbf{e}_n}$-Gröbner basis of syz($\mathbf{a}$). □

Example 54. *Consider the input vector* $\mathbf{a} = \left[1+s^2+s^4,\, 1+s^3+s^4,\, 1+s^4\right] \in \mathbb{Q}[s]^3$. *The HHK algorithm produces a* μ*-basis for* $\mathbf{a}$ *consisting of:*

$$\mathbf{u}_1 = \begin{bmatrix} -s \\ 1 \\ -1+s \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1-2s-2s^2-s^3 \\ 2+2s+s^2+s^3 \\ -3 \end{bmatrix}.$$

With respect to $\mathrm{TOP}_{\mathbf{e}_1<\mathbf{e}_2<\mathbf{e}_3}$ *ordering, we have*

$$LT(\mathbf{u}_1) = \begin{bmatrix} 0 \\ 0 \\ s \end{bmatrix} = s\,\mathbf{e}_3 \text{ and } LT(\mathbf{u}_2) = \begin{bmatrix} 0 \\ s^3 \\ 0 \end{bmatrix} = s^3\,\mathbf{e}_2.$$

Since the basis vectors in these two monomials are distinct, $\mathrm{lcm}(LT(\mathbf{u}_1), LT(\mathbf{u}_2)) = 0$ *and so* $S(\mathbf{u}_1, \mathbf{u}_2) = 0$. *Then, from Lemma 46 and Definition 47, it follows that* $\{\mathbf{u}_1, \mathbf{u}_2\}$ *is a minimal Gröbner basis of* $\mathrm{syz}(\mathbf{a})$. *To show it is reduced, we expand out* $\mathbf{u}_1$ *and* $\mathbf{u}_2$ *into their monomials:*

$$\begin{aligned} \mathbf{u}_1 &= s\,\mathbf{e}_3 - s\,\mathbf{e}_1 - \mathbf{e}_3 + \mathbf{e}_2 \\ \mathbf{u}_2 &= s^3\,\mathbf{e}_2 - s^3\,\mathbf{e}_1 + s^2\mathbf{e}_2 - 2s^2\,\mathbf{e}_1 + 2s\,\mathbf{e}_2 - 2s\,\mathbf{e}_1 - 3\,\mathbf{e}_3 + 2\,\mathbf{e}_2 + \mathbf{e}_1. \end{aligned}$$

Notice that $LT(\mathbf{u}_1) = s\,\mathbf{e}_3$ *does not divide any monomial in* $\mathbf{u}_2$, *and* $LT(\mathbf{u}_2) = s^3\,\mathbf{e}_2$ *does not divide any monomial in* $\mathbf{u}_1$. *Thus,* $\{\mathbf{u}_1, \mathbf{u}_2\}$ *is reduced.*

Remark 55. Using permutations, it is easy to show that the HHK μ-basis algorithm can be used to produced a μ-basis for $\mathbf{a}$ that is the reduced TOP-Gröbner basis of syz($\mathbf{a}$) for any ordering of the standard basis vectors in $\mathbb{K}^n$.

Example 56. *Consider the same input vector* $\mathbf{a} = \left[1+s^2+s^4, 1+s^3+s^4, 1+s^4\right] \in \mathbb{Q}[s]^3$ *as in Example 54. We previously used the HHK algorithm to compute a* μ*-basis for* $\mathbf{a}$ *that is the reduced* $\mathrm{TOP}_{\mathbf{e}_1<\mathbf{e}_2<\mathbf{e}_3}$*-Gröbner basis of* $\mathrm{syz}(\mathbf{a})$. *We will now use the HHK algorithm to compute a* μ*-basis for* $\mathbf{a}$ *that is the reduced* $\mathrm{TOP}_{\mathbf{e}_3<\mathbf{e}_2<\mathbf{e}_1}$*-Gröbner basis of* $\mathrm{syz}(\mathbf{a})$. *Applying the permutation* $(3,2,1)$ *to* $\mathbf{a}$ *yields* $\mathbf{a}' = \left[1+s^4, 1+s^3+s^4, 1+s^2+s^4\right]$. *The HHK algorithm produces a* μ*-basis for* $\mathbf{a}'$ *consisting of:*

$$\mathbf{u}_1' = \begin{bmatrix} 1-s \\ -1 \\ s \end{bmatrix}, \quad \mathbf{u}_2' = \begin{bmatrix} -1-s^2-s^3 \\ s^3 \\ 1 \end{bmatrix}.$$

The inverse of the permutation $(3,2,1)$ *is* $(3,2,1)$. *Applying this inverse permatutation to* $\mathbf{u}_1'$ *and* $\mathbf{u}_2'$

yields a μ-basis for $\mathbf{a}$*:*

$$\mathbf{u}_1 = \begin{bmatrix} s \\ -1 \\ 1-s \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ s^3 \\ -1-s^2-s^3 \end{bmatrix}.$$

With respect to $\mathrm{TOP}_{\mathbf{e}_3<\mathbf{e}_2<\mathbf{e}_1}$ *ordering, we have*

$$LT(\mathbf{u}_1) = \begin{bmatrix} s \\ 0 \\ 0 \end{bmatrix} = s\mathbf{e}_1 \text{ and } LT(\mathbf{u}_2) = \begin{bmatrix} 0 \\ s^3 \\ 0 \end{bmatrix} = s^3 e_2.$$

Since the basis vectors in these two monomials are distinct, $\operatorname{lcm}(LT(\mathbf{u}_1), LT(\mathbf{u}_2)) = 0$ *and so* $S(\mathbf{u}_1, \mathbf{u}_2) = 0$*. Then, from Lemma 46 and Definition 47, it follows that* $\{\mathbf{u}_1, \mathbf{u}_2\}$ *is a minimal Gröbner basis of* $\operatorname{syz}(\mathbf{a})$*. To show it is reduced, we expand out* $\mathbf{u}_1$ *and* $\mathbf{u}_2$ *into their monomials:*

$$\begin{aligned} \mathbf{u}_1 &= s\mathbf{e}_1 - s\mathbf{e}_3 - \mathbf{e}_2 + \mathbf{e}_3 \\ \mathbf{u}_2 &= s^3\mathbf{e}_2 - s^3\mathbf{e}_3 - s^2\mathbf{e}_3 + \mathbf{e}_1 - \mathbf{e}_3 \end{aligned}$$

Notice that $LT(\mathbf{u}_1) = s\mathbf{e}_1$ *does not divide any monomial in* $\mathbf{u}_2$*, and* $LT(\mathbf{u}_2) = s^3\mathbf{e}_2$ *does not divide any monomial in* $\mathbf{u}_1$*. Thus,* $\{\mathbf{u}_1, \mathbf{u}_2\}$ *is reduced.*

3.3.4 Every minimal TOP Gröbner basis is a μ-basis

Theorem 57. *Let* $\mathbf{a} \in \mathbb{K}[s]^n \setminus \{\mathbf{0}\}$ *and* $U \subset \mathbb{K}[s]^n$*. Let*

(A) *U is a μ-basis of* $\operatorname{syz}(\mathbf{a})$.

(B) *U is a minimal* TOP_B*-Gröbner basis of* $\operatorname{syz}(\mathbf{a})$ *for some ordered basis B of* $\mathbb{K}^n$*.*

Then $(B) \Longrightarrow (A)$*.* [2]

Proof. Let $\mathbf{a} \in \mathbb{K}[s]^n \setminus \{\mathbf{0}\}$ and $U = \{\mathbf{u}_1, \ldots, \mathbf{u}_r\} \subset \mathbb{K}[s]^n$. Assume (B), that is, U is a minimal TOP_B-Gröbner basis of syz($\mathbf{a}$) for some ordered basis $B = (\mathbf{b}_1 < \cdots < \mathbf{b}_n)$ of $\mathbb{K}^n$. We need to show (A), that is, U is a μ-basis of syz(a). For this, we check the three conditions for μ-basis in Definition 9.

1. $\mathbf{u}_1, \ldots, \mathbf{u}_r$ generate syz($\mathbf{a}$).

   Immediate from the assumption that $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is a minimal TOP_B-Gröbner basis of syz($\mathbf{a}$)

[2] The theorem is equivalent, though looking quite different, to the following claim: Let $\mathbf{a} \in \mathbb{K}[s]^n$ and B be an ordered basis of $\mathbb{K}^n$. Then every TOP_B-Gröbner basis of syz($\mathbf{a}$) is a μ-basis of syz($\mathbf{a}$).

2. $LV(\mathbf{u}_1),\dots,LV(\mathbf{u}_r)$ are linearly independent over $\mathbb{K}$.

   We will prove it by contradiction, and thus suppose the leading vectors of U are linearly dependent over $\mathbb{K}$. Since U is a minimal TOP_B-Gröbner basis, we have that $LT(\mathbf{u}_k) = s^{d_k}\mathbf{b}_{i_k}$, where $d_k = \deg(\mathbf{u}_k)$ and all r indices i_k, $k = 1,\dots,r$ are distinct. We may assume that U is ordered so that $i_1 < \cdots < i_r$. Then

   $$LV(\mathbf{u}_k) = \mathbf{b}_{i_k} + \sum_{i<i_k} c_{k,i}\mathbf{b}_i, \tag{3.28}$$

   where $c_{k,i}$ is the coefficient of the monomial $s^{d_k}\mathbf{b}_i$ in $\mathbf{u}_k$. Then there exist constants $\alpha_i \in \mathbb{K}$, not all zero, such that

   $$\sum_{j=1}^{r} \alpha_j \, LV(\mathbf{u}_j) = 0. \tag{3.29}$$

   Let $l = \max\{j \mid \alpha_j \neq 0\}$, then (3.29), together with (3.28), imply that $\alpha_l \mathbf{b}_{i_l} = 0$. This is impossible because $\alpha_l \neq 0$ and $\mathbf{b}_{i_l} \neq 0$. Contradiction. Thus we conclude that the leading vectors of U are linearly independent over $\mathbb{K}$.

3. $r = n-1$.

   By Lemma 30, the elements of U are linearly independent over $\mathbb{K}[s]$. Thus U is a basis of the syz($\mathbf{a}$) module. Since syz($\mathbf{a}$) is a free module of rank $n-1$ it follows that U consists of $n-1$ elements.

□

It is worthwhile to point out that a minimal Gröbner basis of syz($\mathbf{a}$) with respect to other monomial orderings (e.g. POT) is not necessarily a μ-basis for $\mathbf{a}$, as the following example shows.

Example 58. *Consider the same vector* $\mathbf{a} = \left[1+s,\ 1+s^2+s^3, 1+s^4\right] \in \mathbb{Q}[s]^3$ *as in Example 51. The reduced (hence minimal)* $\mathrm{POT}_{\mathbf{e}_1<\mathbf{e}_2<\mathbf{e}_3}$*-Gröbner basis of* syz($\mathbf{a}$) *consists of*

$$\mathbf{u}_1 = \begin{bmatrix} -s^3-s^2-1 \\ s+1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -s^3+3s^2-s+1 \\ -2 \\ 1 \end{bmatrix}.$$

We observe that

$$LV(\mathbf{u}_1) = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \text{ and } LV(\mathbf{u}_2) = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}.$$

Clearly these two vectors are not linearly independent over $\mathbb{Q}$*, and therefore* $\{\mathbf{u}_1, \mathbf{u}_2\}$ *is not a* μ*-basis for* $\mathbf{a}$*.*

CHAPTER

4

ALGEBRAIC MOVING FRAMES: THEORY

This chapter examines the theory of algebraic moving frames and degree-optimal moving frames. In Section 4.1, we give precise definitions of an algebraic moving frame, degree-optimal moving frame, a minimal-degree Bézout vector, and a μ-basis. We show the relationships between these objects. In particular, Theorem 74 states that a minimal-degree Bézout vector and a μ-basis are the building blocks of any degree-optimal moving frame. This result, although essential to our theory, is by no means surprising and is easily deducible from known results. Theorem 75 and Proposition 76 establish important relationships between the degrees of a μ-basis and the degree of a minimal Bézout vector. In Section 4.2, by introducing a modified Sylvester-type matrix A, associated with an input vector $\mathbf{a}$, we reduce the problem of constructing a degree-optimal moving frame to a linear algebra problem over $\mathbb{K}$. Theorems 88 and 89 show how a minimal-degree Bézout vector and a μ-basis, respectively, can be constructed from the matrix A. In Section 4.3, we prove new results about the degree of an optimal moving frame. In particular, in Proposition 91, we establish the sharp lower bound $\left\lceil \frac{d}{n-1} \right\rceil$ and the sharp upper bound d for the degree of an optimal moving frame, and in Theorem 95, we prove that for a generic vector $\mathbf{a}$, the degree of every degree-optimal moving frame at $\mathbf{a}$ equals to the sharp lower bound.

4.1 Moving frames, Bézout vectors, and syzygies

In this section, we give the definitions of moving frame and degree-optimal moving frame, and explore the relationships between moving frames, syzygies, and Bézout vectors.

4.1.1 Basic definitions and notation

Throughout the chapter, $\mathbb{K}$ is an arbitrary field, $\overline{\mathbb{K}}$ is its algebraic closure, and $\mathbb{K}[s]$ is the ring of univariate polynomials over $\mathbb{K}$. For arbitrary natural numbers t and m, by $\mathbb{K}[s]^{t\times m}$ we denote the set of $t \times m$ matrices with polynomial entries. The set of $n \times n$ invertible matrices over $\mathbb{K}[s]$ is denoted as $GL_n(\mathbb{K}[s])$. It is well-known and easy to show that the determinant of such matrices is a nonzero element of $\mathbb{K}$. For a matrix N, we will use notation N_{*i} to denote its i-th column. For a square matrix, $|N|$ denotes its determinant.

By $\mathbb{K}[s]^m$ we denote the set of vectors of length m with polynomial entries. All vectors are implicitly assumed to be *column vectors*, unless specifically stated otherwise. Superscript ^{T} denotes transposition. We remind readers of the definitions of the degree and leading vector of a polynomial vector:

Definition 59 (Degree and Leading Vector)**.** *For* $\mathbf{h} = [h_1, \dots, h_m] \in \mathbb{K}[s]^m$ *we define the degree and the leading vector of* $\mathbf{h}$ *as follows:*

- $\deg(\mathbf{h}) = \max_{i=1,\dots,m} \deg(h_i)$.
- $LV(\mathbf{h}) = [\mathrm{coeff}(h_1, t), \dots, \mathrm{coeff}(h_m, t)]^T \in \mathbb{K}^n$*, where* $t = \deg(\mathbf{h})$ *and* $\mathrm{coeff}(h_i, t)$ *denotes the coefficient of* s^t *in* h_i*.*
- *We will say that a set of polynomial vectors* $\mathbf{h}_1, \dots, \mathbf{h}_k$ *is degree-ordered if* $\deg(\mathbf{h}_1) \le \dots \le \deg(\mathbf{h}_k)$

Example 60. *Let* $\mathbf{h} = \begin{bmatrix} 9-12s-s^2 \\ 8+15s \\ -7-5s+s^2 \end{bmatrix}$*. Then* $\deg(\mathbf{h}) = 2$ *and* $LV(\mathbf{h}) = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

By $\mathbb{K}[s]_t^m$ we denote the set of vectors of length m of degree at most t.

Throughout, $\mathbf{a} \in \mathbb{K}[s]^n$ *is assumed to be a nonzero row vector with* $n > 1$*.*

4.1.2 Algebraic moving frames and degree optimality

Definition 61 (Algebraic Moving Frame)**.** *For a given nonzero row vector* $\mathbf{a} \in \mathbb{K}[s]^n$*, with* $n > 1$*, an (algebraic) moving frame at* $\mathbf{a}$ *is a matrix* $P \in GL_n(\mathbb{K}[s])$*, such that*

$$\mathbf{a}\,P = [\gcd(\mathbf{a}), 0, \dots, 0], \tag{4.1}$$

where $\gcd(\mathbf{a})$ *denotes the greatest monic common devisor of* $\mathbf{a}$.

We clarify that by a zero polynomial we mean a polynomial with all its coefficients equal to zero (recall that when $\mathbb{K}$ is a finite field, there may exist a polynomial with nonzero coefficients, which nonetheless is a zero function on $\mathbb{K}$). As we will show below, a moving frame at $\mathbf{a}$ always exists and is not unique. For instance, if P is a moving frame at $\mathbf{a}$, then a matrix obtained from P by permuting the last $n-1$ columns of P is also a moving frame at $\mathbf{a}$. The set of all moving frames at $\mathbf{a}$ will be denoted $\mathrm{mf}(\mathbf{a})$. We are interested in constructing a moving frame of optimal degree.

Definition 62 (Degree-Optimal Algebraic Moving Frame)**.** *A moving frame* P *at* $\mathbf{a}$ *is called degree-optimal if*

1. $\deg(P_{*2}) \le \cdots \le \deg(P_{*n})$,
2. *if* P' *is another moving frame at* $\mathbf{a}$*, such that* $\deg(P'_{*2}) \le \cdots \le \deg(P'_{*n})$*, then*

$$\deg(P_{*i}) \le \deg(P'_{*i}) \quad \text{for} \quad i = 1, \dots, n.$$

In other words, we require that the last $n-1$ *columns of* P *(which are interchangeable) are degree-ordered, and that all columns of* P *are degree-optimal.*

For simplicity, we will often use the term *optimal moving frame* or *degree-optimal frame* instead of *degree-optimal algebraic moving frame*. A degree-optimal moving frame also is not unique, but it is clear from the definition that all optimal moving frames at $\mathbf{a}$ have the same column-wise degrees.

Example 63 (Running Example)**.** *We will show that* $P = \begin{bmatrix} 2-s & 3-3s-s^2 & 9-12s-s^2 \\ 1+2s & 2+5s+s^2 & 8+15s \\ -1-s & -2-2s & -7-5s+s^2 \end{bmatrix}$ *is a degree-optimal frame at* $\mathbf{a} = \begin{bmatrix} 2+s+s^4 & 3+s^2+s^4 & 6+2s^3+s^4 \end{bmatrix}$.

One can immediately notice that the moving frame is closely related to the Bézout identity and to syzygies of $\mathbf{a}$. We explore and exploit this relationship in the following subsections.

4.1.3 Bézout vectors

Definition 64 (Bézout Vector)**.** *A Bézout vector of a row vector* $\mathbf{a} \in \mathbb{K}[s]^n$ *is a column vector* $\mathbf{h} = [h_1, \ldots, h_n]^T \in \mathbb{K}[s]^n$*, such that*

$$\mathbf{a}\,\mathbf{h} = \gcd(\mathbf{a}).$$

The set of all Bézout vectors of $\mathbf{a}$ is denoted by Bez($\mathbf{a}$) and the set of Bézout vectors of degree at most d is denoted $\mathrm{Bez}_d(\mathbf{a})$.

Definition 65 (Minimal Bézout Vector)**.** *A Bézout vector* $\mathbf{h}$ *of* $\mathbf{a} = [a_1, \ldots, a_n] \in \mathbb{K}[s]^n$ *is said to be of minimal degree if*

$$\deg(\mathbf{h}) = \min_{\mathbf{h}' \in \mathrm{Bez}(\mathbf{a})} \deg(\mathbf{h}').$$

The existence of a Bézout vector can be proven using the extended Euclidean algorithm. Moreover, since the set of the degrees of all Bézout vectors is well-ordered, there is a minimal-degree Bézout vector. It is clear that the first column of a moving frame P at $\mathbf{a}$ is a Bézout vector of $\mathbf{a}$, and therefore, we will end up providing, in particular, a simple linear algebra algorithm to construct a Bézout vector of minimal degree.

4.1.4 Syzygies and μ-bases

Definition 66 (Syzygy)**.** *A syzygy of a nonzero row vector* $\mathbf{a} = [a_1, \ldots, a_n] \in \mathbb{K}[s]^n$*, for* $n > 1$*, is a column vector* $\mathbf{h} \in \mathbb{K}[s]^n$*, such that*

$$\mathbf{a}\,\mathbf{h} = 0.$$

The set of all syzygies of $\mathbf{a}$ is denoted by syz($\mathbf{a}$), and the set of syzygies of degree at most d is denoted $\mathrm{syz}_d(\mathbf{a})$. It is easy to see that syz($\mathbf{a}$) is a module. The next proposition shows that the last $n-1$ columns of a moving frame form a basis of syz($\mathbf{a}$).

Proposition 67 (Basis of Syzygies)**.** *Let* $P \in \mathrm{mf}(\mathbf{a})$*. Then the columns* $P_{*2}, \ldots, P_{*n}$ *form a basis of* syz($\mathbf{a}$)*.*

Proof. We need to show that $P_{*2}, \ldots, P_{*n}$ generate syz($\mathbf{a}$) and that they are linearly independent over $\mathbb{K}[s]$.

1. From (4.1), it follows that $\mathbf{a} P_{*2} = \cdots = \mathbf{a} P_{*n} = 0$. Therefore, $P_{*2}, \ldots, P_{*n} \in \mathrm{syz}(\mathbf{a})$. It remains to show that an arbitrary $\mathbf{h} \in \mathrm{syz}(\mathbf{a})$ can be expressed as a linear combination of $P_{*2}, \ldots, P_{*n} \in \mathrm{syz}(\mathbf{a})$ over $\mathbb{K}[s]$. Trivially we have

$$\mathbf{h} = P(P^{-1}\mathbf{h}). \tag{4.2}$$

From (4.1), it follows that $\mathbf{a} = \left[\begin{array}{cccc} \gcd(\mathbf{a}) & 0 & \cdots & 0 \end{array}\right] P^{-1}$ and, therefore, the first row of P^{-1} is the vector $\tilde{\mathbf{a}} = \mathbf{a}/\gcd(\mathbf{a})$.

Hence, since $\mathbf{a}\mathbf{h} = 0$, then $P^{-1}\mathbf{h} = [0, g_2(s), \dots, g_n(s)]^T$ for some $g_i(s) \in \mathbb{K}[s]$, $i = 2, \dots, n$. Then (4.2) implies:

$$\mathbf{h} = \sum_{i=2}^{n} g_i P_{*i}.$$

Thus $P_{*2}, \dots, P_{*n}$ generate $\operatorname{syz}(\mathbf{a})$.

2. Let $f_2, \dots, f_n \in \mathbb{K}[s]$ be such that

$$f_2\, P_{*2} + \cdots + f_n\, P_{*n} = 0. \tag{4.3}$$

Then $P\mathbf{f} = 0$, where $\mathbf{f} = [0, f_2, ..., f_n]^T$, and, since P is invertible, it follows that $f_2 = \cdots = f_n = 0$.

□

Remark 68. *Note that the proof of Proposition 67 is valid over the ring of polynomials in several variables. Thus, if a moving frame exists in the multivariable case, it follows that its last $n-1$ columns comprise a basis of* $\operatorname{syz}(\mathbf{a})$. *It is well-known that in the multivariable case there exists* $\mathbf{a}$ *for which* $\operatorname{syz}(\mathbf{a})$ *is not free and then, from Proposition 67, it immediately follows that a moving frame at* $\mathbf{a}$ *does not exist.*

In the univariate case, both the existence of an algebraic moving frames and freeness of the syzygy module are well-known. We do not, however, use these results, but as a by-product of developing an algorithm for constructing an optimal-degree moving frame, we produce a self-contained elementary linear algebra proof of their existence. We remind readers of the definition of a μ-basis.

Definition 69 (μ-basis)**.** *For a nonzero row vector* $\mathbf{a} \in \mathbb{K}[s]^n$, *a set of $n-1$ polynomial vectors* $\mathbf{u}_1, \dots, \mathbf{u}_{n-1} \in \mathbb{K}[s]^n$ *is called a μ-basis of* $\mathbf{a}$, *or, equivalently, a μ-basis of* $\operatorname{syz}(\mathbf{a})$, *if the following two properties hold:*

1. *$LV(\mathbf{u}_1), \dots, LV(\mathbf{u}_{n-1})$ are linearly independent over $\mathbb{K}$;*

2. $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ *generate* $\operatorname{syz}(\mathbf{a})$, *the syzygy module of* $\mathbf{a}$.

Recall that a μ-basis is, indeed, a *basis* of $\operatorname{syz}(\mathbf{a})$ as justified by Lemma 30.

In [Cox98b], Hilbert polynomials and the Hilbert Syzygy Theorem were used to show the existence of a basis of $\operatorname{syz}(\mathbf{a})$ with especially nice properties, called a μ-basis. An alternative proof of the existence of a μ-basis based on elementary linear algebra was given in Section 3.1.

In Propositions 70 below, we list and prove some properties of μ-bases, which are equivalent to its definition. These are adapted from Theorems 1 and 2 in [SG09]. For a more comprehensive list of properties of a μ-basis, see [SG09].

Proposition 70 (Equivalent properties)**.** *Let* $\mathbf{u}_1,\dots,\mathbf{u}_{n-1}$ *be a degree-ordered basis of* $\operatorname{syz}(\mathbf{a})$*, i.e.* $\deg(\mathbf{u}_1)\le\cdots\le\deg(\mathbf{u}_{n-1})$*. Then the following statements are equivalent:*

1. *[independence of the leading vectors]* $\mathbf{u}_1,\dots,\mathbf{u}_{n-1}$ *is a* μ*-basis.*

2. *[reduced representation] For every* $\mathbf{h}\in\operatorname{syz}(\mathbf{a})$*, there exist polynomials* $f_1,\dots,f_{n-1}$ *such that* $\deg(f_i\,\mathbf{u}_i)\le\deg(\mathbf{h})$ *and*
$$\mathbf{h}=\sum_{i=1}^{n-1} f_i\,\mathbf{u}_i. \tag{4.4}$$

3. *[optimality of the degrees] If* $\mathbf{h}_1,\dots,\mathbf{h}_{n-1}$ *is another basis of* $\operatorname{syz}(\mathbf{a})$*, such that* $\deg(\mathbf{h}_1)\le\cdots\le\deg(\mathbf{h}_{n-1})$*, then* $\deg(\mathbf{u}_i)\le\deg(\mathbf{h}_i)$ *for* $i=1,\dots,n-1$.

Proof.

(1) $\Longrightarrow$ (2) Since $\mathbf{u}_1,\dots,\mathbf{u}_{n-1}$ is a basis of syz($\mathbf{a}$), then for every $\mathbf{h}\in\operatorname{syz}(\mathbf{a})$ there exist polynomials $f_1,\dots,f_{n-1}$ such that (3.1) holds. Let $t=\max_{i=1,\dots,l}\big(\deg(f_i\,\mathbf{u}_i)\big)$ and let $\mathscr{I}$ be the set of indices on which this maximum is achieved. If $t>\deg(\mathbf{h})$, the equation (3.1) implies that $\sum_{i\in\mathscr{I}} LC(f_i)\,LV(\mathbf{u}_i)=0$, where $LC(f_i)$ is the leading coefficient of f_i and is nonzero for $i\in\mathscr{I}$. This identity contradicts our assumption that $LV(\mathbf{u}_1),\dots,LV(\mathbf{u}_{n-1})$ are linearly independent over $\mathbb{K}$. Thus $t\le\deg(\mathbf{h})$ as desired.

(2) $\Longrightarrow$ (3) Assume there exists a degree-ordered basis $\mathbf{h}_1,\dots,\mathbf{h}_{n-1}$ of syz($\mathbf{a}$) and an integer $k\in\{1,\dots,n-1\}$ such that $\deg(\mathbf{h}_k)<\deg(\mathbf{u}_k)$. Then there exists a matrix $H\in\mathbb{K}[s]^{(n-1)\times(n-1)}$, invertible over $\mathbb{K}[s]$, such that $[\mathbf{h}_1,\dots,\mathbf{h}_{n-1}]=[\mathbf{u}_1,\dots,\mathbf{u}_{n-1}]H$. However, from property (2) it follows that the upper right $k\times(n-k)$ block has only zero entries. This implies that $|H|=0$. Contradiction.

(3) $\Longrightarrow$ (1) Assume that $LV(\mathbf{u}_1)$, …, $LV(\mathbf{u}_{n-1})$ are dependent. Then there exist $\alpha_1,\dots,\alpha_{n-1}\in\mathbb{K}$, not all zero, such that
$$\alpha_1\,LV(\mathbf{u}_1)+\ldots+\alpha_{n-1}\,LV(\mathbf{u}_{n-1})=0. \tag{4.5}$$

Let k be the largest index such that α_k is non zero. Since $LV(\mathbf{u}_1)$ is a nonzero vector, $k>1$. Let

$d_i = \deg(\mathbf{u}_i)$ and consider a syzygy

$$\mathbf{h} = \alpha_k \, \mathbf{u}_k - \sum_{i=1}^{k-1} \alpha_i \, s^{d_k - d_i} \, \mathbf{u}_i.$$

Since $\mathbf{u}_k$ is a linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}, \mathbf{h}$, the set

$$\{\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_{n-1}, \mathbf{h}\} \tag{4.6}$$

also is a basis of syz($\mathbf{a}$). From (4.5) it follows that $\deg(\mathbf{h}) < \deg(\mathbf{u}_k)$. If $\deg(\mathbf{h}) < \deg(\mathbf{u}_1)$, then the generating set $\mathbf{h}, \mathbf{u}_1, \ldots, \mathbf{u}_{k-1}, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_{n-1}$ of syz($\mathbf{a}$) is degree-ordered. This contradicts our assumption that $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ is a degree-optimal basis. If $\deg(\mathbf{h}) \geq \deg(\mathbf{u}_1)$, let $i \in \{1, \ldots, k-1\}$ be maximal such that $\deg(\mathbf{h}) \geq \deg(\mathbf{u}_i)$. Then the set $\mathbf{u}_1, \ldots, \mathbf{u}_i, \mathbf{h}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_{k-1}, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_{n-1}$ is degree-ordered. Since $\deg(\mathbf{u}_{i+1}) > \deg(\mathbf{h})$ we again have a contradiction with our assumption that $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ is a degree-optimal basis.

□

We proceed with proving point-wise linear independence of the vectors in a μ-basis. In Theorem 1 of [SG09], μ-bases of real polynomial vectors were considered, and point-wise independence of the vectors in a μ-basis was proven for every s in $\mathbb{R}$. This proof can be word-by-word adapted to μ-bases of polynomial vectors over $\mathbb{K}$ to show point-wise independence of vectors in a μ-basis for every s in $\mathbb{K}$. To prove Theorem 74 below, however, we need a slightly stronger result: point-wise independence of the vectors in a μ-basis for every s in $\overline{\mathbb{K}}$. To arrive at this result, we first prove the following lemma. In this lemma and the following proposition, we use $\mathrm{syz}_{\mathbb{K}[s]}(\mathbf{a})$ to denote the syzygy module of $\mathbf{a}$ over the polynomial ring $\mathbb{K}[s]$, and $\mathrm{syz}_{\overline{\mathbb{K}}[s]}(\mathbf{a})$ to denote the syzygy module of $\mathbf{a}$ over the polynomial ring $\overline{\mathbb{K}}[s]$. Elsewhere, we use a shorter notation $\mathrm{syz}(\mathbf{a}) = \mathrm{syz}_{\mathbb{K}[s]}(\mathbf{a})$.

Lemma 71. *If* $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ *is a* μ*-basis of* $\mathrm{syz}_{\mathbb{K}[s]}(\mathbf{a})$*, then* $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ *is a* μ*-basis of* $\mathrm{syz}_{\overline{\mathbb{K}}[s]}(\mathbf{a})$.

Proof. Since $LV(\mathbf{u}_1), \ldots, LV(\mathbf{u}_{n-1})$ are independent over $\mathbb{K}$, they also are independent over $\overline{\mathbb{K}}$. Thus, it remains to show that $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ generate $\mathrm{syz}_{\overline{\mathbb{K}}[s]}(\mathbf{a})$. For an arbitrary $\mathbf{h} = [h_1, \ldots, h_n]^T \in \mathrm{syz}_{\overline{\mathbb{K}}[s]}(\mathbf{a})$, consider the field extension $\mathbb{H}$ of $\mathbb{K}$ generated by all the coefficients of the polynomials $h_1, \ldots, h_n$. Then $\mathbb{H}$ is a finite algebraic extension of $\mathbb{K}$ and, therefore, by one of the standard theorems of field theory (see, for example, the first two theorems in Section 41 of [Wae70]), $\mathbb{H}$ is a finite-dimensional vector space over $\mathbb{K}$. Let $\gamma_1, \ldots, \gamma_r \in \mathbb{H} \subset \overline{\mathbb{K}}$ be a vector space basis of $\mathbb{H}$ over $\mathbb{K}$. By expanding each of the coefficients in $\mathbf{h}$ in this basis, we can write $\mathbf{h}$ as

$$\mathbf{h} = \gamma_1 \mathbf{w}_1 + \cdots + \gamma_r \mathbf{w}_r, \tag{4.7}$$

for some $\mathbf{w}_1, \ldots, \mathbf{w}_r \in \mathbb{K}[s]^n$. Multiplying by $\mathbf{a}$ on the left, we get

$$0 = \gamma_1 \mathbf{a}\mathbf{w}_1 + \cdots + \gamma_r \mathbf{a}\mathbf{w}_r. \tag{4.8}$$

Assume there exists $i \in \{1, \ldots, r\}$ such that $\mathbf{a}\mathbf{w}_i \neq 0$. Let $k = \deg(\mathbf{a}\mathbf{w}_i)$ and let $b_j \in \mathbb{K}$ be the coefficient of the monomial s^k in the polynomial $\mathbf{a}\mathbf{w}_j$ for $j = 1, \ldots, r$. Then, from (4.8), we have

$$0 = \gamma_1 b_1 + \cdots + \gamma_r\, b_r.$$

Since $b_i \neq 0$, this contradicts the assumption that $\gamma_1, \ldots, \gamma_k$ is a vector space basis of $\mathbb{H}$ over $\mathbb{K}$. Thus, it must be the case that

$$\mathbf{a}\mathbf{w}_i = 0 \text{ for all } i = 1, \ldots, r$$

and, therefore, (4.7) implies that the module $\operatorname{syz}_{\overline{\mathbb{K}}[s]}(\mathbf{a})$ is generated by $\operatorname{syz}_{\mathbb{K}[s]}(\mathbf{a})$. Since $\operatorname{syz}_{\mathbb{K}[s]}(\mathbf{a})$ is generated by $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$, this completes the proof. □

Proposition 72 (Point-wise independence over $\overline{\mathbb{K}}$)**.** *If* $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ *is a μ-basis of* $\operatorname{syz}_{\mathbb{K}[s]}(\mathbf{a})$*, then for any value $s \in \overline{\mathbb{K}}$, the vectors* $\mathbf{u}_1(s), \ldots, \mathbf{u}_{n-1}(s)$ *are linearly independent over* $\overline{\mathbb{K}}$*.*

Proof. Suppose there exists $s_0 \in \overline{\mathbb{K}}$ such that $\mathbf{u}_1(s_0), \ldots, \mathbf{u}_{n-1}(s_0)$ are linearly dependent over $\overline{\mathbb{K}}$. Then there exist constants $\alpha_1, \ldots, \alpha_{n-1} \in \overline{\mathbb{K}}$, not all zero, such that

$$\alpha_1\, \mathbf{u}_1(s_0) + \cdots + \alpha_{n-1}\, \mathbf{u}_{n-1}(s_0) = 0.$$

Let $i = \max\{j \,|\, \alpha_j \neq 0\}$ and let

$$\mathbf{h} = \alpha_1\, \mathbf{u}_1 + \cdots + \alpha_i\, \mathbf{u}_i.$$

Then $\mathbf{h} \in \operatorname{syz}_{\overline{\mathbb{K}}[s]}(\mathbf{a})$ and is not identically zero, but $\mathbf{h}(s_0) = 0$. It follows that $\gcd(\mathbf{h}) \neq 1$ in $\overline{\mathbb{K}}[s]$ and, therefore, $\tilde{\mathbf{h}} = \frac{1}{\gcd(\mathbf{h})}\mathbf{h}$ belongs to $\operatorname{syz}_{\overline{\mathbb{K}}[s]}(\mathbf{a})$ and has degree strictly less than the degree of $\mathbf{h}$.
By Lemma 71, $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ is a μ-basis of $\operatorname{syz}_{\overline{\mathbb{K}}[s]}(\mathbf{a})$ and, since

$$\mathbf{u}_i = \frac{1}{\alpha_i}\left(\gcd(\mathbf{h})\tilde{\mathbf{h}} - \alpha_1\, \mathbf{u}_1 - \cdots - \alpha_{i-1}\, \mathbf{u}_{i-1}\right),$$

the set of syzygies

$$\{\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_{n-1}, \tilde{\mathbf{h}}\}$$

is a basis of $\operatorname{syz}_{\overline{\mathbb{K}}[s]}(\mathbf{a})$. Ordering it by degree and observing that $\deg(\tilde{\mathbf{h}}) < \deg(\mathbf{h}) = \deg(\mathbf{u}_i)$ leads to a contradiction with the degree optimality property of a μ-basis. □

4.1.5 The building blocks of a degree-optimal moving frame

From the discussions of the last subsection, it does not come as unexpected that a Bézout vector and a set of point-wise independent syzygies can serve as building blocks for a moving frame.

Proposition 73 (Building blocks of a moving frame)**.** *For a nonzero* $\mathbf{a} \in \mathbb{K}[s]^n$*, let* $\mathbf{h}_1, \dots, \mathbf{h}_{n-1}$ *be elements of* syz($\mathbf{a}$) *such that, for every* $s \in \overline{\mathbb{K}}$*, vectors* $\mathbf{h}_1(s), \dots, \mathbf{h}_{n-1}(s)$ *are linearly independent over* $\overline{\mathbb{K}}$*, and let* $\mathbf{h}_0$ *be a Bézout vector of* $\mathbf{a}$*. Then the matrix*

$$P = [\mathbf{h}_0, \mathbf{h}_1, \dots, \mathbf{h}_{n-1}]$$

is a moving frame at $\mathbf{a}$.

Proof. Clearly $\mathbf{a}P = [\gcd(\mathbf{a}), 0, \dots, 0]$. Let $\tilde{\mathbf{a}} = \frac{1}{\gcd(\mathbf{a})}\mathbf{a}$, then

$$\tilde{\mathbf{a}}P = [1, 0, \dots, 0]. \tag{4.9}$$

Assume that the determinant of P does not equal to a nonzero constant. Then there exists $s_0 \in \overline{\mathbb{K}}$ such that $|\mathbf{h}_0(s_0), \mathbf{h}_1(s_0), \dots, \mathbf{h}_{n-1}(s_0)| = 0$ and, therefore, there exist constants $\alpha_0, \dots, \alpha_n \in \overline{\mathbb{K}}$, not all zero, such that

$$\alpha_0 \mathbf{h}_0(s_0) + \alpha_1 \mathbf{h}_1(s_0) + \cdots + \alpha_{n-1} \mathbf{h}_{n-1}(s_0) = 0.$$

Multiplying on the left by $\tilde{\mathbf{a}}(s_0)$ and using (4.9), we get $\alpha_0 = 0$. Then

$$\alpha_1 \mathbf{h}_1(s_0) + \cdots + \alpha_{n-1} \mathbf{h}_{n-1}(s_0) = 0$$

for some set of constants $\alpha_1, \dots, \alpha_{n-1} \in \overline{\mathbb{K}}$, not all zero. But this contradicts our assumption that for every $s \in \overline{\mathbb{K}}$, vectors $\mathbf{h}_1(s), \dots, \mathbf{h}_{n-1}(s)$ are linearly independent over $\overline{\mathbb{K}}$. Thus, the determinant of P equals to a nonzero constant, and therefore P is a moving frame. □

Theorem 74. *A matrix* P *is a degree-optimal moving frame at* $\mathbf{a}$ *if and only if* P_{*1} *is a Bézout vector of* $\mathbf{a}$ *of minimal degree and* $P_{*2}, \dots, P_{*n}$ *is a* μ*-basis of* $\mathbf{a}$.

Proof.

($\Longrightarrow$) Let P be a degree-optimal moving frame at $\mathbf{a}$. From Definition 62, it immediately follows that P_{*1} is a Bézout vector of $\mathbf{a}$ of minimal degree. From Proposition 67, it follows that $P_{*2}, \dots, P_{*n}$ is a basis of syz($\mathbf{a}$). Assume $P_{*2}, \dots, P_{*n}$ is not a μ-basis of $\mathbf{a}$, and let $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ be a μ-basis. From Proposition 72, it follows that the vectors $\mathbf{u}_1(s), \dots, \mathbf{u}_{n-1}(s)$ are independent for all $s \in \overline{\mathbb{K}}$.

By Proposition 73, the matrix $P' = [P_{*1}, \mathbf{u}_1, \dots, \mathbf{u}_{n-1}]$ is a moving frame at **a**. On the other hand, since $P_{*2}, \dots, P_{*n}$ is not a μ-basis, then by Proposition 70, it is not a basis of optimal degree, and, therefore, there exists $k \in \{1, \dots, n-1\}$, such that $\deg(\mathbf{u}_k) < \deg(P_{*k+1})$. This contradicts our assumption that P is degree-optimal. Therefore, $P_{*2}, \dots, P_{*n}$ is a μ-basis.

($\Longleftarrow$) Assume P_{*1} is a Bézout vector of **a** of minimal degree and $P_{*2}, \dots, P_{*n}$ is a μ-basis of **a**. Then Proposition 72 implies that the vectors $P_{*2}(s), \dots, P_{*n}(s)$ are independent for all $s \in \overline{\mathbb{K}}$ and so P is a moving frame due to Proposition 73. Assume there exists a moving frame P' and an integer $k \in \{1, \dots, n\}$, such that $\deg(P'_{*k}) < \deg(P_{*k})$. If $k = 1$, then we have a contradiction with the assumption that P_{*1} is a Bézout vector of minimal degree. If $k > 1$, we have a contradiction with the degree optimality property of a μ-basis. Thus P satisfies Definition 62 of a degree-optimal moving frame.

□

Theorem 74 implies the following three-step process for constructing a degree-optimal moving frame at **a**.

1. Construct a Bézout vector **b** of **a** of minimal degree.
2. Construct a μ-basis $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ of **a**.
3. Let $P = [\mathbf{b}, \mathbf{u}_1, \dots, \mathbf{u}_{n-1}]$.

However, by exploiting the relationship between these building blocks, we develop, in Section 5.1, an algorithm that *simultaneously constructs a Bézout vector of minimal degree and a μ-basis*, avoiding redundancies embedded in the above three-step procedure.

4.1.6 The (β, μ)-type of a polynomial vector

The degree-optimality property of a μ-basis, stated in Proposition 70, insures that, although a μ-basis of **a** is not unique, the ordered list of the degrees of a μ-basis of **a** is unique. This list is called the μ-type of **a**. Thus the set of polynomial vectors can be split into classes according to their μ-type. An analysis of the μ-strata of the set of polynomial vectors is given by D'Andrea [D'A04], Cox and Iarrobino [CI15]. Similarly, although a minimal-degree Bézout vector for **a** is not unique, its degree is unique. If we denote this degree by β, we can refine the classification of polynomial vectors by studying their (β, μ)-strata. In this section, we explore the relationship between the μ-type and the β-type of a polynomial vector.

We start by showing that the degree of a minimal-degree Bézout vector of $\mathbf{a}$ is bounded by the maximal degree of a μ-basis of $\mathbf{a}$. This result is repeatedly used in this chapter.

Theorem 75. *For any nonzero* $\mathbf{a} \in \mathbb{K}[s]^n$*, and for any minimal-degree Bézout vector* $\mathbf{b}$ *and any* μ*-basis* $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ *of* $\mathbf{a}$*, we have*

1. *if* $\deg(\mathbf{a}) = \deg\big(\gcd(\mathbf{a})\big)$*, then* $\deg(\mathbf{b}) = 0$ *and* $\deg(\mathbf{u}_i) = 0$ *for* $i = 1, \ldots, n-1$.
2. *otherwise* $\deg(\mathbf{b}) < \max_j\{\deg(\mathbf{u}_j)\}$.

Proof.

1. The condition $\deg(\mathbf{a}) = \deg\big(\gcd(\mathbf{a})\big)$ implies that $\mathbf{a} = \gcd(\mathbf{a})\, v$, where v is a constant non-zero vector. In this case, it is obvious how to construct $\mathbf{b}$ and $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$, each with constant components.

2. In this case, $\deg(\mathbf{a}) > \deg\big(\gcd(\mathbf{a})\big)$. The coefficient of $\mathbf{ab}$ for $s^{\deg(\mathbf{a})+\deg(\mathbf{b})}$ is $LV(\mathbf{a})LV(\mathbf{b})$. By definition of Bézout vector, $\mathbf{ab} = \gcd(\mathbf{a})$. Therefore, by our assumption, $\deg(\mathbf{ab}) < \deg(\mathbf{a})$. Thus $LV(\mathbf{a})LV(\mathbf{b}) = 0$ or, in other words, $LV(\mathbf{b}) \in LV(\mathbf{a})^{\perp}$. Let $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ be a μ-basis of $\mathbf{a}$. By a similar argument, since $\mathbf{a}\mathbf{u}_j = 0$, we have $LV(\mathbf{u}_j) \in LV(\mathbf{a})^{\perp}$ for $j = 1, \ldots, n-1$. By definition of a μ-basis, $LV(\mathbf{u}_j)$ are linearly independent, and so they form a basis for $LV(\mathbf{a})^{\perp}$. Therefore, there exist constants $\alpha_1, \ldots, \alpha_{n-1}$ such that $LV(\mathbf{b}) = \sum_{j=1}^{n-1} \alpha_j LV(\mathbf{u}_j)$. Suppose that $\deg(\mathbf{b}) \geq \max_j\{\deg(\mathbf{u}_j)\}$. Define $\tilde{\mathbf{b}} = \mathbf{b} - \sum_{j=1}^{n-1} \alpha_j \mathbf{u}_j s^{\deg(\mathbf{b})-\deg(\mathbf{u}_j)}$. Then $\mathbf{a}\tilde{\mathbf{b}} = \gcd(\mathbf{a})$ and $\deg(\tilde{\mathbf{b}}) < \deg(\mathbf{b})$, a contradiction to the minimality of $\deg(\mathbf{b})$. Therefore, $\deg(\mathbf{b}) < \max_j\{\deg(\mathbf{u}_j)\}$.

□

In the next proposition, we show that, except for the upper bound provided by $\mu_{n-1} - 1$, no other additional restrictions on the degree of the minimal Bézout vector are imposed by the μ-type, and therefore the β-type provides an essentially new characteristic of a polynomial vector.

Proposition 76. *Fix* $n \geq 2$*. For all ordered lists of nonnegative integers* $\mu_1 \leq \cdots \leq \mu_{n-1}$*, with* $\mu_{n-1} \neq 0$*, and for all* $j \in \big\{0, \ldots, \mu_{n-1} - 1\big\}$*, there exists* $\mathbf{a} \in \mathbb{K}[s]^n$ *such that* $\gcd(\mathbf{a}) = 1$ *and*

1. *for any* μ*-basis* $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ *of* $\mathbf{a}$*, we have* $\deg(\mathbf{u}_i) = \mu_i$*,* $i = 1, \ldots, n-1$.
2. *for any minimal-degree Bézout vector* $\mathbf{b}$ *of* $\mathbf{a}$*, we have* $\deg(\mathbf{b}) = j$.

Proof. In the case when $n = 2$, given a non-negative integer μ_1 and an integer $j \in \{0, \dots, \mu_1 - 1\}$, take $\mathbf{a} = \left[s^{\mu_1 - j}, s^{\mu_1} + 1\right]$. Then, obviously $\gcd(\mathbf{a}) = 1$, vector $\mathbf{b} = [-s^j, 1]^T$ is a minimal-degree Bézout vector, and vector $\mathbf{u}_1 = \left[s^{\mu_1} + 1, -s^{\mu_1 - j}\right]^T$ is the minimal-degree syzygy, which in this case comprises a μ-basis of $\mathbf{a}$. Thus $\mathbf{a}$ has the required properties.

In the case when $n \geq 3$, for the set of integers $\mu_1, \dots, \mu_{n-1}, j$ described in the proposition, take

$$\mathbf{a} = \left[s^{\mu_{n-1} - j}, s^{\mu_{n-1} - j + \mu_1}, s^{\mu_{n-1} - j + \mu_1 + \mu_2}, \dots, s^{\mu_{n-1} - j + \mu_1 + \cdots + \mu_{n-2}}, s^{\mu_1 + \cdots + \mu_{n-1}} + 1\right].$$

Observe that $\gcd(\mathbf{a}) = 1$, and consider the matrix

$$P = \begin{bmatrix} & s^{\mu_1} & & & 1 \\ & -1 & s^{\mu_2} & & \\ & & -1 & \ddots & \\ -s^j & & & \ddots & s^{\mu_{n-1}} \\ 1 & & & & -s^{\mu_{n-1} - j} \end{bmatrix}.$$

It is easy to see that $\mathbf{a}P = [1, 0, \dots, 0]$ and $|P| = \pm 1$, so P is a moving frame at $\mathbf{a}$ according to Definition 61. Therefore, the first column of P, i.e vector $\mathbf{b} = P_{*1}$, is a Bézout vector of $\mathbf{a}$, while the remaining columns $\mathbf{u}_1 = P_{*2}, \dots, \mathbf{u}_{n-1} = P_{*n}$ comprise a basis for the syzygy module of $\mathbf{a}$ according to Proposition 67. Clearly $\deg(\mathbf{b}) = j$, while $\deg \mathbf{u}_i = \mu_i$ for $i = 1, \dots, n-1$.

The leading vectors of $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ are linearly independent and, therefore, vectors $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ comprise a μ-basis of $\mathbf{a}$. To prove that $\mathbf{b}$ is of minimal degree, suppose, for the sake of contradiction, that there exists a vector $\mathbf{f} = \left[f_1, \dots, f_n\right]^T \in \mathbb{K}[s]^n$ with $\deg(\mathbf{f}) < j$ such that

$$f_1(s)\, a_1(s) + \ldots + f_n(s)\, a_n(s) = 1 \text{ for all } s. \tag{4.10}$$

We observe that, since $\mu_{n-1} > 0$ and $j < \mu_{n-1}$, then $a_i(0) = 0$ for $i = 1, \dots, n-1$ and $a_n(0) = 1$. Then, by substituting $s = 0$ in (4.10), we get $f_n(0) = 1$ and, therefore, $f_n(s)$ is not a zero polynomial. This implies that $\deg(f_n a_n) = \mu_1 + \cdots + \mu_{n-1} + \deg(f_n)$. Therefore, in order for the Bézout identity (4.10) to hold, at least one of the remaining $f_i a_i$, $i = 1, \dots, n-1$, must contain a monomial of degree $\mu_1 + \cdots + \mu_{n-1} + \deg(f_n)$ as well. However, we assumed that $\deg(f_i) < j$ for all i, which implies that $\deg(f_i a_i) < \mu_1 + \cdots + \mu_{n-1}$ for $i = 1, \dots, n-1$. Contradiction. We thus conclude that $\mathbf{a}$ has the required properties. □

4.2 Reduction to a linear algebra problem over $\mathbb{K}$

In this section, we show that for a vector $\mathbf{a} \in \mathbb{K}[s]_d^n$ such that $\gcd(\mathbf{a}) = 1$, a Bézout vector of $\mathbf{a}$ of minimal degree and a μ-basis of $\mathbf{a}$ can be obtained from linear relationships among certain columns of a $(2d+1) \times (nd+n+1)$ matrix over $\mathbb{K}$. Since essentially the same matrix has been used to construct a μ-basis in Section 3.1, we later use this result to develop a degree-optimal moving frame algorithm that simultaneously constructs a μ-basis and a minimal-degree Bézout vector.

4.2.1 Sylvester-type matrix A and its properties

For a nonzero polynomial row vector

$$\mathbf{a} = \sum_{0 \le i \le d} [c_{i1}, \dots, c_{in}] s^i \tag{4.11}$$

of length n and degree d, we correspond a $\mathbb{K}^{(2d+1)\times n(d+1)}$ matrix

$$A = \begin{bmatrix}
c_{01} & \cdots & c_{0n} & & & & & & & \\
\vdots & \cdots & \vdots & c_{01} & \cdots & c_{0n} & & & & \\
\vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \ddots & & & \\
c_{d1} & \cdots & c_{dn} & \vdots & \cdots & \vdots & \ddots & c_{01} & \cdots & c_{0n} \\
 & & & c_{d1} & \cdots & c_{dn} & \ddots & \vdots & \cdots & \vdots \\
 & & & & & & \ddots & \vdots & \cdots & \vdots \\
 & & & & & & & c_{d1} & \cdots & c_{dn}
\end{bmatrix} \tag{4.12}$$

with the blank spaces filled by zeros. In other words, matrix A is obtained by taking $d+1$ copies of a $(d+1) \times n$ block of the coefficients of polynomials in $\mathbf{a}$. The blocks are repeated horizontally from left to right, and each block is shifted down by one relative to the previous one. Matrix A is related to the *generalized resultant matrix* R, appearing on page 333 of [VS78]. Indeed, if one takes the top-left $\mathbb{K}^{2d\times nd}$ submatrix of A, transposes this submatrix, and then permutes certain rows, one obtains R. However, the size and shape of the matrix A turns out to be crucial to our construction.

Example 77. *For the row vector* $\mathbf{a}$ *in the running example (Example 63), we have* $n = 3$, $d = 4$,

$$c_0 = [2,3,6],\ c_1 = [1,0,0],\ c_2 = [0,1,0],\ c_3 = [0,0,2],\ c_4 = [1,1,1]$$

and

$$A = \left[\begin{array}{ccccccccccccccc}
2 & 3 & 6 & & & & & & & & & & & & \\
1 & 0 & 0 & 2 & 3 & 6 & & & & & & & & & \\
0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & & & & \\
0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & \\
1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 \\
 & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 0 & 2 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right].$$

A visual periodicity of the matrix A is reflected in the periodicity property of its non-pivotal columns which we are going to precisely define and exploit below. We remind readers the of the definition of pivotal and non-pivotal columns.

Definition 78. *A column of any matrix N is called pivotal if it is either the first column and is nonzero or it is linearly independent of all previous columns. The rest of the columns of N are called non-pivotal. The index of a pivotal (non-pivotal) column is called a pivotal (non-pivotal) index.*

From this definition, it follows that every non-pivotal column can be written as a linear combination of the preceding *pivotal columns*.

We denote the set of pivotal indices of A as p and the set of its non-pivotal indices as q. We remind readers of how the specific structure of the matrix A is reflected in the structure of the set of non-pivotal indices q.

Lemma 79 (Periodicity)**.** *If $j \in q$ then $j + kn \in q$ for $0 \le k \le \left\lfloor \frac{n(d+1)-j}{n} \right\rfloor$. Moreover,*

$$A_{*j} = \sum_{r<j} \alpha_r A_{*r} \quad \Longrightarrow \quad A_{*j+kn} = \sum_{r<j} \alpha_r A_{*r+kn}, \tag{4.13}$$

where A_{*j} denotes the j-th column of A.

Definition 80. *Let q be the set of non-pivotal indices. Let $q/(n)$ denote the set of equivalence classes of q modulo n. Then the set $\tilde{q} = \{\min \varrho \,|\, \varrho \in q/(n)\}$ will be called the set of basic non-pivotal indices. The remaining indices in q will be called periodic non-pivotal indices.*

Example 81. *For the matrix A in Example 77, we have $n = 3$ and $q = \{8, 9, 11, 12, 14, 15\}$. Then $q/(n) = \big\{\{8, 11, 14\}, \{9, 12, 15\}\}\big\}$ and $\tilde{q} = \{8, 9\}$.*

4.2.2 Isomorphism between $\mathbb{K}[s]_t^m$ and $\mathbb{K}^{m(t+1)}$

The second ingredient that we use to reduce our problem to a linear algebra problem over $\mathbb{K}$ is an explicit isomorphism between vector spaces $\mathbb{K}[s]_t^m$ and $\mathbb{K}^{m(t+1)}$. This is the same isomorphism used in Section 3.1, and we reprint it here for convenience. Any polynomial m-vector $\mathbf{h}$ of degree at most t can be written as $\mathbf{h} = w_0 + s\,w_1 + \cdots + s^t\,w_t$ where $w_i = [w_{1i}, \ldots, w_{mi}]^T \in \mathbb{K}^m$. It is clear that the map

$$\sharp_t^m : \mathbb{K}[s]_t^m \to \mathbb{K}^{m(t+1)}$$

$$\mathbf{h} \to \mathbf{h}^{\sharp_t^m} = \begin{bmatrix} w_0 \\ \vdots \\ w_t \end{bmatrix} \tag{4.14}$$

is linear. It is easy to check that the inverse of this map

$$\flat_t^m : \mathbb{K}^{m(t+1)} \to \mathbb{K}[s]_t^m$$

is given by a linear map:

$$v \to v^{\flat_t^m} = S_t^m\, v \tag{4.15}$$

where

$$S_t^m = \begin{bmatrix} I_m & s\,I_m & \cdots s^t\,I_m \end{bmatrix} \in \mathbb{K}[s]^{m\times m(t+1)}.$$

Here I_m denotes the $m \times m$ identity matrix. For the sake of notational simplicity, we will often write $\sharp$, $\flat$ and S instead of $\sharp_t^m$, $\flat_t^m$ and S_t^m when the values of m and t are clear from the context.

Example 82. *For* $\mathbf{h} \in \mathbb{Q}_3^3[s]$ *given by*

$$\mathbf{h} = \begin{bmatrix} 9-12s-s^2 \\ 8+15s \\ -7-5s+s^2 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \\ -7 \end{bmatrix} + s \begin{bmatrix} -12 \\ 15 \\ -5 \end{bmatrix} + s^2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},$$

we have

$$\mathbf{h}^\sharp = [9, 8, -7, -12, 15, -5, -1, 0, 1]^T.$$

Note that

$$\mathbf{h} = (\mathbf{h}^\sharp)^\flat = S\,\mathbf{h}^\sharp = \begin{bmatrix} I_3 & s I_3 & s^2 I_3 \end{bmatrix} \mathbf{h}^\sharp.$$

With respect to the isomorphisms $\sharp$ and $\flat$, the $\mathbb{K}$-linear map $\mathbf{a}\colon \mathbb{K}[s]_d^n \to \mathbb{K}[s]_{2d}$ corresponds to the $\mathbb{K}$ linear map $A\colon \mathbb{K}^{n(d+1)} \to \mathbb{K}^{2d+1}$ in the following sense:

Lemma 83. *Let* $\mathbf{a} = \sum_{0 \le j \le d} c_j s^j \in \mathbb{K}_d^n[s]$ *and* $A \in \mathbb{K}^{(2d+1)\times n(d+1)}$ *defined as in* (4.12)*. Then for any* $v \in \mathbb{K}^{n(d+1)}$ *and any* $\mathbf{h} \in \mathbb{K}[s]_d^n$*:*

$$\mathbf{a}\, v^{\flat_d^n} = (Av)^{\flat_{2d}^1} \text{ and } (\mathbf{a}\mathbf{h})^{\sharp_{2d}^1} = A\mathbf{h}^{\sharp_d^n}. \tag{4.16}$$

The proof of Lemma 83 is straightforward. The proof of the first equality is explicitly spelled out in Section 3.1 (see Lemma 14). The second equality follows from the first and the fact that $\flat_t^m$ and $\sharp_t^m$ are mutually inverse maps.

Example 84. *Consider the row vector* $\mathbf{a}$ *in the running example (Example 63) and its associated matrix* A *(Example 77). Let* $v = [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15]^T$*. Then*

$$Av = [26, 60, 98, 143, 194, 57, 62, 63, 42]^T$$

and so

$$(Av)^{\flat_{2d}^1} = S_8^1(Av) = 26 + 60s + 98s^2 + 143s^3 + 194s^4 + 57s^5 + 62s^6 + 63s^7 + 42s^8.$$

On the other hand, since

$$v^{\flat_d^n} = S_4^3 v = \begin{bmatrix} 1 + 4s + 7s^2 + 10s^3 + 13s^4 \\ 2 + 5s + 8s^2 + 11s^3 + 14s^4 \\ 3 + 6s + 9s^2 + 12s^3 + 15s^4 \end{bmatrix},$$

we have

$$\begin{aligned} \mathbf{a}v^{\flat_d^n} &= \begin{bmatrix} 2 + s + s^4 & 3 + s^2 + s^4 & 6 + 2s^3 + s^4 \end{bmatrix} \begin{bmatrix} 1 + 4s + 7s^2 + 10s^3 + 13s^4 \\ 2 + 5s + 8s^2 + 11s^3 + 14s^4 \\ 3 + 6s + 9s^2 + 12s^3 + 15s^4 \end{bmatrix} \\ &= 42s^8 + 63s^7 + 62s^6 + 57s^5 + 194s^4 + 143s^3 + 98s^2 + 60s + 26. \end{aligned}$$

We observe that

$$\mathbf{a}v^{\flat_d^n} = (Av)^{\flat_{2d}^1}.$$

We proceed by showing that if $\gcd(\mathbf{a}) = 1$, then the matrix A has full rank. This statement can be deduced from the results about the rank of a different Sylvester-type matrix, R, given in Section 2 of [VS78]. We, however, give a short independent proof using the following lemma, which also will be used in other parts of the chapter.

Lemma 85. *For all* $\mathbf{a} \in \mathbb{K}[s]^n$ *with* $\gcd(\mathbf{a}) = 1$ *and* $\deg(\mathbf{a}) = d$ *and all* $i = 0, \dots, 2d$*, there exist vectors* $\mathbf{h}_i \in \mathbb{K}[s]^n$ *such that* $\deg(\mathbf{h}_i) \le d$ *and* $\mathbf{a}\mathbf{h}_i = s^i$*.*

Proof. Let $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ be a μ-basis of $\mathbf{a}$. We will proceed by induction on i.
Induction basis: For $i = 0$, the statement follows immediately from Theorem 75 and the well-known fact that $\operatorname{syz}(\mathbf{a})$ can be generated by vectors of degree at most d (see, for example, [Hon17] or [SG09]).
Induction step: Assume the statement is true in the i-th case i.e. there exists $\mathbf{h}_i \in \mathbb{K}[s]^n$ with $\deg(\mathbf{h}_i) \le d$ such that $\mathbf{a}\mathbf{h}_i = s^i$ $(i \le 2d-1)$. Then $\mathbf{a}(s\mathbf{h}_i) = s^{i+1}$. Let $\tilde{\mathbf{h}} = s\mathbf{h}_i$. Since $\deg(\mathbf{h}_i) \le d$, it follows that $\deg(\tilde{\mathbf{h}}) \le d+1$. If $\deg(\tilde{\mathbf{h}}) \le d$, let $\mathbf{h}_{i+1} = \tilde{\mathbf{h}}$ and we are done. Otherwise, $\deg(\tilde{\mathbf{h}}) = d+1$. Following a similar argument as in Theorem 75, the coefficient of $\mathbf{a}\tilde{\mathbf{h}}$ for s^{2d+1} is $LV(\mathbf{a})\,LV(\tilde{\mathbf{h}})$, and since we assumed $i \le 2d-1$, it must be that $LV(\mathbf{a})LV(\tilde{\mathbf{h}}) = 0$. Thus, there exist constants $\alpha_1, \dots, \alpha_{n-1}$ such that $LV(\tilde{\mathbf{h}}) = \sum_{j=1}^{n-1} \alpha_j LV(\mathbf{u}_j)$. Define

$$\mathbf{h}_{i+1} = \tilde{\mathbf{h}} - \sum_{j=1}^{n-1} \alpha_j \mathbf{u}_j s^{d+1-\deg(\mathbf{u}_j)}.$$

Then $\mathbf{a}\mathbf{h}_{i+1} = s^{i+1}$ and $\deg(\mathbf{h}_{i+1}) < \deg(\tilde{\mathbf{h}})$, which means $\deg(\mathbf{h}_{i+1}) \le d$. □

Proposition 86 (Full Rank)**.** *For a nonzero polynomial vector* $\mathbf{a}$ *of degree* d*, defined by* (4.11)*, such that* $\gcd(\mathbf{a}) = 1$*, the corresponding matrix* A*, defined by* (4.12)*, has rank* $2d+1$*.*

Proof. By Lemma 85, for all $i = 0, \dots, 2d$, there exist vectors $\mathbf{h}_i \in \mathbb{K}[s]^n$ with $\deg(\mathbf{h}_i) \le d$ such that $\mathbf{a}\mathbf{h}_i = s^i$. Observe that $(s^i)^\sharp = e_{i+1}$. Since $(\mathbf{a}\mathbf{h}_i)^\sharp = A\mathbf{h}_i^\sharp$, it follows that there exist vectors $\mathbf{h}_i^\sharp \in \mathbb{K}^{n(d+1)}$ such that $A\mathbf{h}_i^\sharp = e_j$ for all $j = 1, \dots, 2d+1$. This means the range of A is $\mathbb{K}^{2d+1}$ and hence $\operatorname{rank}(A) = 2d+1$. □

4.2.3 The minimal Bézout vector theorem

In this section, we construct a Bézout vector of $\mathbf{a}$ of minimal degree by finding an appropriate solution to the linear equation

$$A\,v = e_1, \text{ where } e_1 = [1, 0, \dots, 0]^T \in \mathbb{K}^{2d+1}. \tag{4.17}$$

The following lemma establishes a one-to-one correspondence between the set $\operatorname{Bez}_d(\mathbf{a})$ of Bézout vectors of $\mathbf{a}$ of degree at most d and the set of solutions to (4.17).

Lemma 87. *Let* $\mathbf{a} \in \mathbb{K}[s]_d^n$ *be a nonzero vector such that* $\gcd(\mathbf{a}) = 1$*. Then* $\mathbf{b} \in \mathbb{K}[s]_d^n$ *belongs to* $\operatorname{Bez}_d(\mathbf{a})$ *if and only if* $\mathbf{b}^\sharp$ *is a solution of* (4.17)*. Also* $v \in \mathbb{K}^{n(d+1)}$ *solves* (4.17) *if and only if* $v^\flat$ *belongs to* $\operatorname{Bez}_d(\mathbf{a})$*.*

Proof. Follows immediately from (4.16) and the observation that $e_1^{\flat^1_{2d}} = 1$. □

Thus, our goal is to construct a solution v of (4.17), such that $v^\flat$ is a Bézout vector of **a** of minimal degree. To accomplish this, we recall that, when $\gcd(\mathbf{a}) = 1$, Proposition 86 asserts that $\operatorname{rank}(A) = 2d+1$. Therefore, A has exactly $2d+1$ pivotal indices, which we can list in the increasing order $p = \{p_1, \dots, p_{2d+1}\}$. The corresponding columns of matrix A form a basis of $\mathbb{K}^{2d+1}$ and, therefore, $e_1 \in \mathbb{K}^{2d+1}$ can be expressed as a unique linear combination of the pivotal columns:

$$e_1 = \sum_{j=1}^{2d+1} \alpha_j A_{*p_j}. \tag{4.18}$$

Define vector $v \in \mathbb{K}^{2d+1}$ by setting its p_j-th element to be α_j and all other elements to be 0. We prove that $\mathbf{b} = v^\flat$ is a Bézout vector of **a** of minimal degree.

Theorem 88 (Minimal-Degree Bézout Vector)**.** *Let* $\mathbf{a} \in \mathbb{K}[s]^n_d$ *be a polynomial vector with* $\gcd(\mathbf{a}) = 1$*, and let A be the corresponding matrix defined by* (4.12)*. Let $p = \{p_1, \dots, p_{2d+1}\}$ be the pivotal indices of A, and let $\alpha_1, \dots, \alpha_{2d+1} \in \mathbb{K}$ be defined by the unique expression* (4.18) *of the vector $e_1 \in \mathbb{K}^{2d+1}$ as a linear combination of the pivotal columns of A. Define vector $v \in \mathbb{K}^{2d+1}$ by setting its p_j-th element to be α_j for $j = 1, \dots, 2d+1$ and all other elements to be 0, and let* $\mathbf{b} = v^\flat$*. Then*

1. $\mathbf{b} \in \operatorname{Bez}_d(\mathbf{a})$

2. $\deg(\mathbf{b}) = \min\limits_{\mathbf{b}' \in \operatorname{Bez}(\mathbf{a})} \deg(\mathbf{b}')$.

Proof.

1. From (4.18), it follows immediately that $Av = e_1$. Therefore, by Lemma 87, we have that $\mathbf{b} = v^\flat \in \operatorname{Bez}_d(\mathbf{a})$.

2. To show that **b** is of minimal degree, we rewrite (4.18) as

$$e_1 = \sum_{j=1}^{k} \alpha_j A_{*p_j}, \tag{4.19}$$

where k is the largest integer between 1 and $2d+1$, such that $\alpha_k \neq 0$. Then the last nonzero entry of v appears in p_k-th position and, therefore,

$$\deg(\mathbf{b}) = \deg(v^\flat) = \lceil p_k/n \rceil - 1. \tag{4.20}$$

Assume that $\mathbf{b}' \in \mathrm{Bez}(\mathbf{a})$ is such that $\deg(\mathbf{b}') < \deg(\mathbf{b})$. Then $\mathbf{b}' \in \mathrm{Bez}_d(\mathbf{a})$ and therefore $A\,v' = e_1$, for $v' = \mathbf{b}'^{\sharp} = [v'_1, \ldots, v'_{n(d+1)}] \in \mathbb{K}^{n(d+1)}$. Then

$$e_1 = \sum_{j=1}^{n(d+1)} v'_j A_{*j} = \sum_{j=1}^{r} v'_j A_{*j}, \tag{4.21}$$

where r is the largest integer between 1 and $n(d+1)$, such that $v'_r \neq 0$. Then

$$\deg(\mathbf{b}') = \lceil r/n \rceil - 1 \tag{4.22}$$

and since we assumed that $\deg(\mathbf{b}') < \deg(\mathbf{b})$, we conclude from (4.20) and (4.22) that $r < p_k$. On the other hand, since all non-pivotal columns are linear combinations of the preceding pivotal columns, we can rewrite (4.21) as

$$e_1 = \sum_{j \in \{1,\ldots,2d \,|\, p_j \leq r < p_k\}} \alpha'_j A_{*p_j} = \sum_{j=1}^{k-1} \alpha'_j A_{*p_j}. \tag{4.23}$$

By the uniqueness of the representation of e_1 as a linear combination of the A_{*p_j}, the coefficients in the expansions (4.19) and (4.23) must be equal, but $\alpha_k \neq 0$ in (4.19). Contradiction.

□

In the algorithm presented in Section 5.1, we exploit the fact that the coefficients α's in (4.19) needed to construct a minimal-degree Bézout vector of $\mathbf{a}$ can be read off the reduced row echelon form $[\hat{A}|\hat{v}]$ of the augmented matrix $[A|e_1]$. On the other hand, as was shown in Section 3.1 and reviewed in the next subsection, the coefficients of a μ-basis of $\mathbf{a}$ also can be read off the matrix $\hat{A}$. Therefore, a μ-basis is constructed as a byproduct of our algorithm for constructing a Bézout vector of minimal degree.

4.2.4 The μ-bases theorem

In Section 3.1, we showed that the coefficients of a μ-basis of $\mathbf{a}$ can be read off the basic non-pivotal columns of matrix A (recall Definition 80). Recall that according to Lemma 22, the matrix A has exactly $n-1$ basic non-pivotal columns.

Theorem 89 (μ-Basis)**.** *Let $\mathbf{a} \in \mathbb{K}[s]_d^n$ be a polynomial vector, and let A be the corresponding matrix defined by* (4.12)*. Let $\tilde{q} = [\tilde{q}_1, \ldots, \tilde{q}_{n-1}]$ be the basic non-pivotal indices of A, ordered increasingly.*

For $i = 1, \dots, n-1$, a basic non-pivotal column $A_{*\tilde{q}_i}$ is a linear combination of the previous pivotal columns:

$$A_{*\tilde{q}_i} = \sum_{\{r \in p \,|\, r < \tilde{q}_i\}} \alpha_{i\,r} A_{*r}, \tag{4.24}$$

for some $\alpha_{i\,r} \in \mathbb{K}$. Define vector $b_i \in \mathbb{K}^{2d+1}$ by setting its $\tilde{q}_i$-th element to be 1, its r-th element to be $-\alpha_{i\,r}$ for $r \in p$ such that $p_j < \tilde{q}_i$, and all other elements to be 0. Then the set of polynomial vectors

$$\mathbf{u}_1 = b_1^\flat, \quad \dots \quad , \mathbf{u}_{n-1} = b_{n-1}^\flat$$

is a degree-ordered μ-basis of $\mathbf{a}$.

Proof. The fact that $\mathbf{u}_1 = b_1^\flat, \quad \dots \quad , \mathbf{u}_{n-1} = b_{n-1}^\flat$ is a μ-basis of $\mathbf{a}$ is the statement of Theorem 31. By construction, the last nonzero entry of vector b_i is in the $\tilde{q}_i$-th position, and therefore for $i = 1, \dots, n-1$,

$$\deg(\mathbf{u}_i) = \deg(b_i^\flat) = \lceil \tilde{q}_i / n \rceil - 1.$$

Since the indices in $\tilde{q}$ are ordered increasingly, the vectors $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ are degree-ordered. □

The algorithm presented in Section 5.1 exploits the fact that the coefficients α's in (4.24) are already computed in the process of computing a Bézout vector of $\mathbf{a}$.

4.3 The degree of an optimal moving frame

Similarly to the degree of a polynomial vector (Definition 6), we define the degree of a polynomial matrix to be the maximum of the degrees of its entries. Obviously, for a given vector $\mathbf{a}$, all degree-optimal moving frames have the same degree. In this section, we establish the sharp upper and lower bounds on the degree of optimal moving frames. We also show that, for generic inputs, the degree of an optimal moving frame equals to the lower bound. An alternative simple proof of the bounds could be given using the fact that, when $\gcd(\mathbf{a}) = 1$, the sum of the degrees of a μ-basis of $\mathbf{a}$ equals to $\deg(\mathbf{a})$ (see Theorem 2 in [SG09]), along with the result relating the degree of a minimal-degree Bézout vector and the maximal degree of a μ-basis in Theorem 75 in this chapter. For the sharpness of the lower bound and its generality, one could use Proposition 3.3 of [CI15], determining the dimension of the set of vectors of a given μ-type, again combined with Theorem 75 in this chapter. Our results on the upper bound differ from what can be deduced from [CI15], because we allow components of $\mathbf{a}$ to be linearly dependent over $\mathbb{K}$. To keep the presentation self-contained, we give the proofs based on the results of this chapter. We will repeatedly use the following lemma.

Lemma 90. *Let* $\mathbf{a} \in \mathbb{K}[s]^n$ *be nonzero and let* A *be the corresponding matrix* (4.12). *Furthermore, let* k *be the maximum among the basic non-pivotal indices of* A*. Then the degree of any optimal moving frame at* $\mathbf{a}$ *equals to* $\lceil \frac{k}{n} \rceil - 1$.

Proof. It is straightforward to check that the maximal degree of the μ-basis, constructed in Theorem 89, has degree $\lceil \frac{k}{n} \rceil - 1$. From the optimality of the degrees property in Proposition 70, it follows that for any two degree-ordered μ-bases $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ and $\mathbf{u}'_1, \dots, \mathbf{u}'_{n-1}$ of $\mathbf{a}$ and for $i = 1, \dots, n-1$, we have $\deg(\mathbf{u}_i) = \deg(\mathbf{u}'_i)$. Therefore, the maximum degree of vectors in any μ-basis equals to $\lceil \frac{k}{n} \rceil - 1$. Theorem 75 implies that the degree of any optimal moving frame equals to the maximal degree of a μ-basis. □

Proposition 91 (Sharp Degree Bounds.)**.** *Let* $\mathbf{a} \in \mathbb{K}[s]^n$ *with* $\deg(\mathbf{a}) = d$ *and* $\gcd(\mathbf{a}) = 1$*. Then for every degree-optimal moving frame* P *at* $\mathbf{a}$*, we have* $\lceil \frac{d}{n-1} \rceil \le \deg(P) \le d$*, and these degree bounds are sharp. By sharp, we mean that for all* $n > 1$ *and* $d > 0$*, there exists an* $\mathbf{a} \in \mathbb{K}[s]^n$ *with* $\deg(\mathbf{a}) = d$ *and* $\gcd(\mathbf{a}) = 1$ *such that, for every degree-optimal moving frame* P *at* $\mathbf{a}$*, we have* $\deg(P) = \lceil \frac{d}{n-1} \rceil$*. Likewise, for all* $n > 1$ *and* $d > 0$*, there exists an* $\mathbf{a} \in \mathbb{K}[s]^n$ *with* $\deg(\mathbf{a}) = d$ *and* $\gcd(\mathbf{a}) = 1$ *such that, for every degree-optimal moving frame* P *at* $\mathbf{a}$*, we have* $\deg(P) = d$.

Proof.

1. (lower bound): Let P be a degree-optimal moving frame at $\mathbf{a}$. Then $\mathbf{a}P = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$. and so from Cramer's rule:
$$\mathbf{a}_i = \frac{(-1)^{i+1}}{|P|} \left| P_{i,1} \right| \quad i = 1, \dots n,$$
where $P_{i,1}$ denotes the submatrix of P obtained by removing the 1-st column and the i-th row. We remind the reader that $|P|$ is a nonzero constant. Assume for the sake of contradiction that $\deg(P) < \lceil \frac{d}{n-1} \rceil$. Then $\deg(P) < \frac{d}{n-1}$. Since $\left| P_{i,1} \right|$ is the determinant of an $(n-1)\times(n-1)$ submatrix of P, we have $\deg(\mathbf{a}_i) = \deg(\left| P_{i,1} \right|) < (n-1)\frac{d}{n-1} = d$ for all $i = 1 \dots, n$. This contradicts the assumption that $\deg(\mathbf{a}) = d$. Thus, $\deg(P) \ge \lceil \frac{d}{n-1} \rceil$.

   We will prove that the lower bound $\lceil \frac{d}{n-1} \rceil$ is sharp by showing that, for all $n > 1$ and $d > 0$, the following matrix

$$P=\left[\begin{array}{ccc|ccccc} 1 & & & -s^{d-k\lceil \frac{d}{n-1}\rceil} & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ \hline & & & 1 & -s^{\lceil\frac{d}{n-1}\rceil} & & & \\ & & & & 1 & -s^{\lceil\frac{d}{n-1}\rceil} & & \\ & & & & & \ddots & \ddots & \\ & & & & & & \ddots & -s^{\lceil\frac{d}{n-1}\rceil} \\ & & & & & & & 1 \end{array}\right] \tag{4.25}$$

has degree $\lceil\frac{d}{n-1}\rceil$ and is a degree-optimal moving frame at the vector

$$\mathbf{a}=\left[1,0,\dots,0,s^{d-k\cdot\lceil\frac{d}{n-1}\rceil},\dots,s^{d-2\cdot\lceil\frac{d}{n-1}\rceil},s^{d-1\cdot\lceil\frac{d}{n-1}\rceil},s^{d-0\cdot\lceil\frac{d}{n-1}\rceil}\right]. \tag{4.26}$$

Here $k\in\mathbb{N}$ is the maximal such that $d>k\lceil\frac{d}{n-1}\rceil$ (explicitly $k=\left\lceil\frac{d}{\lceil\frac{d}{n-1}\rceil}\right\rceil-1$), the number of zeros in $\mathbf{a}$ is $n-k-2$, the upper-left block of P is of the size $(n-k-1)\times(n-k-1)$, the lower-right block is of the size $(k+1)\times(k+1)$, and the other two blocks are of the appropriate sizes.

First, we show that such $\mathbf{a}$ and P actually exist (not just optically). That is, the number of zeros in $\mathbf{a}$ is non-negative, and the upper-left block in P exists; in other words, $n-1\geq k+1$. Suppose otherwise. Then we would have

$$d-k\left\lceil\frac{d}{n-1}\right\rceil\leq d-(n-1)\left\lceil\frac{d}{n-1}\right\rceil\leq 0$$

which contradicts the condition $d>k\lceil\frac{d}{n-1}\rceil$.

Second, P is a degree-optimal moving frame at $\mathbf{a}$. Namely,

(a) $\mathbf{a}P=[1,0,\dots,0]$, so P is a moving frame at $\mathbf{a}$.

(b) The first column of P, $[1,0,\dots,0]^T$, is a minimal-degree Bézout vector of $\mathbf{a}$.

(c) The last $n-1$ columns of P are syzygies of $\mathbf{a}$, and since $P\in\mathrm{mf}(\mathbf{a})$, by Proposition 67, they form a basis of syz($\mathbf{a}$). It is easy to see that these columns have linearly independent leading vectors as well. Thus, they form a μ-basis of $\mathbf{a}$.

Finally, we show that the degree of P is the lower bound, i.e. $\lceil\frac{d}{n-1}\rceil$. From inspection of the entries of P, we see immediately that

$$\deg(P) = \max\left\{ d - k\left\lceil \frac{d}{n-1} \right\rceil, \left\lceil \frac{d}{n-1} \right\rceil \right\}.$$

It remains to show that $d - k\lceil \frac{d}{n-1} \rceil \leq \lceil \frac{d}{n-1} \rceil$. Suppose not. Then

$$d > (k+1)\left\lceil \frac{d}{n-1} \right\rceil,$$

a contradiction to the maximality of k. Thus, $\deg(P) = \lceil \frac{d}{n-1} \rceil$. Hence, we have proved that the lower bound is sharp.

2. (upper bound): From Theorems 88 and 89, it follows immediately that d is an upper bound of a degree-optimal moving frames. We will prove that the upper bound d is sharp by showing that, for all $n > 1$ and $d > 0$, the following matrix of degree d

$$P = \begin{bmatrix} 1 & & & -s^d \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots \\ & & & 1 \end{bmatrix}. \tag{4.27}$$

is a degree-optimal moving frame for the vector

$$\mathbf{a} = [1, 0, \ldots, 0, s^d]$$

Indeed:

(a) $\mathbf{a}P = [1, 0, \ldots, 0]$ and so P is a moving frame at $\mathbf{a}$.

(b) The first column of P, $[1, 0, \ldots, 0]^T$, is a minimal-degree Bézout vector of $\mathbf{a}$.

(c) The last $n-1$ columns of P are syzygies of $\mathbf{a}$, and since $P \in \mathrm{mf}(\mathbf{a})$, by Proposition 67, they form a basis of $\mathrm{syz}(\mathbf{a})$. It is easy to see that these columns have linearly independent leading vectors as well. Thus, they form a μ-basis of $\mathbf{a}$.

□

In Theorem 95 below, we show that for generic $\mathbf{a} \in \mathbb{K}[s]^n$ with $\deg(\mathbf{a}) = d$ and $\gcd(\mathbf{a}) = 1$, and for all degree-optimal moving frames P at $\mathbf{a}$, $\deg(P) = \lceil \frac{d}{n-1} \rceil$. To prove the theorem, we need the following lemmas, where we will use notation

$$k = \mathrm{quo}(d, n-1) \text{ and } r = \mathrm{rem}(d, n-1).$$

Lemma 92. *For arbitrary* $\mathbf{a} \in \mathbb{K}[s]^n$ *with* $\deg(\mathbf{a}) = d$ *and* $\gcd(\mathbf{a}) = 1$*, the principal* $d+k+1$ *submatrix of the associated matrix* A *has the form*

$$C = \begin{bmatrix} c_{01} & \cdots & \cdots & c_{0n} & & & & & & & & \\ \vdots & \cdots & \cdots & \vdots & c_{01} & \cdots & \cdots & c_{0n} & & & & \\ \vdots & \cdots & \cdots & \vdots & \vdots & \cdots & \cdots & \vdots & \ddots & & & \\ c_{d1} & \cdots & \cdots & c_{dn} & \vdots & \cdots & \cdots & \vdots & \ddots & c_{01} & \cdots & c_{0,r+1} \\ & & & & c_{d1} & \cdots & \cdots & c_{dn} & \ddots & \vdots & \cdots & \vdots \\ & & & & & & & & \ddots & \vdots & \cdots & \vdots \\ & & & & & & & & & c_{d1} & \cdots & c_{d,r+1} \end{bmatrix}, \tag{4.28}$$

where C *consists of* k *full* $(d+1)\times n$ *size blocks and 1 partial block of size* $(d+1)\times(r+1)$*.*

Proof. If we take k full $(d+1)\times n$ blocks and 1 partial $(d+1)\times(r+1)$ block, then the number of columns of C is $nk+r+1 = (n-1)k+r+k+1 = d+k+1$, as desired. Furthemore, since the leftmost block takes up the first $d+1$ rows of C, and we shift the block down by 1 a total of k times, the number of rows of C is $d+k+1$ as well. □

Lemma 93. *Let* $\mathbf{a} \in \mathbb{K}[s]^n$ *with* $\deg(\mathbf{a}) = d$ *and* $\gcd(\mathbf{a}) = 1$*, and let* C *be the principal* $d+k+1$ *submatrix of* A *given by* (4.28)*. If* C *is nonsingular, then for any degree-optimal moving frame* P *at* $\mathbf{a}$*, we have* $\deg(P) = \lceil \frac{d}{n-1} \rceil$*.*

Proof. If C is nonsingular, then first $d+k+1$ columns of the matrix A are pivotal columns. Since $\mathrm{rank}(A) = 2d+1$, there are $d-k$ additional pivotal columns in A and, from the structure of A, each of the last $d-k$ blocks of A contain exactly one of these additional pivotal columns. All other columns in A are non-pivotal. We now consider two cases:

1) If $n-1$ divides d, then $r=0$ and $k = \frac{d}{n-1} = \lceil \frac{d}{n-1} \rceil$. Thus, there is one column in the partial block in C, and so the remaining $n-1$ columns in this $(k+1)$-th block of A are basic non-pivotal columns. Since in total there are $n-1$ basis non-pivotal columns, the largest basic non-pivotal index equals to $n(k+1)$, and therefore by Lemma 90, the degree of any optimal moving frame at $\mathbf{a}$ is $\lceil \frac{d}{n-1} \rceil$.

2) If $n-1$ does not divide d, then $r > 0$ and $k = \lfloor \frac{d}{n-1} \rfloor$. Thus, there are at least two columns in the partial block in C, and so there are at most $n-2$ basic non-pivotal columns in the $(k+1)$-th block of A. Since there are a total of $n-1$ basis non-pivotal columns, and all but one of the columns in the $(k+2)$-th block are non-pivotal, the largest basic non-pivotal column index will appear in the $(k+2)$-th block. Therefore, this largest index equals to $n(k+1)+j$ for some $1 \le j \le n$. By Lemma 90, the degree of any optimal moving frame at $\mathbf{a}$ equals to $\lceil \frac{n(k+1)+j}{n} \rceil - 1 = k+1 = \lfloor \frac{d}{n-1} \rfloor + 1 = \lceil \frac{d}{n-1} \rceil$.

□

Lemma 94. *For all* $n > 1$ *and* $d > 0$*, there exists a vector* $\mathbf{a} \in \mathbb{K}[s]^n$ *with* $\deg(\mathbf{a}) = d$ *and* $\gcd(\mathbf{a}) = 1$ *such that* $\det(C) \neq 0$.

Proof. Let $n > 1$ and $d > 0$. We will find a suitable witness for $\mathbf{a}$. Recalling the relation $d = k(n-1)+r$, we will consider the following three cases:

1) If $n-1 > d$, we claim that the following is a witness:

$$\mathbf{a} = \left[s^d, s^{d-1}, \ldots, s, 1, \ldots, 1\right]$$

Note that there is at least one 1 at the end. Thus $\deg(\mathbf{a}) = d$ and $\gcd(\mathbf{a}) = 1$. It remains to show that $|C| \neq 0$. Note that $k = 0$ and $r = d$. Thus, the matrix C is a $(d+1) \times (d+1)$ partial block that looks like

$$C = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}.$$

Therefore, $|C| = \pm 1$.

2) If $n-1 \le d$ and $n-1$ divides d, we claim that the following is a witness:

$$\mathbf{a} = \left[s^d, s^{d-k}, \ldots, s^{d-(n-1)k}\right]$$

Note that the last component is $s^{d-(n-1)k} = s^0 = 1$. Thus $\deg(\mathbf{a}) = d$ and $\gcd(\mathbf{a}) = 1$. It remains to show that $|C| \neq 0$. To do this, we examine the shape of C. To get intuition, consider the instance where $n = 3$ and $d = 6$. Note that $k = 3$ and $r = 0$. Thus, we have

$$\begin{aligned} a &= \left[s^6, s^3, s^0\right] \\ C &= \left[\begin{array}{cccccccccc} 0 & 0 & \mathbf{1} & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \\ 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & \mathbf{1} & 0 & 0 & 0 \\ & & & & & & & & & \mathbf{1} \end{array}\right] \end{aligned}$$

All the empty spaces are zeros. Note that C is a permutation matrix (each row has only one 1 and each column has only one 1). Therefore, $|C| = \pm 1$. It is easy to see that the same observation holds in general.

3) If $n-1 \le d$ and $n-1$ does not divide d, we claim that the following is a witness:

$$\mathbf{a} = \left[s^d, s^{d-(1k+1)}, s^{d-(2k+2)} \ldots, s^{d-(rk+r)}, s^{d-((r+1)k+r)}, \ldots, s^{d-((n-1)k+r)}\right]$$

Note that the last component is $s^{d-((n-1)k+r)} = s^0 = 1$. Thus $\deg(\mathbf{a}) = d$ and $\gcd(\mathbf{a}) = 1$. It remains to show that $|C| \neq 0$. To do this, we examine the shape of C. To get intuition, consider the case $n = 5$ and $d = 14$. Note that $k = 3$ and $r = 2$. Thus, we have

$$\begin{aligned} a &= \left[s^{14}, s^{14-(1\cdot3+1)}, s^{14-(2\cdot3+2)}, s^{14-(3\cdot3+2)}, s^{14-(4\cdot3+2)}\right] = \left[s^{14}, s^{10}, s^{6}, s^{3}, s^{0}\right] \\ C &= \begin{bmatrix} 0&0&0&0&\mathbf{1}&&&&&&&&&&&&& \\ 0&0&0&0&0&0&0&0&0&\mathbf{1}&&&&&&&& \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&&& \\ 0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0 \\ 0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&\mathbf{1} \\ 0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&0 \\ \mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ &&&&&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0 \\ &&&&&&&&&&\mathbf{1}&0&0&0&0&0&0&0 \\ &&&&&&&&&&&&&&&\mathbf{1}&0&0 \end{bmatrix} \end{aligned}$$

All the empty spaces are zeros. Note that C is a permutation matrix (each row has only one 1 and each column has only one 1). Therefore, $|C| = \pm 1$. It is easy to see that the same observation holds in general.

□

Theorem 95 (Generic Degree.). *Let $\mathbb{K}$ be an infinite field. For generic $\mathbf{a} \in \mathbb{K}[s]^n$ with $\deg(\mathbf{a}) = d$ and $\gcd(\mathbf{a}) = 1$, for every degree-optimal moving frame P at $\mathbf{a}$, we have $\deg(P) = \left\lceil \frac{d}{n-1} \right\rceil$.*

Proof. From Lemma 94, it follows that $\det(C)$ is a nonzero polynomial on the $n(d+1)$-dimensional vector space $\mathbb{K}[s]^n$ over $\mathbb{K}$. Thus, the condition $\det(C) \neq 0$ defines a proper Zariski open subset of $\mathbb{K}[s]^n$. Lemma 93 implies that for every $\mathbf{a}$ in this Zariski open subset, every degree-optimal moving frame P at $\mathbf{a}$ has degree $\left\lceil \frac{d}{n-1} \right\rceil$. If we assume $\mathbb{K}$ is an infinite field, then the complement of any proper Zariski open subset is of measure zero, and we can say that for a generic $\mathbf{a}$, the degree of every degree-optimal moving frame at $\mathbf{a}$ equals the sharp lower bound $\left\lceil \frac{d}{n-1} \right\rceil$. □

Remark 96. *Some simple consequences of the general results about the degrees are worthwhile recording. From Proposition 91, it follows that, when $d \geq n$, the degree of an optimal moving frame* *is always strictly greater than 1. From the above theorem and Theorem 75, it follows that when $d < n$ and $\mathbb{K}$ is infinite, then for a generic input, the degree of an optimal moving frame is 1 and the minimal-degree Bézout vector is a constant vector.*

CHAPTER

5

ALGEBRAIC MOVING FRAMES: COMPUTATION

Having established, in Chapter 4, the definitions, background, and building blocks of moving frames, we now turn our attention to computing algebraic moving frames and degree-optimal moving frames. In particular, in Section 5.1 we present an OMF algorithm built upon the theory described in Sections 4.1 and 4.2. The algorithm exploits the fact that the construction procedures for a minimal-degree Bézout vector and for a μ-basis, suggested by Theorems 88 and 89, can be accomplished simultaneously by a single partial row-echelon reduction of a $(2d+1)\times(nd+n+1)$ matrix over $\mathbb{K}$. In Proposition 100, we prove that the theoretical (worst-case asymptotic) complexity of the OMF algorithm equals to $O(d^2n+d^3+n^2)$, and we trace the algorithm on our running example. In Section 5.2, we present a slight modification of the OMF algorithm that produces degree-optimal moving frames for input vectors with non-trivial gcd. In Section 5.3, we show that important algebraic properties of the frames produced by the OMF algorithm can be enhanced by a group-equivariant property which plays a crucial role in geometric moving frame theory. We also show that a simple modification of any deterministic algorithm for producing a degree-optimal algebraic moving frame leads to an algorithm that produces a $GL_n(\mathbb{K})$-equivariant degree-optimal moving frame.

We then turn our attention to alternative moving frame algorithms. In Section 5.4, we discuss other possible approaches for computing degree-optimal moving frames. As mentioned above, there are no existing algorithms for simultaneously computing a minimal-degree Bézout vector and a μ-basis. We describe instead a two-step approach that involves performing reduction on non-optimal moving frames. In Sections 5.5, 5.6, 5.7, and 5.8, we present algorithms for computing algebraic moving frames that are not necessarily optimal. In Sections 5.9 and 5.10, we present new algorithms for computing a degree-optimal moving frame simultaneously, using reduced TOP-Gröbner basis computations and μ-basis computations, respectively. In Section 5.11, we consider the natural generalization of the moving frame problem to unimodular matrix inputs $\mathbf{A}$.

5.1 The OMF Algorithm

The theory developed in Chapter 4, namely Sections 4.1 and 4.2, can be recast into an algorithm for computing a degree-optimal moving frame. After giving its informal outline, we provide an optimized version of the algorithm, trace it, and analyze its theoretical complexity.

5.1.1 Informal outline

Before stating a rigorous and optimized version of the algorithm, we make an informal outline and indicate how the optimization is done:

1. For an input vector $\mathbf{a} \in \mathbb{K}[s]_d^n$ such that $\gcd(\mathbf{a}) = 1$, construct the augmented matrix $W = [A \mid e_1] \in \mathbb{K}^{(2d+1)\times(nd+n+1)}$, where $A \in \mathbb{K}^{(2d+1)\times n(d+1)}$ is given by (4.12) and $e_1 = [1, 0, \dots, 0]^T \in \mathbb{K}^{2d+1}$.

2. Compute the reduced row-echelon form $E = [\hat{A} \mid \hat{v}]$ of W.

3. Construct a matrix $P \in \mathbb{K}[s]^{n\times n}$ whose first column is a Bézout vector of $\mathbf{a}$ of minimal degree and whose last $n-1$ columns form a μ-basis of $\mathbf{a}$, as follows:

   (a) Construct the matrix $V \in \mathbb{K}^{n(d+1)\times n}$ whose first column solves $\hat{A}v = \hat{v}$ and whose last $n-1$ columns are the null vectors of A corresponding to its basic non-pivotal columns. Here $p = [p_1, \dots, p_{2d+1}]$ is the list of the pivotal indices and $\tilde{q} = [\tilde{q}_1, \dots, \tilde{q}_{n-1}]$ is the list of the basic non-pivotal indices of A.

   - $V_{p_r,1} = \hat{v}[r]$ for $r = 1, \dots, 2d+1$
   - $V_{\tilde{q}_{j-1},j} = 1$ for $j = 2, \dots, n$
   - $V_{p_r,j} = -E_{r,\tilde{q}_{j-1}}$ for $j = 2, \dots, n$ and $r \in \{1, \dots, 2d+1 \mid p_r < \tilde{q}_{j-1}\}$

- All other entries are zero

(b) Use the isomorphism $\flat$ to convert matrix V into $P = [V_{*1}^{\flat}, \ldots, V_{*n}^{\flat}]$.

However, steps 2 and 3 do some wasteful operations and they can be improved, as follows:

- Note that step 2 constructs the entire reduced row-echelon form of W, even though we only need $n-1$ null vectors corresponding to its basic non-pivot columns and the single solution vector. Hence, it is natural to optimize this step so that only the $n-1$ null vectors and the single solution vector are constructed: instead of using Gauss-Jordan elimination to compute the entire reduced row-echelon form, one performs operations column by column only on the pivot columns, basic non-pivot columns, and augmented column.

- Note that step 3 constructs the entire matrix V even though many entries are zero. Hence, it is natural to optimize this step so that we bypass constructing the matrix V, but instead construct the matrix P directly from the matrix E. This is possible because the matrix E contains all the information about the matrix V.

5.1.2 Formal algorithm and proof

In this section, $\text{quo}(i, j)$ denotes the quotient and $\text{rem}(i, j)$ denotes the remainder generated by dividing an integer i by an integer j.

Algorithm 97 (**OMF**)**.**

Input: $\mathbf{a} \neq 0 \in \mathbb{K}[s]^n$*, row vector, where* $n > 1$, $\gcd(\mathbf{a}) = 1$*, and* $\mathbb{K}$ *a computable field*

Output: $P \in \mathbb{K}[s]^{n \times n}$*, a degree-optimal moving frame at* $\mathbf{a}$

1. *Construct a matrix* $W \in \mathbb{K}^{(2d+1)\times(nd+n+1)}$*, whose left* $(2d+1)\times(nd+n)$ *block is matrix* (4.12) *and whose last column is* e_1*.*

   (a) $d \longleftarrow \deg(\mathbf{a})$

   (b) Identify the row vectors $c_0 = [c_{01}, \ldots c_{0n}], \ldots, c_d = [c_{d1}, \ldots c_{dn}]$ *such that* $\mathbf{a} = c_0 + c_1 s + \cdots + c_d s^d$*.*

   (c) $$W \longleftarrow \left[\begin{array}{ccc|c} c_0 & & & 1 \\ \vdots & \ddots & & 0 \\ c_d & \vdots & c_0 & \\ & \ddots & \vdots & \vdots \\ & & c_d & 0 \end{array}\right] \in \mathbb{K}^{(2d+1)\times(nd+n+1)}$$

2. *Construct the "partial" reduced row-echelon form E of W.*

   This can be done by using Gauss-Jordan elimination (forward elimination, backward elimination, and normalization), with the following optimizations:

   - *Skip over periodic non-pivot columns.*
   - *Carry out the row operations only on the required columns.*

3. *Construct a matrix $P \in \mathbb{K}[s]^{n\times n}$ whose first column is a Bézout vector of* **a** *of minimal degree and whose last $n-1$ columns form a μ-basis of* **a**.

   Let p be the list of the pivotal indices and let $\tilde{q}$ be the list of the basic non-pivotal indices of E.

```
(a) Initialize an $n \times n$ matrix $P$ with $0$ in every entry.
(b) For $j = 2, \dots, n$
        $r \leftarrow \mathrm{rem}\big(\tilde{q}_{j-1} - 1, n\big) + 1$
        $k \leftarrow \mathrm{quo}\big(\tilde{q}_{j-1} - 1, n\big)$
        $P_{r,j} \leftarrow P_{r,j} + s^k$
(c) For $i = 1, \dots, 2d+1$
        $r \leftarrow \mathrm{rem}\big(p_i - 1, n\big) + 1$
        $k \leftarrow \mathrm{quo}\big(p_i - 1, n\big)$
        $P_{r,1} \leftarrow P_{r,1} + E_{i,nd+n+1} s^k$
        For $j = 2, \dots, n$
            $P_{r,j} \leftarrow P_{r,j} - E_{i,\tilde{q}_{j-1}} s^k$
```

Theorem 98. *The output of the OMF Algorithm is a degree-optimal moving frame at* $\mathbf{a}$, *where* $\mathbf{a}$ *is the input vector* $\mathbf{a} \in \mathbb{K}[s]^n$ *such that* $n > 1$ *and* $\gcd(\mathbf{a}) = 1$.

Proof. In step 1, we construct a matrix $W = [A \mid e_1] \in \mathbb{K}^{(2d+1)\times(nd+n+1)}$ whose left $(2d+1)\times(nd+n)$ block is matrix (4.12) and whose last column is $e_1 = [1,0,\dots,0]^T$. Under isomorphism $\flat$, the null space of A corresponds to $\mathrm{syz}_d(\mathbf{a})$, and the solutions to $Av = [1,0,\dots,0]^T$ correspond to $\mathrm{Bez}_d(\mathbf{a})$. From Proposition 86, we know that $\mathrm{rank}(A) = 2d+1$, and thus all pivotal columns of W are the pivotal columns of A. In step 2, we perform partial Gauss-Jordan operations on W to identify the coefficients α's appearing in (4.24) and (4.18), that express the $n-1$ basic non-pivotal columns of A and the vector e_1, respectively, as linear combinations of pivotal columns of A. These coefficients will appear in the basic non-pivotal columns and the last column of the partial reduced row-echelon form E of W. In Step 3, we use these coefficients to construct a minimal-degree Bézout vector of $\mathbf{a}$ and a degree-ordered μ-basis of $\mathbf{a}$, as prescribed by Theorems 88 and 89. We place these vectors as the columns of matrix P, and the resulting matrix is, indeed, a degree-optimal moving frame according to Theorem 74. □

Example 99. *We trace the algorithm on the input vector*

$$\mathbf{a} = \left[\begin{array}{ccc} 2+s+s^4 & 3+s^2+s^4 & 6+2s^3+s^4 \end{array}\right] \in \mathbb{Q}[s]^3$$

1. *Construct matrix W:*

   (a) $d \longleftarrow 4$

   (b) $c_0, c_1, c_2, c_3, c_4 \longleftarrow [\,2\ 3\ 6\,], [\,1\ 0\ 0\,], [\,0\ 1\ 0\,], [\,0\ 0\ 2\,], [\,1\ 1\ 1\,]$

   (c) $$W \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc|c}
2 & 3 & 6 & & & & & & & & & & & & & 1 \\
1 & 0 & 0 & 2 & 3 & 6 & & & & & & & & & & \\
0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & & & & & \\
0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & & \\
1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & \\
 & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & \\
 & & & & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 0 & 2 & \\
 & & & & & & & & & & & & 1 & 1 & 1 &
\end{array}\right]$$

2. *Construct the "partial" reduced row-echelon form E of W.*

   $$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc|c}
1 & & & & & & & -3 & -9 & & & & & & & 2 \\
 & 1 & & & & & & -2 & -8 & & & & & & & 1 \\
 & & 1 & & & & & 2 & 7 & & & & & & & -1 \\
 & & & 1 & & & & 3 & 12 & & 3 & 6 & & & & -1 \\
 & & & & 1 & & & -5 & -15 & & 0 & 0 & & 3 & 6 & 2 \\
 & & & & & 1 & & 2 & 5 & & 1 & 0 & & 0 & 0 & -1 \\
 & & & & & & 1 & 1 & 1 & & 0 & 2 & & 1 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & & 0 & 2 & 0 \\
 & & & & & & & & & & & & 1 & 1 & 1 & 0
\end{array}\right]$$

   Here, blue denotes pivotal columns, red denotes basic non-pivotal columns, brown denotes periodic non-pivotal columns, and grey denotes the solution column.

3. *Construct a matrix $P \in \mathbb{K}[s]^{n\times n}$ whose first column consists of a minimal-degree Bézout vector of $\mathbf{a}$ and whose last $n-1$ columns form a μ-basis of $\mathbf{a}$.*

(a) $P \longleftarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(b) $P \longleftarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & s^2 & 0 \\ 0 & 0 & s^2 \end{bmatrix}$

(c) $P \longleftarrow \begin{bmatrix} 2-s & 3-3s-s^2 & 9-12s-s^2 \\ 1+2s & 2+5s+s^2 & 8+15s \\ -1-s & -2-2s & -7-5s+s^2 \end{bmatrix}$

Proposition 100 (Theoretical Complexity)**.** *Under the assumption that the time for any arithmetic operation is constant, the complexity of the OMF algorithm is* $O(d^2n + d^3 + n^2)$.

Proof. We will trace the theoretical complexity for each step of the algorithm.

1. (a) To determine d, we scan through each of the n polynomials in a to identify the highest degree term, which is always $\leq d$. Thus, the complexity for this step is $O(dn)$.

   (b) We identify $n(d+1)$ values to make up $c_0, \ldots, c_d$. Thus, the complexity for this step is $O(dn)$.

   (c) We construct a matrix with $(2d+1)(nd+n+1)$ entries. Thus, the complexity for this step is $O(d^2n)$.

2. With the partial Gauss-Jordan elimination, we perform row operations only on the $2d+1$ pivot columns of A, the $n-1$ basic non-pivot columns of A, and the augmented column e_1. Thus, we perform Gauss-Jordan elimination on a $(2d+1)\times(2d+n+1)$ matrix. In general, for a $k \times l$ matrix, Gauss-Jordan elimination has complexity $O(k^2l)$. Thus, the complexity for this step is $O(d^2(d+n))$.

3. (a) We fill 0 into the entries of an $n \times n$ matrix P. Thus, the complexity for this step is $O(n^2)$.

   (b) We update entries of the matrix $n-1$ times. Thus, the complexity for this step is $O(n)$.

   (c) We update entries of the matrix $(2d+1)(n-1)$ times. Thus, the complexity for this step is $O(dn)$.

By summing up, we have $O\left(dn + dn + d^2n + d^2(d+n) + n^2 + n + dn\right) = O\left(d^2n + d^3 + n^2\right)$ □

Remark 101. *Note that the n^2 term in the above complexity is solely due to step 3(a), where the matrix P is initialized with zeros. If one uses a sparse representation of the matrix (storing only nonzero elements), then one can skip the initialization of the matrix P. As a result, the complexity can be improved to $O\left(d^2 n + d^3\right)$.*

It turns out that the theoretical complexity of the OMF algorithm is exactly the same as that of the μ-basis algorithm presented in Section 3.1. This is unsurprising, because the μ-basis algorithm presented in Section 3.1 is based on partial Gauss-Jordan elimination of matrix A, while the OMF algorithm is based on partial Gauss-Jordan elimination of the matrix obtained by appending to A a single column e_1.

5.2 Case when gcd $\neq 1$

In this section, we make a slight modification to the OMF algorithm, so that it produces degree-optimal moving frames in the case when $\gcd(\mathbf{a}) \neq 1$. As before, we accomplish this by reducing to a linear algebra problem.

Let $\mathbf{a} = \sum_{0 \le j \le d} c_j s^j \in \mathbb{K}[s]_d^n$, where $c_j = [c_{1j}, \ldots, c_{nj}] \in \mathbb{K}^n$ are row vectors, and suppose $\deg(\gcd(\mathbf{a})) = D$. Define

$$A' = \begin{bmatrix} c_D & \cdots & c_0 & & \\ \vdots & \ddots & \vdots & \ddots & \\ \vdots & \vdots & c_D & & c_0 \\ c_d & \vdots & \vdots & \ddots & \vdots \\ & \ddots & \vdots & & c_D \\ & & c_d & & \vdots \\ & & & \ddots & \vdots \\ & & & & c_d \end{bmatrix} \in \mathbb{K}^{(2d+1-D)\times n(d+1)} \tag{5.1}$$

and observe that A' is our usual matrix A given in (4.12) with the first D rows removed. We then have the following.

Lemma 102. *For any $v \in \mathbb{K}^{n(d+1)}$,*

$$\mathbf{a} v^{\flat_d^n} = \gcd(\mathbf{a})(A' v)^{\flat_{2d-D}^1}. \tag{5.2}$$

Proof. A vector $v \in \mathbb{K}^{n(d+1)}$ can be split into $(d+1)$ blocks

$$\begin{bmatrix} w_0 \\ \vdots \\ w_d \end{bmatrix},$$

where $w_i \in \mathbb{K}^n$ are column vectors. For $j < 0$ and $j > d$, let us define $c_j = 0 \in \mathbb{K}^n$. Then $A'v$ is a $(2d+1-D)$-vector with $(k+1)$-th entry

$$(A'v)_{k+1} = c_{k+D}\, w_0 + c_{k+D-1}\, w_1 + \cdots + c_{k+D-d}\, w_d = \sum_{0 \le i \le d} c_{k+D-i}\, w_i,$$

where $k = 0, \ldots, 2d-D$. Then

$$\begin{aligned} \mathbf{a}v^\flat &= \mathbf{a}S_d^n\, v = \left(\sum_{0\le j\le d} c_j s^j\right)\left(\sum_{0\le i\le d} w_i\, s^i\right) = \sum_{0\le i,j\le d} c_j\, w_i s^{i+j} \\ &= \sum_{0\le k\le 2d} s^k \left(\sum_{0\le i\le d} c_{k-i}\, w_i\right) \end{aligned}$$

Since $\deg(\gcd(\mathbf{a})) = D$, we can factor out $\gcd(\mathbf{a})$ to obtain

$$\begin{aligned} \mathbf{a}v^\flat &= \gcd(\mathbf{a}) \sum_{0\le k\le 2d-D} s^k \left(\sum_{0\le i\le d} c_{k+D-i}\, w_i\right) = \gcd(\mathbf{a}) \sum_{0\le k\le 2d-D} s^k\, (A'v)_{k+1} \\ &= \gcd(\mathbf{a}) S^1_{2d-D}\, (A'v) = \gcd(\mathbf{a})(A'v)^{\flat^1_{2d-D}}. \end{aligned}$$

□

Lemma 103. $v^\flat \in \mathrm{Bez}_d(\mathbf{a})$ *if and only if* v *solves the equation* $A'v = e_1$*, where* $e_1 = [1, 0, \ldots, 0]^T \in \mathbb{K}^{2d+1-D}$.

Proof. Follows immediately from (5.2) and the observation that $e_1^{\flat^1_{2d-D}} = 1$. □

Thus, we've once again reduced our problem to solving a matrix equation, in this case $A'v = e_1$. To set up this equation, though, we need some way to determine the value of D. This can be done using the fact (see [VS78]) that the rank deficiency of the original matrix A given by (4.12) is D, or in other words, the number of pivotal columns of A is $2d+1-D$. This suggests the following informal algorithm:

1. Given $\mathbf{a} \in \mathbb{K}[s]_d^n$, form the matrix $A \in \mathbb{K}^{(2d+1)\times n(d+1)}$ as given by (4.12).

2. Compute the partial reduced row-echelon form E of A, keeping track of the list p of pivotal columns and list $\tilde{q}$ of basic non-pivotal columns.

3. Construct the augmented matrix $W = [A' \mid e_1] \in \mathbb{K}^{(2d+1-D)\times(nd+n+1)}$, where $D = 2d+1-|p|$ and A' is given by (5.1).

4. Compute the reduced row-echelon form $E' = [\widehat{A'}, \hat{v}]$ of W.

5. Construct $P \in \mathbb{K}[s]^{n\times n}$ from $\hat{v}$ and the columns of E indexed by $\tilde{q}$.

Recall from Section 3.1 that a μ-basis can be formed from the partial RREF of matrix A (denoted by E) even when $\gcd(\mathbf{a}) \neq 1$. Moreover, step 3 of the above procedure can be modified by observing that A and A' have the same rank and the same pivotal indices. Thus, instead of using the entire matrix A', we can simply use the pivotal columns that we determine in step 2 and call the truncated matrix A'_p. We now present the formal algorithm.

Algorithm 104 (**GOMF**)**.**

Input: $\mathbf{a} \neq 0 \in \mathbb{K}[s]^n$*, row vector, where* $n > 1$*, and* $\mathbb{K}$ *a computable field*

Output: $P \in \mathbb{K}[s]^{n\times n}$*, a degree-optimal moving frame at* $\mathbf{a}$

1. *Construct matrix* $A \in \mathbb{K}^{(2d+1)\times n(d+1)}$

   (a) $d \longleftarrow \deg(\mathbf{a})$

   (b) *Identify the row vectors* $c_0 = [c_{01}, \dots c_{0n}], \dots, c_d = [c_{d1}, \dots c_{dn}]$ *such that* $\mathbf{a} = c_0 + c_1 s + \cdots + c_d s^d$.

   (c) $A \longleftarrow \begin{bmatrix} c_0 & & \\ \vdots & \ddots & \\ c_d & \vdots & c_0 \\ & \ddots & \vdots \\ & & c_d \end{bmatrix} \in \mathbb{K}^{(2d+1)\times n(d+1)}$

2. *Construct the "partial" reduced row-echelon form E of A.*

   This can be done by using Gauss-Jordan elimination (forward elimination, backward elimination, and normalization), with the following optimizations:

- *Skip over periodic non-pivot columns.*
- *Carry out the row operations only on the required columns.*

3. *Construct augmented matrix* $W = [A'_p \mid e_1] \in \mathbb{K}^{(2d+1-D)\times(2d+1-D+1)}$

   (a) $p \longleftarrow$ *the list of pivotal indices of* E

   (b) $D \longleftarrow 2d+1-|p|$

   (c) $W \longleftarrow [A'_p \mid e_1]$, *where* A'_p *is the submatrix of* $\begin{bmatrix} c_D & \cdots & c_0 & & \\ \vdots & \ddots & \vdots & \ddots & \\ \vdots & \vdots & c_D & & c_0 \\ c_d & \vdots & \vdots & \ddots & \vdots \\ & \ddots & \vdots & & c_D \\ & & c_d & & \vdots \\ & & & \ddots & \vdots \\ & & & & c_d \end{bmatrix}$ *indexed by* p.

4. *Construct the reduced row-echelon form* E' *of* W

   This can be done using regular Gauss-Jordan elimination.

5. *Construct a matrix* $P \in \mathbb{K}[s]^{n\times n}$ *whose first column is a Bézout vector of* $\mathbf{a}$ *of minimal degree and whose last* $n-1$ *columns form a* μ*-basis of* $\mathbf{a}$.

   Let p *be the list of the pivotal indices and let* $\tilde{q}$ *be the list of the basic non-pivotal indices of* E.

   (a) *Initialize an* $n \times n$ *matrix* P *with* 0 *in every entry.*

   (b) *For* $j = 2, \ldots, n$

   ```
           $r \leftarrow \mathrm{rem}(\tilde{q}_{j-1} - 1, n) + 1$
           $k \leftarrow \mathrm{quo}(\tilde{q}_{j-1} - 1, n)$
           $P_{r,j} \leftarrow P_{r,j} + s^k$
   ```

   (c) *For* $i = 1, \ldots, 2d+1-D$

   ```
           $r \leftarrow \mathrm{rem}(p_i - 1, n) + 1$
           $k \leftarrow \mathrm{quo}(p_i - 1, n)$
           $P_{r,1} \leftarrow P_{r,1} + E'_{i,nd+n+1} s^k$
           For $j = 2, \ldots, n$
               $P_{r,j} \leftarrow P_{r,j} - E_{i,\tilde{q}_{j-1}} s^k$
   ```

Example 105. *We trace the algorithm on the input vector*

$$\begin{aligned}\mathbf{a} &= \left[\begin{array}{ccc} 2+3s+s^2+s^4+s^5 & 3+3s+s^2+s^3+s^4+s^5 & 6+6s+2s^3+3s^4+s^5 \end{array}\right] \\ &= (1+s)\left[\begin{array}{ccc} 2+s+s^4 & 3+s^2+s^4 & 6+2s^3+s^4 \end{array}\right] \in \mathbb{Q}[s]^3\end{aligned}$$

1. *Construct matrix* $A \in \mathbb{K}^{(2d+1)\times n(d+1)}$
   - *(a)* $d \longleftarrow 5$
   - *(b)* $c_0, c_1, c_2, c_3, c_4, c_5 \longleftarrow [\,2\ 3\ 6\,], [\,3\ 3\ 6\,], [\,1\ 1\ 0\,], [\,0\ 1\ 2\,], [\,1\ 1\ 3\,], [\,1\ 1\ 1\,]$
   - *(c)*

$$A \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc|ccc}
2 & 3 & 6 & & & & & & & & & & & & & & & \\
3 & 3 & 6 & 2 & 3 & 6 & & & & & & & & & & & & \\
1 & 1 & 0 & 3 & 3 & 6 & 2 & 3 & 6 & & & & & & & & & \\
0 & 1 & 2 & 1 & 1 & 0 & 3 & 3 & 6 & 2 & 3 & 6 & & & & & & \\
1 & 1 & 3 & 0 & 1 & 2 & 1 & 1 & 0 & 3 & 3 & 6 & 2 & 3 & 6 & & & \\
1 & 1 & 1 & 1 & 1 & 3 & 0 & 1 & 2 & 1 & 1 & 0 & 3 & 3 & 6 & 2 & 3 & 6 \\
 & & & 1 & 1 & 1 & 1 & 1 & 3 & 0 & 1 & 2 & 1 & 1 & 0 & 3 & 3 & 6 \\
 & & & & & & 1 & 1 & 1 & 1 & 1 & 3 & 0 & 1 & 2 & 1 & 1 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 1 & 1 & 3 & 0 & 1 & 2 \\
 & & & & & & & & & & & & 1 & 1 & 1 & 1 & 1 & 3 \\
 & & & & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

2. *Construct the "partial" reduced row-echelon form* E *of* A.

$$E \longleftarrow \left[\begin{array}{ccc|ccc|ccc|ccc|ccc|ccc}
1 & & & & & & & -3 & -9 & & & & & & & & & \\
 & 1 & & & & & & -2 & -8 & & & & & & & & & \\
 & & 1 & & & & & 2 & 7 & & & & & & & & & \\
 & & & 1 & & & & 3 & 12 & & 3 & 6 & & & & & & \\
 & & & & 1 & & & -5 & -15 & & 3 & 6 & & 3 & 6 & & & \\
 & & & & & 1 & & 2 & 5 & & 1 & 0 & & 3 & 6 & & 3 & 6 \\
 & & & & & & 1 & 1 & 1 & & 1 & 2 & & 1 & 0 & & 3 & 6 \\
 & & & & & & & & & 1 & 1 & 3 & & 1 & 2 & & 1 & 0 \\
 & & & & & & & & & & 1 & 1 & 1 & 1 & 3 & & 1 & 2 \\
 & & & & & & & & & & & & & 1 & 1 & 1 & 1 & 3 \\
 & & & & & & & & & & & & & & & & 1 & 1
\end{array}\right]$$

   Here, blue denotes pivotal columns, red denotes basic non-pivotal columns, and brown denotes periodic non-pivotal columns.

3. *Construct augmented matrix* $W = [A'_p \mid e_1] \in \mathbb{K}^{(2d+1-D)\times(2d+1-D+1)}$
   - *(a)* $p \longleftarrow [1,2,3,4,5,6,7,10,13,16]$

(b) $D \longleftarrow 1$

(c) $$W \longleftarrow \left[\begin{array}{cccccccccc|c} 3 & 3 & 6 & 2 & 3 & 6 & & & & & 1 \\ 1 & 1 & 0 & 3 & 3 & 6 & 2 & & & & \\ 0 & 1 & 2 & 1 & 1 & 0 & 3 & 2 & & & \\ 1 & 1 & 3 & 0 & 1 & 2 & 1 & 3 & 2 & & \\ 1 & 1 & 1 & 1 & 1 & 3 & 0 & 1 & 3 & 2 & \\ & & & 1 & 1 & 1 & 1 & 0 & 1 & 3 & \\ & & & & & & 1 & 1 & 0 & 1 & \\ & & & & & & & 1 & 1 & 0 & \\ & & & & & & & & 1 & 1 & \\ & & & & & & & & & 1 & \end{array}\right]$$

4. *Construct the reduced row-echelon form E' of W.*

$$E' \longleftarrow \left[\begin{array}{cccccccccc|c} 1 & & & & & & & & & & 2 \\ & 1 & & & & & & & & & 1 \\ & & 1 & & & & & & & & -1 \\ & & & 1 & & & & & & & -1 \\ & & & & 1 & & & & & & 2 \\ & & & & & 1 & & & & & -1 \\ & & & & & & 1 & & & & 0 \\ & & & & & & & 1 & & & 0 \\ & & & & & & & & 1 & & 0 \\ & & & & & & & & & 1 & 0 \end{array}\right]$$

5. *Construct a matrix $P \in \mathbb{K}[s]^{n\times n}$ whose first column consists of a minimal-degree Bézout vector of $\mathbf{a}$ and whose last $n-1$ columns form a μ-basis of $\mathbf{a}$.*

(a) $$P \longleftarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(b) $$P \longleftarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & s^2 & 0 \\ 0 & 0 & s^2 \end{bmatrix}$$

(c) $$P \longleftarrow \begin{bmatrix} 2-s & 3-3\,s-s^2 & 9-12\,s-s^2 \\ 1+2\,s & 2+5\,s+s^2 & 8+15\,s \\ -1-s & -2-2\,s & -7-5\,s+s^2 \end{bmatrix}$$

Unsurprisingly, the matrix P computed in the example above is the same P computed in Example 99. As far as theoretical complexity goes, the complexity of Algorithm 104 is the same as that of the original OMF Algorithm 97. This is because the primary new step (step 4) in Algorithm 104 is a row-echelon reduction of a $(2d+1-D)\times(2d+1-D+1)$ matrix, and so the worst-case complexity of this step is $O(d^3)$. Adding this to the complexity obtained in Proposition 100 yields a complexity of $O(d^2n+d^3+n^2)$. Of course, we still expect Algorithm 104 to run slower than the original OMF algorithm 97 in real time due to the extra reduction step.

5.3 Equivariance

In Chapter 4, we justified the term moving frame by picturing it as a coordinate system moving along a curve. This point of view is reminiscent of classical geometric frames, such as the Frenet-Serret frame. However, the frames we've discussed here were defined by suitable algebraic properties, not its geometric properties. It is then natural to ask if it is possible to combine algebraic properties of Definition 62 with some essential geometric properties, in particular with the group-equivariance property. In this section, we show that any deterministic algorithm for computing an optimal moving frame can be augmented to obtain an algorithm that computes a $GL_n(\mathbb{K})$-equivariant moving frame.

The majority of frames in differential geometry have a group-equivariance property. For a curve in the three dimensional space, the Frenet frame is a classical example of a Euclidean group-equivariant frame. However, alternative geometric frames, in particular rotation minimizing frames, appear in applications in computer aided geometric design, geometric modeling, and computer graphics (see, for instance, [Gug89], [Wan08], [Far14], [Far16] and references therein). A method for deriving equivariant moving frames for higher-dimensional objects and for non-Euclidean geometries has been developed by Cartan (such as in [Car35]), who used moving frames to solve various group-equivalence problems (see [Gug63], [IL16], [Cle17] for modern introduction into Cartan's approach). The moving frame method was further developed and generalized by Griffiths [Gri74], Green [Gre78], Fels and Olver [FO99], and many others. Group-equivariant moving frames have a wide range of applications to problems in mathematics, science, and engineering (see [Olv15] for an overview).

The group-equivariance property is essential for the majority of frames arising in differential geometry. For the Frenet-Serret frame it is manifested as follows. We recall that for a smooth curve γ in $\mathbb{R}^3$, the Frenet-Serret frame at a point $p\in\gamma$ consists of the unit tangent vector T, the unit normal vector N and the unit binormal vector B to the curve at p. Consider the action of Euclidean group $E(3)$ (consisting of rotations, reflections, and translations) on $\mathbb{R}^3$. This action induces and action of the curves in $\mathbb{R}^3$ and on the vectors. It is easy to see that, for any $g\in E(3)$, the vectors $g\,T$, $g\,N$ and $g\,B$ are the unit tangent, the unit normal and the unit binormal, respectively, at the point $g\,p$ of the curve $g\gamma$. Thus, if we define $F_\gamma(p) = [T, N, P]$, then we can record the equivariance property as:

$$F_{g\gamma}(g\,p) = g\,F_\gamma(p) \quad \text{for all } \gamma \subset \mathbb{R}^3, p \in \gamma \text{ and } g \in E(3). \tag{5.3}$$

In the case of the algebraic moving frames we consider, we are interested in developing an algorithm that for $\mathbf{a} \in \mathbb{K}[s]^n \backslash \{\mathbf{0}\}$ produces an optimal moving frame $P_\mathbf{a}$ (recall Definition 62) with the additional $GL_n(\mathbb{K})$-equivariance property:

$$P_{\mathbf{a}g}(s) = g^{-1}\,P_\mathbf{a}(s) \text{ for all } \mathbf{a} \in \mathbb{K}[s]^n \backslash \{\mathbf{0}\}, s \in \mathbb{K} \text{ and } g \in GL_n(\mathbb{K}). \tag{5.4}$$

We observe that on the right-hand side of (5.3) the frame is multiplied by g, while on the right-hand side of (5.4) the frame is multiplied by g^{-1}. This means that the columns of P comprise a *right* equivariant moving frame, while the Frenet-Serret frame is a *left* moving frame (see Definition 3.1 in [FO99] and the subsequent discussion).

To give a precise definition of a $GL_n(\mathbb{K})$-right-equivariant algebraic moving frame algorithm, consider the set $M = \mathbb{K} \times (\mathbb{K}[s]^n \backslash \{\mathbf{0}\})$, and view an algorithm producing an algebraic moving frame as a map $\rho : M \to GL_n(\mathbb{K})$ such that, for a fixed $\mathbf{a}$, the matrix $P_\mathbf{a}(s) = \rho(s, \mathbf{a})$ is polynomial in s and satisfies Definition 62. Then the $GL_n(\mathbb{K})$-property (5.4) is equivalent to the commutativity of the following diagram:

$$\begin{tikzcd}
GL_n(\mathbb{K}) \arrow[r, "L_g^{-1}"] & GL_n(\mathbb{K}) \\
M \arrow[u, "\rho"] \arrow[r, "g"] & M \arrow[u, "\rho"]
\end{tikzcd}$$

On the top of the diagram, L_g^{-1} indicates the right action of $g \in GL_n(\mathbb{K})$ on $GL_n(\mathbb{K})$ defined by multiplication from the left by g^{-1}, while on the bottom the right action is defined by $g \cdot (s, \mathbf{a}) = (s, \mathbf{a}\,g)$.

We observe further that if the columns of P comprise a right equivariant moving frame, then the rows of P^{-1} comprise a left frame. The inverse algebraic frame has an easy geometric interpretation: the first row of $P_\mathbf{a}^{-1}$ equals to the position vector $\mathbf{a}$ and together with the last $n-1$ rows forms an n-dimensional parallelepiped whose volume does not change as the frame moves along the curve.

It is easy to find an instance of g and $\mathbf{a}$ to show that $P_\mathbf{a} = OMF(\mathbf{a})$, where $OMF(\mathbf{a})$ is produced by Algorithm 97, does not satisfy (5.4) and, therefore, the OMF algorithm is not a $GL_n(\mathbb{K})$-equivariant algorithm. However, for input vectors $\mathbf{a} = [a_1, \ldots, a_n]$ such that $a_1, \ldots, a_n$ are independent over $\mathbb{K}$, the OMF algorithm can be augmented into a $GL_n(\mathbb{K})$-equivariant algorithm as follows:

Algorithm 106 (**EOMF**)**.**

Input: $\mathbf{a} = [a_1, \dots, a_n] \neq 0 \in \mathbb{K}[s]^n$, *row vector, where* $n > 1$, $\gcd(\mathbf{a}) = 1$, $\mathbb{K}$ *a computable field, and components of* $\mathbf{a}$ *are linearly independent over* $\mathbb{K}$.

Output: $P \in \mathbb{K}[s]^{n\times n}$, *a degree-optimal moving frame at* $\mathbf{a}$

1. *Construct an* $n \times n$ *invertible submatrix of the coefficient matrix of* $\mathbf{a}$.

   (a) $d \longleftarrow \deg(\mathbf{a})$

   (b) *Identify the row vectors* $c_0 = [c_{01}, \dots c_{0n}], \dots, c_d = [c_{d1}, \dots c_{dn}]$ *such that* $\mathbf{a} = c_0 + c_1 s + \dots + c_d s^d$.

   (c) $I = [i_1, \dots, i_n] \longleftarrow$ *lexicographically smallest vector of integers between* 0 *and* d, *such that vectors* $c_{i_1}, \dots, c_{i_n}$ *are independent over* $\mathbb{K}$.

   (d) $\widehat{C} \longleftarrow \begin{bmatrix} c_{i_1} \\ \vdots \\ c_{i_n} \end{bmatrix}$

2. *Compute an optimal moving frame for a canonical representative of the* $GL_n(\mathbb{K})$*-orbit of* $\mathbf{a}$.

$$\widehat{P} \longleftarrow OMF(\mathbf{a}\widehat{C}^{-1})$$

3. *Revise the moving frame* $\widehat{P}$ *so that the algorithm has the equivariant property* (5.4).

$$P \longleftarrow \widehat{C}^{-1}\widehat{P}.$$

To prove the algorithm we need the following proposition.

Proposition 107. *Let* P *be a degree-optimal moving frame at a nonzero polynomial vector* $\mathbf{a}$. *Then, for any* $g \in GL_n(\mathbb{K})$, *the matrix* $g^{-1}P$ *is a degree-optimal moving frame at the vector* $\mathbf{a}g$.

Proof. By definition, $\mathbf{a}P = [\gcd(\mathbf{a}), 0, \dots, 0]$ and, therefore, for any $g \in GL_n(\mathbb{K})$ we have:

$$(\mathbf{a}g)\,g^{-1}P = [\gcd(\mathbf{a}), 0, \dots, 0].$$

From this, we conclude that $\gcd(\mathbf{a}g) = \gcd(\mathbf{a})$ and that $g^{-1}P$ is a moving frame at $\mathbf{a}g$. We note that the rows of the matrix $g^{-1}P$ are linear combinations over $\mathbb{K}$ of the rows of the matrix P. Therefore, the degrees of the columns of $g^{-1}P$ are less than or equal to the degrees of the corresponding columns of P.

Assume that $g^{-1}P$ is not a degree-optimal moving frame at $\mathbf{a}g$. Then there exists a moving frame P' at $\mathbf{a}g$ such that at least one of the columns of P', say the j-th column, has degree strictly less than the j-th column of $g^{-1}P$. Then, from the paragraph above, the j-th column of P' has degree strictly less than the degree of the j-th column of P.

By the same argument, gP' is a moving frame at $\mathbf{a}$ such that its j-th column has degree less than or equal to the degree of the j-th column of P', which is strictly less than the degree of the j-th column of P. This contradicts our assumption that P is degree-optimal. □

Proof of the Algorithm 106. We first note that, since polynomials $a_1, \ldots, a_n$ are linearly independent over $\mathbb{K}$, then the coefficient matrix C contains n independent rows and, therefore, Step 1 of the algorithm can be accomplished. Let $\widehat{\mathbf{a}} = \mathbf{a}\widehat{C}^{-1}$, then $\mathbf{a} = \widehat{\mathbf{a}}\widehat{C}$ and P is an optimal moving frame at $\mathbf{a}$ by Proposition 107. To show (5.4), for an arbitrary input $\mathbf{a}_1$ and an arbitrary $g \in GL_n(\mathbb{K})$, let $\mathbf{a}_2 = \mathbf{a}_1 g$. Then $\widehat{C_{\mathbf{a}_2}} = \widehat{C_{\mathbf{a}_1}}\, g$ and so

$$EOMF(\mathbf{a}_2) = \widehat{C_{\mathbf{a}_2}}^{-1} OMF(\mathbf{a}_2 \widehat{C_{\mathbf{a}_2}}^{-1}) = g^{-1}\widehat{C_{\mathbf{a}_1}}^{-1} OMF(\mathbf{a}_1 g\, g^{-1} \widehat{C_{\mathbf{a}_1}}^{-1}) = g^{-1}\, EOMF(\mathbf{a}_1).$$

□

Remark 108. *It is clear from the above proof that if, in Step 2 of Algorithm 106, we replace OMF with any (not necessarily degree-optimal) algorithm, then (not necessarily degree-optimal) frames produced by Algorithm 106 will have the $GL_n(\mathbb{K})$-equivariant property* (5.4).

5.4 Other approaches

In Section 4.1, we outlined an informal algorithm for producing a degree-optimal moving frame at $\mathbf{a}$. Namely, construct a minimal-degree Bézout vector of $\mathbf{a}$, construct a μ-basis of $\mathbf{a}$, and then combine them in a matrix. However, we are not aware of any algorithms for constructing a minimal-degree Bézout vector. Using the ideas presented in the proof of Theorem 75, it is possible to reduce the degree of a Bézout vector of $\mathbf{a}$ using a μ-basis of $\mathbf{a}$. However, such a process (construct Bézout vector, construct μ-basis, reduce) is inefficient, and it still does not guarantee that the Bézout vector after reduction will be minimal-degree. The advantage of the OMF algorithm is that not only does it construct a minimal-degree Bézout vector of $\mathbf{a}$ and a μ-basis of $\mathbf{a}$, but it does so simultaneously and with just a single "partial" Gauss-Jordan elimination.

Using Gröbner bases, it is possible to produce a degree-optimal moving frame, as follows:

```
Input: $\mathbf{a} \in \mathbb{K}[s]^n$, $\mathbf{a} \neq 0$
Output: $P$, a degree-optimal moving frame at $\mathbf{a}$

  1. $\mathbf{b} \longleftarrow$ a Bézout vector of $\mathbf{a}$.
  2. $\{\mathbf{g}_1, \dots, \mathbf{g}_{n-1}\} \longleftarrow$ a reduced Gröbner basis for syz($\mathbf{a}$) with respect to TOP ordering.
  3. $\mathbf{r} \longleftarrow$ the normal form of $\mathbf{b}$ with respect to $\{\mathbf{g}_1, \dots, \mathbf{g}_{n-1}\}$.
  4. $P \longleftarrow [\mathbf{r}, \mathbf{g}_1, \dots, \mathbf{g}_{n-1}]$.
  5. Return $P$.
```

The primary drawback to the above algorithm is that Gröbner basis computations tend to be quite costly. Moreover, the general process (construct Bézout vector, construct Gröbner basis, reduce) is inefficient. We would like to compare the OMF algorithm with other approaches that construct both components of a degree-optimal moving frame (minimal-degree Bézout vector and μ-basis) simultaneously. Unfortunately, we are not aware of any previously-published algorithm for such degree-optimal moving frames. Hence, we cannot compare the algorithm OMF with any existing algorithms. However, there are approaches for constructing both components of a not-necessarily-optimal moving frame (Bézout vector and pointwise independent basis of syzygies) simultaneously. So instead, we will compare with a not yet published, but tempting alternative approach. The approach consists of two steps: (1) Compute a moving frame. (2) Reduce the degree to obtain a degree-optimal moving frame. We elaborate on this *two-step* approach.

(1) *Compute a moving frame.* The problem of constructing an algebraic moving frame is a particular case of the well-known problem of providing a constructive proof of the Quillen-Suslin theorem [FG90], [LS92], [Can93], [PW95], [LY05], [FQ07]. In those papers, the multivariate problem is reduced inductively to the univariate case, and then an algorithm for the univariate case is proposed. Those univariate algorithms produce moving frames. As far as we are aware, the produced moving frames are *usually not* degree-optimal. However, the algorithms are very efficient. We discuss some of these algorithms in subsequent sections.

(2) *Reduce the degree to obtain a degree-optimal moving frame.* There are several different ways to carry out degree reduction: Popov form ([Bec99], [Bec06]), column reduced form [CL07] and matrix GCD [BL00]. As far as we are aware, the Popov form algorithm [Bec06] is the only one with a publicly available Maple implementation. Thus, we will use it for comparison. We explain how to use Popov form to reduce the degree.

(a) Compute the Popov normal form of the last $n-1$ columns of a non-optimal moving frame P.

(b) Reduce the degree of the first column of P (a Bézout vector) by the Popov normal form of the last $n-1$ columns.

5.5 Fabianska-Quadrat algorithm

We now discuss some algorithms for computing not-necessarily-optimal algebraic moving frames. We will start with one such algorithm used by Fabianska and Quadrat in [FQ07], because it has the least computational complexity among algorithms of which we are aware. Furthermore, the algorithm has been implemented by the authors in MAPLE, and the package can be obtained from `http://wwwb.math.rwth-aachen.de/QuillenSuslin/`. We discussed this algorithm in Section 2.2.3. For the readers' convenience, we outline their algorithm (for univariate case) again below.

Input: $\mathbf{a} \neq 0 \in \mathbb{K}[s]^n$ with $\gcd(\mathbf{a}) = 1$
Output: A moving frame at $\mathbf{a}$

1. Find constants $k_3, \dots, k_n$ such that $\gcd(a_1 + k_3 a_3 + \cdots + k_n a_n, a_2) = 1$.

2. Find $f_1, f_2 \in \mathbb{K}[s]$ such that $(a_1 + k_3 a_3 + \cdots + k_n a_n) f_1 + a_2 f_2 = 1$. This can be done by using the Extended Euclidean Algorithm.

3. $$P \longleftarrow \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ k_3 & & 1 & & \\ \vdots & & & \ddots & \\ k_n & & & & 1 \end{bmatrix} \begin{bmatrix} f_1 & -a_2 & & & \\ f_2 & a_1' & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -a_3 & \cdots & -a_n \\ 0 & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix},$$
   where $a_1' = a_1 + k_3 a_3 + \cdots + k_n a_n$.

The complexity of this algorithm is $O(d^2 + n^3)$, where d^2 comes from the Extended Euclidean Algorithm and n^3 comes from forming the matrix P, which is much better than the complexity of the OMF algorithm. We note, however, that the output of the Fabianska-Quadrat algorithm has degree at least d, while the output of the OMF algorithm has degree at most d and generically $\left\lceil \frac{d}{n-1} \right\rceil$. The same can be said about the Fitchas-Galligo algorithm, as discussed in Section 2.2.1.

We compared the computing times of the algorithm OMF and the alternative two-step approach as described in Section 5.4. Both algorithms are implemented in Maple (2016) and were executed on Apple iMac (Intel i 7-2600, 3.4 GHz, 16GB). The inputs polynomial vectors were generated as follows.

The coefficients were randomly taken from $[-10, 10]$. The degrees d of the polynomials ranged from 3 to 15. The length n of the vectors also ranged from 3 to 15.

Figure 5.1 shows the timings. The horizontal axes correspond to n and d and the vertical axis corresponds to computing time t in seconds. Each dot (d, n, t) represents an experimental timing. The red dots indicate the experimental timing of the algorithm OMF, while the blue dots indicate the experimental timing of the two-step approach described in Section 5.4.

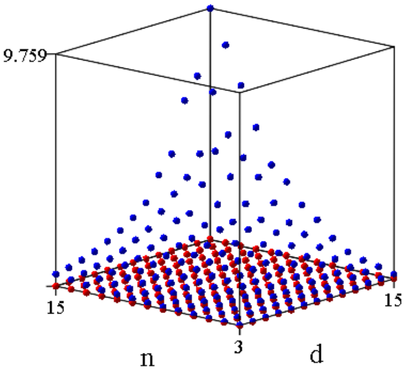

Figure 5.1 Timing comparison: OMF vs. Two-step approach

As can be seen, the OMF algorithm runs significantly more efficiently. This is due primarily to the cost of computing the Popov form of the last $n-1$ columns of the non-optimal moving frame. As described in [Bec06], the complexity of this step is $O(d^3n^7)$, which is bigger than $O(d^2n+d^3+n^2)$, the complexity of the OMF algorithm (Proposition 100). Although other algorithms and implementations for Popov form computations may be more efficient than the one currently implemented in Maple, we still expect OMF to significantly outperform any similar two-step procedure, because the degree reduction step is essentially similar to a TOP reduced Gröbner basis computation for a module, which is computationally expensive.

5.6 Algorithm based on generalized extended GCD

In Section 2.2.2, we highlighted another simple and elegant algorithm for constructing not-necessarily-optimal moving frames, based on a generalized version of Euclid's extended gcd algorithm. Such an algorithm has been mentioned in [LS92], [PW95], [PW98], and [Elk12]. For the readers' convenience, we describe this algorithm again in our notation (as a recursive program for simple presentation). We will refer to this algorithm as MF_GE (abbreviation of “Moving Frame by Generalized Euclid’s algorithm”).

```
Input:  $\mathbf{a} \in \mathbb{K}[s]^n$, $\mathbf{a} \neq 0$
Output: $P$, a moving frame at $\mathbf{a}$

  1. Let $k$ be such that $\mathbf{a} = \begin{bmatrix} a_1 & \cdots & a_k & 0 & \cdots & 0 \end{bmatrix}$ where $a_k \neq 0$.

  2. If $k = 1$ then set
        $P = \begin{bmatrix} \frac{1}{\mathrm{lc}(a_1)} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$
     and return $P$. (Here, $\mathrm{lc}(a_1)$ denotes the leading coefficient of $a_1$.)

  3. (Find $q_2, \ldots, q_k, r \in \mathbb{K}[s]^n$ such that $a_1 = q_2 a_2 + \ldots + q_k a_k + r$.)
     (a) $r \leftarrow a_1$
     (b) For $i = 2, \ldots, k$ do
             $q_i \leftarrow \mathrm{quo}(r, a_i)$
             $r \leftarrow \mathrm{rem}(r, a_i)$

  4. $\mathbf{a}' \leftarrow \begin{bmatrix} a_2 & \cdots & a_k & r & 0 & \cdots & 0 \end{bmatrix}$.

  5. $T \leftarrow \begin{bmatrix} & & & 1 & & & \\ 1 & & & -q_2 & & & \\ & \ddots & & \vdots & & & \\ & & 1 & -q_k & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \in \mathbb{K}[s]^{n \times n}$,
     where the $q$'s are placed in the $k$-th column

  6. $P' \leftarrow \mathrm{MF\_GE}(\mathbf{a}')$.

  7. $P \leftarrow T P'$.
```

8. Return P.

Remark 109. *We make a few remarks on the MF_GE algorithm.*

- *The MF_GE algorithm is elegant and very simple. The correctness is immediate from the fact that P is a product of uni-modular matrices (up to a sign).*

- *The MF_GE algorithm works even when* $\gcd(\mathbf{a}) \neq 1$, *unlike our algorithm OMF.*

- *For $n = 2$, the MF_GE algorithm always produces a degree-optimal moving frame. This is unsurprising, since when $n = 2$, the algorithm returns a matrix whose first column is the output of the standard extended Euclidean algorithm which, for two polynomials, is known to produce a minimal-degree Bézout vector. The second column of the output is the obvious lowest degree syzygy $\frac{1}{\gcd(\mathbf{a})}[-a_2, a_1]$.*

- *For $n > 2$, the MF_GE algorithm does not always produce a degree-optimal moving frame. For instance, for the example used in Section 5.1, $\mathbf{a} = \left[2+s+s^4, 3+s^2+s^4, 6+2s^3+s^4\right]$, the MF_GE output is*

$$\begin{bmatrix} \frac{1}{2}+\frac{1}{2}s+\frac{1}{2}s^2 & -\frac{3}{2}s+\frac{1}{2}s^2-\frac{1}{2}s^3 & \frac{3}{2}-\frac{3}{2}s-\frac{1}{2}s^2 \\ -\frac{1}{2}s-\frac{1}{2}s^2 & 2+s-\frac{1}{2}s^2+\frac{1}{2}s^3 & 1+\frac{5}{2}s+\frac{1}{2}s^2 \\ 0 & -1 & -1-s \end{bmatrix}$$

  and the OMF output is

$$\begin{bmatrix} 2-s & 3-3s-s^2 & 9-12s-s^2 \\ 1+2s & 2+5s+s^2 & 8+15s \\ -1-s & -2-2s & -7-5s+s^2 \end{bmatrix}.$$

  Observe that the degree of the Bézout vector column (i.e. the first column) for MF_GE is 2, while the degree of the Bézout vector column for OMF is 1. Likewise, the degrees of the syzygy columns for MF_GE are 2 and 3, while the degrees of the syzygy columns for OMF are 2 and 2.

5.7 Unimodular multipliers

In Sections 5.4 and 5.5, we discussed the concept of Popov form and how it can be used to reduce a non-optimal moving frame to a degree-optimal moving frame. The Popov form of a polynomial matrix was first introduced in [Pop70] and further developed in [Pop72], [Bec99], [Bec06]. In [Bec99], algorithms for computing the Popov form (polynomial echelon form) were described. This paper also included the notion of unimodular multipliers, i.e. unimodular polynomial matrices that bring a matrix to its Popov form. In the case of a single row vector $\mathbf{a} \in \mathbb{K}[s]^n$, the Popov form of $\mathbf{a}$ is $\left[\gcd(\mathbf{a}), 0, \dots, 0\right]$, and the multiplier is an algebraic moving frame. However, the multiplier is not necessarily degree optimal, as the following example shows.

Example 110. *Consider the input vector* $\mathbf{a} = \left[1+s^2, 1+s^2+s^3, 1+s^4\right] \in \mathbb{Q}[s]^3$. *The multiplier for* $\mathbf{a}$ *is*

$$\begin{bmatrix} -s^2-s+1 & s^2+2s-1 & -s^3-s^2-1 \\ s & -2s & s^2+1 \\ 0 & 1 & 0 \end{bmatrix}$$

and the OMF output is

$$\begin{bmatrix} s & s^2+2s-1 & 3s \\ -s & -2s & s^2-2s-1 \\ 1 & 1 & 1-s \end{bmatrix}.$$

Observe that the degree of the Bézout vector column (i.e. the first column) for the multiplier is 2, while the degree of the Bézout vector column for OMF is 1. Likewise, the degree of the syzygy columns for the multiplier are 2 and 3, while the degrees of the syzygy columns for OMF are 2 and 2.

Using additional column operations, a multiplier can be reduced to a degree-optimal moving frame (a “minimal multiplier" in the language of [Bec06]). The computations are similar to those performed when computing a TOP Gröbner basis, and as such this method is less efficient than methods discussed in previous sections. Faster algorithms for computing multipliers based on these ideas can be found in [Zho12], although a code is not available for comparison.

5.8 Using POT Gröbner basis computations

In Section 3.3, we presented an important relationship between μ-bases and Gröbner bases for the syzygy module of a polynomial vector. Since μ-bases form a key component of degree-optimal moving frames, it is thus reasonable to inquire whether Gröbner basis computations can be used to compute degree-optimal moving frames. We discussed an informal method to do so in Section 5.4, but this method does not produce both components of a moving frame simultaneously. In the next two sections, we show how to use Gröbner basis computations to compute moving frames simultaneously. The first algorithm is courtesy of David Cox and Thomas Sederberg. They work in the module $\mathbb{K}[s]^{n+1}$ with the standard basis $\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_n$, and they use POT Gröbner basis computations to construct a non-optimal moving frame. We present the algorithm here and prove its correctness. In the next section, we present a modified algorithm which will produce a degree-optimal moving frame.

Algorithm 111 (MF_POT)**.**

Input: $\mathbf{a} \in \mathbb{K}[s]^n$

Output: *a moving frame at* $\mathbf{a}$ *and the* gcd *of* $\mathbf{a}$.

1. $\mathbf{m}_i \longleftarrow a_i \mathbf{e}_0 + \mathbf{e}_i$ *for* $i = 1, \dots, n$.
2. *Compute the reduced Gröbner basis for the module* $M \subset \mathbb{K}[s]^{n+1}$ *generated by the* $\mathbf{m}_1, \dots, \mathbf{m}_n$ *with respect to* $\mathrm{POT}_{\mathbf{e}_0 > \cdots > \mathbf{e}_n}$ *ordering, obtaining*

$$\left[\begin{array}{c} \left[\begin{array}{c} g \\ \mathbf{b} \end{array}\right], \left[\begin{array}{c} 0 \\ \mathbf{g}_1 \end{array}\right], \left[\begin{array}{c} 0 \\ \mathbf{g}_2 \end{array}\right], \dots, \left[\begin{array}{c} 0 \\ \mathbf{g}_{n-1} \end{array}\right] \end{array}\right]$$

3. *Return* $[\mathbf{b}, \mathbf{g}_1, \dots, \mathbf{g}_{n-1}]$ *and* g.

Note that the algorithm works even when $\mathrm{gcd}(\mathbf{a}) \neq 1$ and can be used to compute $\mathrm{gcd}(\mathbf{a})$ as well. David Cox provided the proof that $g = \mathrm{gcd}(\mathbf{a})$, $\mathbf{b}$ is a Bézout vector of $\mathbf{a}$, and $\mathbf{g}_1, \dots, \mathbf{g}_{n-1}$ is a reduced $\mathrm{POT}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ Gróbner basis of syz($\mathbf{a}$). We provide the details here. We also add some explicit details to show that the output is a moving frame.

Consider $\mathbb{K}[s]^{n+1}$ as the free module with standard basis $\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_n$. Consider the submodule $M \subset \mathbb{K}[s]^{n+1}$ generated by $\mathbf{m}_1, \dots, \mathbf{m}_n$, where

$$\mathbf{m}_i = a_i \mathbf{e}_0 + \mathbf{e}_i = [a_i, 0, \dots, 0, 1, 0, \dots, 0]^T$$

It is easy to see that $\mathbf{m}_1, \dots, \mathbf{m}_n$ are linearly independent and thus they form a basis of M, making M free with rank n. Lemma 113 below shows that M contains an isomorphic copy of syz($\mathbf{a}$).

Lemma 112. $\left[\begin{array}{c} g \\ \mathbf{b} \end{array}\right] \in M$ *if and only if* $g = \mathbf{a}\mathbf{b}$.

Proof. Note $\left[\begin{array}{c} g \\ \mathbf{b} \end{array}\right] \in M \iff \left[\begin{array}{c} g \\ \mathbf{b} \end{array}\right] = \sum_{i=1}^{n} b_i \mathbf{m}_i \iff g = \mathbf{a}\mathbf{b}$. □

Lemma 113. $M \cap (\{0\} \times \mathbb{K}[s]^n) = \{0\} \times \mathrm{syz}(\mathbf{a})$.

Proof. From Lemma 112, we have $\begin{bmatrix} 0 \\ \mathbf{b} \end{bmatrix} \in M \iff 0 = \mathbf{ab}$. Thus the lemma follows. □

The following lemma provides a key result.

Lemma 114. *Let G be the reduced Gröbner basis for the module $M \subset \mathbb{K}[s]^{n+1}$ with respect to* $\mathrm{POT}_{\mathbf{e}_0 > \cdots > \mathbf{e}_n}$ *ordering. Then*

1. *G contains a vector of the form $\mathbf{w}_0 = \begin{bmatrix} \mathrm{gcd}(\mathbf{a}) \\ \mathbf{b} \end{bmatrix}$, such that $\mathrm{LT}\,\mathbf{w}_0 = s^{\deg \mathrm{gcd}(\mathbf{a})}\mathbf{e}_0$, and $\mathbf{b}$ is a Bézout vector of $\mathbf{a}$.*

2. *Apart from $\mathbf{w}_0$, there are $n-1$ remaining vectors in G, and they all have 0 in the first component.*

3. *If $\mathbf{w}_1 = \begin{bmatrix} 0 \\ \mathbf{g}_1 \end{bmatrix}, \ldots, \mathbf{w}_{n-1} = \begin{bmatrix} 0 \\ \mathbf{g}_{n-1} \end{bmatrix}$ are the $n-1$ vectors described in 2, then $\{\mathbf{g}_1, \ldots, \mathbf{g}_{n-1}\}$ is the reduced Gröbner basis for* $\mathrm{syz}(\mathbf{a})$ *with respect to* $\mathrm{POT}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ *ordering.*

Proof. Let $h = \mathrm{gcd}(\mathbf{a})$. Throughout the proof, we repeatedly use an observation that for any nonzero vector $\widehat{\mathbf{h}} = \begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix} \in \mathbb{K}[s]^{n+1}$ with a zero first component, $\mathrm{LT}\,\widehat{\mathbf{h}}$ relative to $\mathrm{POT}_{\mathbf{e}_0 > \mathbf{e}_1 > \cdots > \mathbf{e}_n}$ monomial ordering on $\mathbb{K}[s]^{n+1}$ equals to $\mathrm{LT}\,\mathbf{h}$ relative to $\mathrm{POT}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ monomial ordering on $\mathbb{K}[s]^n$. In this situation, we will simply write $\mathrm{LT}\,\widehat{\mathbf{h}} = \mathrm{LT}\,\mathbf{h}$, without explicitly mentioning that these vectors belong to two different modules. We will prove each claim one-by-one.

1. We split the first claim into two sub-claims:

   (a) *There exists $\mathbf{t} \in M$ such that $\mathrm{LT}\,\mathbf{t} = \begin{bmatrix} s^{\deg h} \\ \mathbf{0} \end{bmatrix}$.*
   Let $\mathbf{b}'$ be a Bézout vector of $\mathbf{a}$. Then $\mathbf{ab}' = h$, and so by Lemma 112

   $$\begin{bmatrix} h \\ \mathbf{b}' \end{bmatrix} \in M.$$

   With respect to $\mathrm{POT}_{\mathbf{e}_0 > \cdots > \mathbf{e}_n}$ ordering, we have

   $$\mathrm{LT}\begin{bmatrix} h \\ \mathbf{b}' \end{bmatrix} = \begin{bmatrix} s^{\deg h} \\ \mathbf{0} \end{bmatrix}$$

(b) *G contains a vector of the form* $\mathbf{w}_0 = \begin{bmatrix} h \\ \mathbf{b} \end{bmatrix}$*, such that* $\operatorname{LT}\mathbf{w}_0 = s^{\deg h}\,\mathbf{e}_0$*, and* $\mathbf{b}$ *is a Bézout vector of* $\mathbf{a}$*.*

The Gröbner basis G must contain a vector $\mathbf{w}_0$, whose leading term divides the leading term of vector $\mathbf{t} \in M$, described in part (a). Thus $\mathbf{w}_0 = \begin{bmatrix} g \\ \mathbf{b} \end{bmatrix}$, where $g \neq 0$, $\operatorname{LT}\mathbf{w}_0 = s^{\deg g}\,\mathbf{e}_0$ and $\deg g \leq \deg h$. On the other hand, by Lemma 112, we have $g = \mathbf{ab}$ and, therefore, since $\deg \mathbf{ab} \geq \deg h$, we have $\deg g = \deg \mathbf{ab} \geq \deg h$. Thus $\deg g = \deg h$. By uniqueness of gcd, we conclude $g = h$ and so $\mathbf{w}_0 = \begin{bmatrix} \gcd(\mathbf{a}) \\ \mathbf{b} \end{bmatrix}$. By Lemma 112, we have $\gcd(\mathbf{a}) = \mathbf{ab}$, and so $\mathbf{b}$ is a Bézout vector of $\mathbf{a}$.

2. The fact that, apart from $\mathbf{w}_0$, the reduced Gröbner basis G contains exactly $n-1$ vectors follows immediately from the fact that M is a free module of rank n.

   To show that the $n-1$ remaining vectors all have 0 in the first component, let $\mathbf{q} = \begin{bmatrix} q \\ \mathbf{v} \end{bmatrix} \in G \backslash \{\mathbf{w}_0\}$ and assume that $q \neq 0$. Then, by Lemma 112, we have $q = \mathbf{av}$. Then $\mathbf{q}$ contains a monomial $s^{\deg \mathbf{av}}\,\mathbf{e}_0$ with a nonzero coefficient. Since $\deg \mathbf{av} \geq \deg h$, this monomial is divisible by $\operatorname{LT}\mathbf{w}_0$. This is in contradiction with our assumption that G is the reduced Gröbner basis. Thus, $q = 0$.

3. Let $\mathbf{w}_1 = \begin{bmatrix} 0 \\ \mathbf{g}_1 \end{bmatrix}, \ldots, \mathbf{w}_{n-1} = \begin{bmatrix} 0 \\ \mathbf{g}_{n-1} \end{bmatrix}$ be the $n-1$ vectors in G whose first component is zero. We split the claim into three sub-claims.

   (a) $\mathbf{g}_1, \ldots, \mathbf{g}_{n-1}$ *generate* $\operatorname{syz}(\mathbf{a})$.

   By Lemma 113, for any nonzero $\mathbf{h} \in \operatorname{syz}(\mathbf{a})$, vector $\begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix} \in M$. Then $\begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix}$ is a linear combination over $\mathbb{K}[s]$ of vectors in G. From the previous two parts, we have that, apart from $\mathbf{w}_1, \cdots \mathbf{w}_{n-1}$, there is only one vector in G with nonzero first component, and thus $\begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix}$ has to be a linear combination, over $\mathbb{K}[s]$, of vectors $\mathbf{w}_1, \cdots \mathbf{w}_{n-1}$. Since, by Lemma 113, $\{\mathbf{g}_1, \ldots, \mathbf{g}_{n-1}\} \subset \operatorname{syz}(\mathbf{a})$, we conclude that $\operatorname{syz}(\mathbf{a}) = \langle \mathbf{g}_1, \ldots, \mathbf{g}_{n-1} \rangle$.

   (b) $\operatorname{LT}\mathbf{g}_1, \ldots, \operatorname{LT}\mathbf{g}_{n-1}$ *with respect to* $\operatorname{POT}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ *monomial ordering on* $\mathbb{K}[s]^n$*, generate the leading monomial ideal of* $\operatorname{syz}(\mathbf{a})$*.*

   We need to show that for an arbitrary nonzero $\mathbf{h} \in \operatorname{syz}(\mathbf{a})$, there exists $i \in 1, \ldots, n-1$ such that, relative to $\operatorname{POT}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ ordering, $\operatorname{LT}\mathbf{h}$ is divisible by $\operatorname{LT}\mathbf{g}_i$. Assume that $\operatorname{LT}\mathbf{h} =$

$c\,s^t\,\mathbf{e}_k$, for some constant $c \in \mathbb{K}$, non-negative integer t and integer $k \in \{1,\dots,n-1\}$. By Lemma 113, vector $\widehat{\mathbf{h}} = \begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix} \in M$. Relative to $\mathrm{POT}_{\mathbf{e}_0 > \mathbf{e}_1 > \dots > \mathbf{e}_n}$ monomial ordering on $\mathbb{K}[s]^{n+1}$, we have $\mathrm{LT}\,\widehat{\mathbf{h}} = c\,s^t\,\mathbf{e}_k$. Since G is a Gröbner basis for M, there exists an element of G whose leading term divides $\mathrm{LT}\,\widehat{\mathbf{h}}$. Since, by parts 1 and 2, $\mathbf{w}_1, \cdots \mathbf{w}_{n-1}$ are the only elements of G, whose leading terms can involve basis vector $\mathbf{e}_k$ with $k > 0$, there exists $i \in 1,\dots,n-1$ such that $\mathrm{LT}\,\mathbf{w}_i$ divides $\mathrm{LT}\,\widehat{\mathbf{h}}$. The argument is concluded by an observation that $\mathrm{LT}\,\mathbf{w}_i = \mathrm{LT}\,\mathbf{g}_i$ and $\mathrm{LT}\,\widehat{\mathbf{h}} = \mathrm{LT}\,\mathbf{h}$.

(c) $\mathbf{g}_1,\dots,\mathbf{g}_{n-1}$ *comprise a reduced Gröbner basis for* syz($\mathbf{a}$). Parts (a) and (b) above imply that, by definition, $\mathbf{g}_1,\dots,\mathbf{g}_{n-1}$ comprise a Gröbner basis of syz($\mathbf{a}$). Since $\mathrm{LT}\;\mathbf{g}_i = \mathrm{LT}\;\mathbf{w}_i$ and G is reduced, it follows that that for all $j \neq i$, none of the monomials in $\mathbf{g}_j$ is divisible by $\mathrm{LT}\;\mathbf{g}_i$. Thus $(\mathbf{g}_1,\dots,\mathbf{g}_{n-1})$ is a reduced Gröbner basis for syz($\mathbf{a}$).

□

We proceed with showing that the output of Algorithm 111 is a moving frame. This requires the following lemma which shows that $\mathbf{g}_1,\dots,\mathbf{g}_{n-1}$ are point-wise linearly independent over $\overline{\mathbb{K}}$.

Lemma 115. *Let* $\mathbf{a} \in \mathbb{K}[s]^n$*, and let* $\mathbf{g}_1,\dots,\mathbf{g}_{n-1}$ *be a reduced Gröbner basis for* syz($\mathbf{a}$) *with respect to* $\mathrm{POT}_{\mathbf{e}_1 > \dots > \mathbf{e}_n}$. *Then* $\mathbf{g}_1,\dots,\mathbf{g}_{n-1}$ *are point-wise linearly independent over* $\overline{\mathbb{K}}$.

Proof. Assume $\mathrm{LT}\,\mathbf{g}_1 < \cdots < \mathrm{LT}\,\mathbf{g}_{n-1}$. Since the $\mathbf{g}_j$ form a reduced $\mathrm{POT}_{e_1 > \dots > e_n}$ Gröbner basis, they have a triangular structure similar to the following

$$\begin{bmatrix} \\ \\ \\ * \\ * \end{bmatrix}, \begin{bmatrix} \\ \\ * \\ * \\ * \end{bmatrix}, \dots, \begin{bmatrix} * \\ * \\ * \\ * \\ * \end{bmatrix}.$$

Suppose there exists $s_0 \in \overline{\mathbb{K}}$ such that $\mathbf{g}_1(s_0),\dots,\mathbf{g}_{n-1}(s_0)$ are linearly dependent over $\overline{\mathbb{K}}$. Then there exist constants $\alpha_1,\dots,\alpha_{n-1} \in \overline{\mathbb{K}}$, not all zero, such that

$$\alpha_1\,\mathbf{g}_1(s_0) + \cdots + \alpha_{n-1}\,\mathbf{g}_{n-1}(s_0) = 0.$$

Let $i = \max\{j \mid \alpha_j \neq 0\}$ and let

$$\mathbf{h} = \alpha_1\,\mathbf{g}_1 + \cdots + \alpha_i\,\mathbf{g}_i.$$

Then, since the $\mathbf{g}_j$ are reduced with leading terms ordered as above and we are using $\mathrm{POT}_{\mathbf{e}_1>\cdots>\mathbf{e}_n}$ ordering, $\mathrm{LT}\,\mathbf{h} = \alpha_i\,\mathrm{LT}\,\mathbf{g}_i$. Moreover, $\mathbf{h} \in \mathrm{syz}(\mathbf{a})$ and is not identically zero, but $\mathbf{h}(s_0) = 0$. It follows that $\gcd(\mathbf{h}) \neq 1$ in $\overline{\mathbb{K}}[s]$ and, therefore, $\tilde{\mathbf{h}} = \frac{1}{\gcd(\mathbf{h})}\mathbf{h}$ belongs to $\mathrm{syz}(\mathbf{a})$ and has leading term with degree strictly less than $\mathrm{LT}\,\mathbf{h}$. Thus, there exists $\mathbf{g}_j$ such that $\mathrm{LT}\,\mathbf{g}_j$ divides $\mathrm{LT}\,\tilde{\mathbf{h}}$. However, by the triangular structure discussed above, the only vector whose leading term depends on the same basis vector as $\mathrm{LT}\,\tilde{\mathbf{h}}$ is $\mathbf{g}_i$, and $\mathrm{LT}\,\mathbf{g}_i$ has degree greater than $\mathrm{LT}\,\tilde{\mathbf{h}}$ because $\deg(\gcd(\mathbf{h})) > 0$. Hence, division of $\mathrm{LT}\,\tilde{\mathbf{h}}$ by one of the $\mathrm{LT}\,\mathbf{g}_j$ is not possible. Contradiction implies the $\mathbf{g}_j$ must be point-wise linearly independent. □

Proof of the correctness of Algorithm 111

In Lemma 114, it is shown that the first column of P is a Bézout vector of $\mathbf{a}$, $g = \gcd(\mathbf{a})$, and the last $n-1$ columns of P form a reduced Gröbner basis for $\mathrm{syz}(\mathbf{a})$ with respect to $\mathrm{POT}_{\mathbf{e}_1>\cdots>\mathbf{e}_n}$ ordering. Lemma 115 shows that these last $n-1$ columns are point-wise linearly independent syzygies, and thus P is a moving frame by Proposition 73. □

The following example illustrates this algorithm does not necessarily provide a degree-optimal moving frame:

Example 116. *Let* $\mathbf{a} = [2+s+s^4, 3+s^2+s^4, 6+2s^3+s^4] \in \mathbb{Q}[s]^3$. *Then the output of the Algorithm 111 is*

$$\begin{bmatrix} 0 & 0 & 15 \\ -4s^3-9s^2+4s+21 & s^4+2s^3+6 & 3s^3+8s^2+7s-12 \\ 4s^3+s^2-2s-8 & -s^4-s^2-3 & -3s^3-2s^2-6s+1 \end{bmatrix},$$

while the OMF output is

$$\begin{bmatrix} 2-s & 3-3s-s^2 & 9-12s-s^2 \\ 1+2s & 2+5s+s^2 & 8+15s \\ -1-s & -2-2s & -7-5s+s^2 \end{bmatrix}.$$

Observe that the Bézout vector (i.e. first column) of the MF_POT algorithm output has degree 3, compared to 1 for OMF. Likewise, the basis of $\mathrm{syz}(\mathbf{a})$ *in the MF_POT algorithm has degrees 4 and 3, compared to 2 and 2 for OMF.*

5.9 Using TOP Gröbner basis computations

In this section, we present a modification of the algorithm from Section 5.8 and prove that it outputs a *degree-optimal* moving frame. The difference is highlighted in red color.

Algorithm 117 (MF_TOP)**.**

Input: $\mathbf{a} \in \mathbb{K}[s]^n$.

Output: *a degree-optimal moving frame at* $\mathbf{a}$ *and the* gcd *of* $\mathbf{a}$.

1. $\mathbf{m}_i \longleftarrow s^d a_i \mathbf{e}_0 + \mathbf{e}_i$ *for* $i = 1, \ldots, n$, *where* d *is the degree of* $\mathbf{a}$.

2. *Compute the reduced Gröbner basis for the module* $M \subset \mathbb{K}[s]^{n+1}$ *generated by the* $\mathbf{m}_1, \ldots, \mathbf{m}_n$ *with respect to* $\mathrm{TOP}_{\mathbf{e}_0 > \cdots > \mathbf{e}_n}$ *ordering, obtaining*

$$\begin{bmatrix} g \\ \mathbf{b} \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{g}_1 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{g}_2 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ \mathbf{g}_{n-1} \end{bmatrix}$$

3. *Return* $[\mathbf{b}, \mathbf{g}_1, \ldots, \mathbf{g}_{n-1}]$ *and* $\frac{1}{s^d} g$.

Now we will prove the correctness of the above algorithm. We will use the following notations. Consider $\mathbb{K}[s]^{n+1}$ as the free module with standard basis $\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_n$. Let $d = \deg(\mathbf{a})$, and consider the submodule $M \subset \mathbb{K}[s]^{n+1}$ generated by $\mathbf{m}_1, \ldots, \mathbf{m}_n$, where

$$\mathbf{m}_i = s^d a_i \mathbf{e}_0 + \mathbf{e}_i = [s^d a_i, 0, \ldots, 0, 1, 0, \ldots, 0]^T$$

It is easy to see that $\mathbf{m}_1, \ldots, \mathbf{m}_n$ are linearly independent and thus they form a basis of M, making M free with rank n. Lemma 119 below shows that M contains an isomorphic copy of syz($\mathbf{a}$). This lemma is similar to Lemma 113. The difference is that each basis vector of the submodule M has s^d multiplied to the first position.

Lemma 118. $\begin{bmatrix} g \\ \mathbf{b} \end{bmatrix} \in M$ *if and only if* $g = s^d \mathbf{ab}$.

Proof. Note $\begin{bmatrix} g \\ \mathbf{b} \end{bmatrix} \in M \iff \begin{bmatrix} g \\ \mathbf{b} \end{bmatrix} = \sum_{i=1}^n b_i \mathbf{m}_i \iff g = s^d \mathbf{ab}$. □

Lemma 119. $M \cap (\{0\} \times \mathbb{K}[s]^n) = \{0\} \times \mathrm{syz}(\mathbf{a})$.

Proof. From Lemma 118, we have $\begin{bmatrix} 0 \\ \mathbf{b} \end{bmatrix} \in M \iff 0 = s^d \mathbf{ab} \iff 0 = \mathbf{ab}$. Thus the lemma follows. □

The next lemma plays a key role in the proof of the algorithm:

Lemma 120. *Let G be the reduced Gröbner basis for the module $M \subset \mathbb{K}[s]^{n+1}$ with respect to* $\mathrm{TOP}_{\mathbf{e}_0 > \cdots > \mathbf{e}_n}$ *ordering. Then*

1. *G contains a vector of the form* $\mathbf{w}_0 = \begin{bmatrix} s^d \gcd(\mathbf{a}) \\ \mathbf{b} \end{bmatrix}$*, such that* $\mathrm{LT}\,\mathbf{w}_0 = s^{d+\deg\gcd(\mathbf{a})}\,\mathbf{e}_0$.

2. *Apart from $\mathbf{w}_0$, there are $n-1$ remaining vectors in G, and they all have 0 in the first component.*

3. *If* $\mathbf{w}_1 = \begin{bmatrix} 0 \\ \mathbf{g}_1 \end{bmatrix}, \ldots, \mathbf{w}_{n-1} = \begin{bmatrix} 0 \\ \mathbf{g}_{n-1} \end{bmatrix}$ *are the $n-1$ vectors described in 2, then $\{\mathbf{g}_1, \ldots, \mathbf{g}_{n-1}\}$ is the reduced Gröbner basis for* $\mathrm{syz}(\mathbf{a})$ *with respect to* $\mathrm{TOP}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ *ordering.*

4. *Vector $\mathbf{b} \in \mathbb{K}[s]^n$, appearing in part 1, is a minimal-degree Bézout vector of $\mathbf{a}$.*

Proof. Let $h = \gcd(\mathbf{a})$. Throughout the proof we repeatedly use an observation that for any nonzero vector $\widehat{\mathbf{h}} = \begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix} \in \mathbb{K}[s]^{n+1}$ with a zero first component, $\mathrm{LT}\,\widehat{\mathbf{h}}$ relative to $\mathrm{TOP}_{\mathbf{e}_0 > \mathbf{e}_1 > \cdots > \mathbf{e}_n}$ monomial ordering on $\mathbb{K}[s]^{n+1}$ equals to $\mathrm{LT}\,\mathbf{h}$ relative to $\mathrm{TOP}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ monomial ordering on $\mathbb{K}[s]^n$. In this situation, we will simply write $\mathrm{LT}\,\widehat{\mathbf{h}} = \mathrm{LT}\,\mathbf{h}$, without explicitly mentioning that these vectors belong to two different modules. We will prove each claim one-by-one.

1. We split the first claim into two sub-claims:

   (a) *There exists $\mathbf{t} \in M$ such that* $\mathrm{LT}\,\mathbf{t} = \begin{bmatrix} s^{d+\deg h} \\ \mathbf{0} \end{bmatrix}$.

   Let $\mathbf{b}'$ be a Bézout vector of $\mathbf{a}$ with $\deg \mathbf{b}' < d$ (such $\mathbf{b}'$ exists by Theorem 75). Since $\mathbf{b}'$ is a Bézout vector of $\mathbf{a}$, we have $h = \mathbf{a}\mathbf{b}'$ and, in turn, $s^d h = s^d \mathbf{a}\mathbf{b}'$. Then, by Lemma 118:

   $$\begin{bmatrix} s^d h \\ \mathbf{b}' \end{bmatrix} \in M$$

   Since $\deg \mathbf{b}' < d$ and we are using TOP ordering, we have

   $$\mathrm{LT}\begin{bmatrix} s^d h \\ \mathbf{b}' \end{bmatrix} = \begin{bmatrix} s^{d+\deg h} \\ \mathbf{0} \end{bmatrix}$$

   (b) *G contains a vector of the form* $\mathbf{w}_0 = \begin{bmatrix} s^d h \\ \mathbf{b} \end{bmatrix}$*, such that* $\mathrm{LT}\,\mathbf{w}_0 = s^{d+\deg h}\,\mathbf{e}_0$. The Gröbner basis G must contain a vector $\mathbf{w}_0$, whose leading term divides the leading term of vector

$\mathbf{t} \in M$, described in part (a). Thus $\mathbf{w}_0 = \begin{bmatrix} g \\ \mathbf{b} \end{bmatrix}$, where $g \neq 0$, $\operatorname{LT}\mathbf{w}_0 = s^{\deg g}\mathbf{e}_0$ and $\deg g \leq d + \deg h$. On the other hand, by Lemma 118, we have $g = s^d \mathbf{ab}$ and, therefore, since $\deg \mathbf{ab} \geq \deg h$, we have $\deg g = d + \deg \mathbf{ab} \geq d + \deg h$. Thus $\deg g = d + \deg h$.

2. The fact that, apart from $\mathbf{w}_0$, the reduced Gröbner basis G contains exactly $n-1$ vectors follows immediately from the fact that M is a free module of rank n.

   To show that the $n-1$ remaining vectors all have 0 in the first component, let $\mathbf{q} = \begin{bmatrix} q \\ \mathbf{v} \end{bmatrix} \in G \backslash \{\mathbf{w}_0\}$ and assume that $q \neq 0$. Then, by Lemma 118, we have $q = s^d \mathbf{av}$. Then $\mathbf{q}$ contains a monomial $s^{d+\deg \mathbf{av}}\mathbf{e}_0$ with a nonzero coefficient. Since $\deg \mathbf{av} \geq \deg h$, this monomial is divisible by $\operatorname{LT}\mathbf{w}_0$. This is in contradiction with our assumption that G is the reduced Gröbner basis. Thus, $q = 0$.

3. Let $\mathbf{w}_1 = \begin{bmatrix} 0 \\ \mathbf{g}_1 \end{bmatrix}, \ldots, \mathbf{w}_{n-1} = \begin{bmatrix} 0 \\ \mathbf{g}_{n-1} \end{bmatrix}$ be the $n-1$ vectors in G whose first component is zero. We split the claim into three sub-claims.

   (a) $\mathbf{g}_1, \ldots, \mathbf{g}_{n-1}$ *generate* syz($\mathbf{a}$).

   By Lemma 119, for any nonzero $\mathbf{h} \in$ syz($\mathbf{a}$), vector $\begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix} \in M$. Then $\begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix}$ is a linear combination over $\mathbb{K}[s]$ of vectors in G. From the previous two parts, we have that, apart from $\mathbf{w}_1, \cdots \mathbf{w}_{n-1}$, there is only one vector in G with non-zero first component, and thus $\begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix}$ has to be a linear combination, over $\mathbb{K}[s]$, of vectors $\mathbf{w}_1, \cdots \mathbf{w}_{n-1}$. Since, by Lemma 119, $\{\mathbf{g}_1, \ldots, \mathbf{g}_{n-1}\} \subset$ syz($\mathbf{a}$), we conclude that syz($\mathbf{a}$) $= \langle \mathbf{g}_1, \ldots, \mathbf{g}_{n-1} \rangle$.

   (b) $\operatorname{LT}\mathbf{g}_1, \ldots, \operatorname{LT}\mathbf{g}_{n-1}$ *with respect to* $\mathrm{TOP}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ *monomial ordering on* $\mathbb{K}[s]^n$*, generate the leading monomial ideal of* syz($\mathbf{a}$).

   We need to show that for an arbitrary nonzero $\mathbf{h} \in$ syz($\mathbf{a}$), there exists $i \in 1, \ldots, n-1$ such that, relative to $\mathrm{TOP}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ ordering, $\operatorname{LT}\mathbf{h}$ is divisible by $\operatorname{LT}\mathbf{g}_i$. Assume that $\operatorname{LT}\mathbf{h} = c\, s^t\, \mathbf{e}_k$, for some constant $c \in \mathbb{K}$, non-negative integer t and integer $k \in \{1, \ldots, n-1\}$. By Lemma 119, vector $\widehat{\mathbf{h}} = \begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix} \in M$. Relative to $\mathrm{TOP}_{\mathbf{e}_0 > \mathbf{e}_1 > \cdots > \mathbf{e}_n}$ monomial ordering on $\mathbb{K}[s]^{n+1}$, we have $\operatorname{LT}\widehat{\mathbf{h}} = c\, s^t\, \mathbf{e}_k$. Since G is a Gröbner basis for M, there exists an element of G whose leading term divides $\operatorname{LT}\widehat{\mathbf{h}}$. Since, by parts 1 and 2, $\mathbf{w}_1, \cdots \mathbf{w}_{n-1}$ are the only elements of G, whose leading terms can involve basis vector $\mathbf{e}_k$ with $k > 0$, there exists

$i \in 1, \dots, n-1$ such that $\mathrm{LT}\,\mathbf{w}_i$ divides $\mathrm{LT}\,\widehat{\mathbf{h}}$. The argument is concluded by an observation that $\mathrm{LT}\,\mathbf{w}_i = \mathrm{LT}\,\mathbf{g}_i$ and $\mathrm{LT}\,\widehat{\mathbf{h}} = \mathrm{LT}\,\mathbf{h}$.

(c) $\mathbf{g}_1, \dots, \mathbf{g}_{n-1}$ *comprise a reduced Gröbner basis for* $\mathrm{syz}(\mathbf{a})$. Parts (a) and (b) above imply that, by definition, $\mathbf{g}_1, \dots, \mathbf{g}_{n-1}$ comprise a Gröbner basis of $\mathrm{syz}(\mathbf{a})$. Since $\mathrm{LT}\ \mathbf{g}_i = \mathrm{LT}\ \mathbf{w}_i$ and G is reduced, it follows that that for all $j \neq i$, none of the monomials in $\mathbf{g}_j$ is divisible by $\mathrm{LT}\ \mathbf{g}_i$. Thus $(\mathbf{g}_1, \dots, \mathbf{g}_{n-1})$ is a reduced Gröbner basis for $\mathrm{syz}(\mathbf{a})$.

4. Let $\mathbf{w}_0 = \begin{bmatrix} s^d h \\ \mathbf{b} \end{bmatrix}$ be the vector appearing in part 1. By Lemma 118, we have $s^d h = s^d \mathbf{ab}$. Thus, $h = \mathbf{ab}$, and so $\mathbf{b}$ is a Bézout vector of $\mathbf{a}$.

   To show that $\mathbf{b}$ is of minimal degree, assume that there exists another Bézout vector $\mathbf{b}'$ of $\mathbf{a}$ with $\deg(\mathbf{b}') < \deg(\mathbf{b})$. Then $\mathbf{h} = \mathbf{b} - \mathbf{b}' \in \mathrm{syz}(\mathbf{a})$. Since $\{\mathbf{g}_1, \dots, \mathbf{g}_{n-1}\}$ is a Gröbner basis for $\mathrm{syz}(\mathbf{a})$, there exists $i \in \{1, \dots, n-1\}$, such that $\mathrm{LT}\,\mathbf{g}_i$ divides $\mathrm{LT}\,\mathbf{h}$. Since $\deg \mathbf{b} < \deg \mathbf{b}'$, then $\mathrm{LT}\,\mathbf{h} = \mathrm{LT}\,\mathbf{b}$ relative to the $\mathrm{TOP}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ ordering. Recalling that $\mathrm{LT}\,\mathbf{w}_i = \mathrm{LT}\,\mathbf{g}_i$, it follows that for some $i \in \{1, \dots, n-1\}$, $\mathrm{LT}\,\mathbf{w}_i$ divides a nonzero monomial in $\mathbf{w}_0$, contradicting our assumption that G is a reduced basis.

□

Proof of the correctness of Algorithm 117 By Lemma 120, $\frac{1}{s^d} g = \gcd(\mathbf{a})$, the first column of P is a minimal-degree Bézout vector of $\mathbf{a}$, and the last $n-1$ columns of P form a reduced $\mathrm{TOP}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ Gröbner basis for $\mathrm{syz}(\mathbf{a})$. Theorem 49 implies that a reduced $\mathrm{TOP}_{\mathbf{e}_1 > \cdots > \mathbf{e}_n}$ Gröbner basis for $\mathrm{syz}(\mathbf{a})$ is a μ-basis. Then, by Theorem 74, P is a degree-optimal moving frame. □

Example 121. *Let* $\mathbf{a} = [2 + s + s^4, 3 + s^2 + s^4, 6 + 2s^3 + s^4] \in \mathbb{Q}[s]^3$. *Then the output of Algorithm 117 is*

$$\begin{bmatrix} -s+2 & 9s-6 & s^2+12s-9 \\ 2s+1 & s^2-10s-6 & -15s-8 \\ -s-1 & -s^2+3s+5 & -s^2+5s+7 \end{bmatrix},$$

while the OMF output is

$$\begin{bmatrix} 2-s & 3-3\,s-s^2 & 9-12\,s-s^2 \\ 1+2\,s & 2+5\,s+s^2 & 8+15\,s \\ -1-s & -2-2\,s & -7-5\,s+s^2 \end{bmatrix}.$$

We observe that the degrees for the Bézout vector and the basis of $\mathrm{syz}(\mathbf{a})$ *are the same for both algorithms.*

We compared the computing times of the algorithms OMF and MF_TOP. Both algorithms are implemented in Maple (2016) and were executed on Apple iMac (Intel i 7-2600, 3.4 GHz, 16GB). The inputs polynomial vectors were generated as follows. The coefficients were randomly taken from $[-10, 10]$. The degrees d of the polynomials ranged from 5 to 50. The length n of the vectors also ranged from 5 to 50.

Figure 5.2 shows the timings. The horizontal axes correspond to n and d and the vertical axis corresponds to computing time t in seconds. Each dot (d, n, t) represents an experimental timing. The red dots indicate the experimental timing of the algorithm OMF, while the blue dots indicate the experimental timing of the algorithm MF_TOP. As can be seen, the OMF algorithm runs significantly more efficiently.

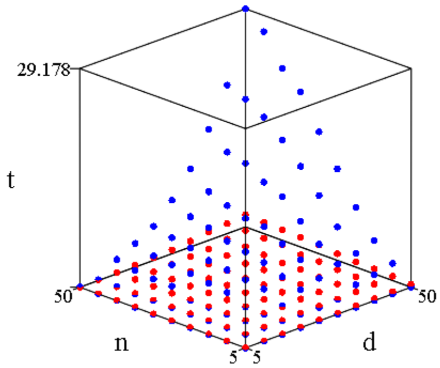

Figure 5.2 Timing comparison: OMF vs. MF_TOP

5.10 OMF via μ-basis algorithm

In the previous section, we computed a degree-optimal moving frame by computing a reduced TOP_E-Gröbner basis of the module M generated by $\mathbf{m}_i = s^d a_i \mathbf{e}_0 + \mathbf{e}_i$, $i = 1, \ldots, n$, where E refers to the standard basis. It is straightforward to show that the module M is the same as the syzygy module of the vector $[-1, s^d a_1, \ldots, s^d a_n]$.

Lemma 122. $M = \text{syz}([-1, s^d a_1, \ldots, s^d a_n])$

Proof. Let $\mathbf{a}' = [-1, s^d a_1, \ldots, s^d a_n]$. It is easy to see that $\mathbf{a}'\mathbf{m}_i = 0$ for all $i = 1, \ldots, n$. Thus $M \subset \text{syz}(\mathbf{a}')$.

Let $\mathbf{h} = [h_0, h_1, \dots, h_n]^T \in \operatorname{syz}(\mathbf{a}')$. Then

$$-h_0 + s^d a_1 h_1 + \cdots + s^d a_n h_n = 0,$$

from which it follows that $h_0 = s^d a_1 h_1 + \cdots + s^d a_n h_n$. Thus,

$$\mathbf{h} = \begin{bmatrix} s^d a_1 h_1 + \cdots + s^d a_n h_n \\ h_1 \\ \vdots \\ h_n \end{bmatrix} = h_1 \mathbf{m}_1 + \cdots + h_n \mathbf{m}_n.$$

Therefore, $\mathbf{h} \in M$ and so $\operatorname{syz}(\mathbf{a}') \subset M$. It follows that $M = \operatorname{syz}(\mathbf{a}')$. □

By this lemma and the results from the previous section, it follows that a degree-optimal moving frame can be computed by computing a reduced TOP_E-Gröbner basis for the syzygy module of the vector $[-1, s^d a_1, \dots, s^d a_n]$. Recall from Proposition 53 that the HHK μ-basis algorithm presented in Section 3.1.4 can be used to compute reduced TOP_E-Gröbner bases for syzygy modules. We thus immediately have the following.

Algorithm 123 (OMF_mu)**.**

Input: $\mathbf{a} \in \mathbb{K}[s]^n$.

Output: *a degree-optimal moving frame at* $\mathbf{a}$ *and the* gcd *of* $\mathbf{a}$.

1. $\mathbf{a}' \longleftarrow [-1, s^d \mathbf{a}] \in \mathbb{K}[s]^{n+1}$, *where* d *is the degree of* $\mathbf{a}$.
2. *Run the HHK* μ*-basis algorithm on* $\mathbf{a}'$, *obtaining*

$$\begin{bmatrix} g & 0 & \cdots & 0 \\ \mathbf{b} & \mathbf{u}_1 & \cdots & \mathbf{u}_{n-1} \end{bmatrix}$$

3. *Return* $[\mathbf{b}, \mathbf{u}_1, \dots, \mathbf{u}_{n-1}]$ *and* $\frac{1}{s^d} g$.

5.11 Matrix inputs

The algorithm presented in Section 5.1 has a natural generalization to unimodular matrix inputs **A**. In the matrix case, partial row echelon reduction is performed on the matrix obtained by stacking together Sylvester-type matrices corresponding to each row of $\mathbf{A}$. We now provide the details of this generalization.

Consider the matrix

$$\mathbf{A} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ \vdots & \vdots & \vdots \\ - & \mathbf{a}_m & - \end{bmatrix} \in \mathbb{K}[s]^{m\times n}$$

of degree d, where $m < n$ and $\mathbf{A}$ is unimodular. This means that the gcd of the $m \times m$ minors of $\mathbf{A}$ is 1. It is easy to show that this implies $\mathbf{A}$ has rank m and $\gcd(\mathbf{a}_i) = 1$ for all $i = 1, \ldots, m$. In fact, unimodular implies that $\mathbf{A}$ has point-wise rank m i.e. $\text{rank}(\mathbf{A}(s_0)) = m$ for all $s_0 \in \overline{\mathbb{K}}$. We wish to find a matrix $P \in \mathbb{K}[s]^{n\times n}$ of degree at most md such that

$$\mathbf{A}P = \begin{bmatrix} I_{m\times m} & 0_{m\times(n-m)} \end{bmatrix}. \tag{5.5}$$

Just as in Section 3.2, we do this by considering appropriate solutions to matrix equations involving the stacked matrix

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} \in \mathbb{K}^{m(md+d+1)\times n(md+1)},$$

where each $A_i \in \mathbb{K}^{(md+d+1)\times n(md+1)}$ is the block coefficient matrix corresponding to $\mathbf{a}_i$.

Clearly, the last $n - m$ columns of P are in the kernel of $\mathbf{A}$, and they can be computed using the minimal basis algorithm presented in Section 3.2. For $i = 1, \ldots, m$, the i-th column of P satisfies $\mathbf{a}_i P_{*i} = 1$ and $\mathbf{a}_j P_{*i} = 0$ for $j \neq i$. In terms of the matrix A and the flat isomorphism (recall Sections 3.1.2 and 4.2), this means $P_{*i} = v_i^\flat$, where $v_i \in \mathbb{K}^{n(md+1)}$ satisfies $A_i v_i = e_1$ and $A_j v_i = 0$ for $j \neq i$. Thus, for the first m columns of P, we are solving a linear system of equations with augmented matrix

$$W = \left[\begin{array}{ccc|cccc}
c_0^1 & & & 1 & & & \\
\vdots & \ddots & & & & & \\
c_d^1 & \vdots & c_0^1 & & & & \\
 & \ddots & \vdots & & & & \\
 & & c_d^1 & & & & \\
c_0^2 & & & & 1 & & \\
\vdots & \ddots & & & & & \\
c_d^2 & \vdots & c_0^2 & & & & \\
 & \ddots & \vdots & & & & \\
 & & c_d^2 & & & \ddots & \\
\vdots & \vdots & \vdots & & & & \\
\vdots & \vdots & \vdots & & & & \\
\vdots & \vdots & \vdots & & & & \\
c_0^m & & & & & & 1 \\
\vdots & \ddots & & & & & \\
c_d^m & \vdots & c_0^m & & & & \\
 & \ddots & \vdots & & & & \\
 & & c_d^m & & & &
\end{array}\right] \in \mathbb{K}^{m(md+d+1)\times(nmd+n+m)}.$$

All of the results in Sections 4.1 and 4.2 can now be readily adapted to show that a minimal-degree solution P to (5.5) can be constructed by computing the partial reduced row-echelon form $[E \mid \hat{v}_1, \dots, \hat{v}_m]$ of W. The first m columns of P are formed from $\hat{v}_1, \dots, \hat{v}_m$ and the last $n-m$ columns of P are formed from the basic non-pivotal columns of E. We thus have the following algorithm.

Algorithm 124.

```
Input: $\mathbf{A} \neq 0 \in \mathbb{K}[s]^{m\times n}$, unimodular where $n > 1$, $m < n$, and $\mathbb{K}$ a computable field

Output: $P \in \mathbb{K}[s]^{n\times n}$, unimodular, such that $\mathbf{A}P = \left[\begin{array}{cc} I_{m\times m} & 0_{m\times(n-m)} \end{array}\right]$

  1. Construct stacked augmented matrix $W \in \mathbb{K}^{m(md+d+1)\times(nmd+n+m)}$.

      (a) $d \longleftarrow \deg(\mathbf{a})$
```

(b) *For each* $i = 1, \ldots, m$, *identify the row vectors* $c_0^i, \ldots, c_d^i \in \mathbb{K}^n$ *such that* $\mathbf{a}_i = c_0^i + c_1^i s + \cdots + c_d^i s^d$.

(c) $$W \longleftarrow \left[\begin{array}{ccc|cccc}
c_0^1 & & & 1 & & & \\
\vdots & \ddots & & & & & \\
c_d^1 & \vdots & c_0^1 & & & & \\
 & \ddots & \vdots & & & & \\
 & & c_d^1 & & & & \\
c_0^2 & & & & 1 & & \\
\vdots & \ddots & & & & & \\
c_d^2 & \vdots & c_0^2 & & & & \\
 & \ddots & \vdots & & & & \\
 & & c_d^2 & & & \ddots & \\
\vdots & \vdots & \vdots & & & & \\
\vdots & \vdots & \vdots & & & & \\
\vdots & \vdots & \vdots & & & & \\
c_0^m & & & & & & 1 \\
\vdots & \ddots & & & & & \\
c_d^m & \vdots & c_0^m & & & & \\
 & \ddots & \vdots & & & & \\
 & & c_d^m & & & &
\end{array}\right] \in \mathbb{K}^{m(md+d+1)\times(nmd+n+m)}$$

2. *Construct the "partial" reduced row-echelon form* E *of* W.

   This can be done by using Gauss-Jordan elimination (forward elimination, backward elimination, and normalization), with the following optimizations:

   - *Skip over periodic non-pivot columns.*
   - *Carry out the row operations only on the required columns.*

3. *Construct a matrix* $P \in \mathbb{K}[s]^{n\times n}$, *unimodular, of minimal-degree that satisfies* (5.5).

   Let p *be the list of the pivotal indices and let* $\tilde{q}$ *be the list of the basic non-pivotal indices of* E.

   (a) *Initialize an* $n \times n$ *matrix* P *with* 0 *in every entry.*

```
(b) For $j = m+1,\dots,n$
        $r \leftarrow \mathrm{rem}\big(\tilde{q}_{j-1} - 1, n\big) + 1$
        $k \leftarrow \mathrm{quo}\big(\tilde{q}_{j-1} - 1, n\big)$
        $P_{r,j} \leftarrow P_{r,j} + s^k$
(c) For $i = 1,\dots,|p|$
        $r \leftarrow \mathrm{rem}\big(p_i - 1, n\big) + 1$
        $k \leftarrow \mathrm{quo}\big(p_i - 1, n\big)$
        For $l = 1,\dots,m$
            $P_{r,l} \leftarrow P_{r,l} + E_{i,nmd+n+l}\, s^k$
        For $j = m+1,\dots,n$
            $P_{r,j} \leftarrow P_{r,j} - E_{i,\tilde{q}_{j-1}}\, s^k$
```

Example 125. *We trace the algorithm on the input matrix*

$$\mathbf{A} = \begin{bmatrix} s^3+s+4 & s^3+s^2+3 & s^3+4 & s^3+s^2+s+4 \\ s^3+s+3 & s^3+s^2+2 & s^3+3 & s^3+6 \end{bmatrix} \in \mathbb{Q}^{2\times 4}$$

1. *Construct stacked augmented matrix* $W \in \mathbb{K}^{m(md+d+1)\times(nmd+n+m)}$.

   (a) $d \longleftarrow 3$

   (b) $c_0^1, c_1^1, c_2^1, c_3^1 \longleftarrow \begin{bmatrix} 4 & 3 & 4 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$
   $c_0^2, c_1^2, c_2^2, c_3^2 \longleftarrow \begin{bmatrix} 3 & 2 & 3 & 6 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$

   (c) $W \longleftarrow$

$$\left[\begin{array}{cccc|cccc|cccc|cccc|cccc|cccc|cccc|cc}
4&3&4&4 & &&& & &&& & &&& & &&& & &&& & &&& & 1 & \\
1&0&0&1 & 4&3&4&4 & &&& & &&& & &&& & &&& & &&& & & \\
0&1&0&1 & 1&0&0&1 & 4&3&4&4 & &&& & &&& & &&& & &&& & & \\
1&1&1&1 & 0&1&0&1 & 1&0&0&1 & 4&3&4&4 & &&& & &&& & &&& & & \\
&&& & 1&1&1&1 & 0&1&0&1 & 1&0&0&1 & 4&3&4&4 & &&& & &&& & & \\
&&& & &&& & 1&1&1&1 & 0&1&0&1 & 1&0&0&1 & 4&3&4&4 & &&& & & \\
&&& & &&& & &&& & 1&1&1&1 & 0&1&0&1 & 1&0&0&1 & 4&3&4&4 & & \\
&&& & &&& & &&& & &&& & 1&1&1&1 & 0&1&0&1 & 1&0&0&1 & & \\
&&& & &&& & &&& & &&& & &&& & 1&1&1&1 & 0&1&0&1 & & \\
&&& & &&& & &&& & &&& & &&& & &&& & 1&1&1&1 & & \\
3&2&3&6 & &&& & &&& & &&& & &&& & &&& & &&& & & 1 \\
1&0&0&0 & 3&2&3&6 & &&& & &&& & &&& & &&& & &&& & & \\
0&1&0&0 & 1&0&0&0 & 3&2&3&6 & &&& & &&& & &&& & &&& & & \\
1&1&1&1 & 0&1&0&0 & 1&0&0&0 & 3&2&3&6 & &&& & &&& & &&& & & \\
&&& & 1&1&1&1 & 0&1&0&0 & 1&0&0&0 & 3&2&3&6 & &&& & &&& & & \\
&&& & &&& & 1&1&1&1 & 0&1&0&0 & 1&0&0&0 & 3&2&3&6 & &&& & & \\
&&& & &&& & &&& & 1&1&1&1 & 0&1&0&0 & 1&0&0&0 & 3&2&3&6 & & \\
&&& & &&& & &&& & &&& & 1&1&1&1 & 0&1&0&0 & 1&0&0&0 & & \\
&&& & &&& & &&& & &&& & &&& & 1&1&1&1 & 0&1&0&0 & & \\
&&& & &&& & &&& & &&& & &&& & &&& & 1&1&1&1 & &
\end{array}\right]$$

2. *Construct the “partial" reduced row-echelon form E of A*

   $E \longleftarrow$

$$\left[\begin{array}{cccc|cccc|cccc|cccc|cccc|cccc|cccc|cc}
1 & & & & & & & & & & -1 & & & -4 & & & & & & & & & & & & & & & -1 & 1 \\
 & 1 & & & & & & & & & 0 & & & -12 & & & & & & & & & & & & & & & 3 & -4 \\
 & & 1 & & & & & & & & 1 & & & 14 & & & & & & & & & & & & & & & -1 & 2 \\
 & & & 1 & & & & & & & 0 & & & -1 & 4 & & & & & & & & & & & & & & 0 & 0 \\
 & & & & 1 & & & & & & 0 & & & 7 & 0 & & & 3 & 4 & & & & & & & & & & -3 & 4 \\
 & & & & & 1 & & & & & -1 & & & -1 & 0 & & & 0 & 0 & & & 3 & 4 & & & & & & -1 & 1 \\
 & & & & & & 1 & & & & 1 & & & -5 & 1 & & & 1 & 0 & & & 0 & 0 & & & 3 & 4 & & 4 & -5 \\
 & & & & & & & 1 & & & 0 & & & 0 & & & & 1 & 1 & & & 1 & 0 & & & 0 & 0 & & & \\
 & & & & & & & & 1 & & 1 & & & 3 & & & & & & & & 1 & 1 & & & 1 & 0 & & & \\
 & & & & & & & & & 1 & 0 & & & -2 & & & & & & & & & & & & 1 & 1 & & & \\
 & & & & & & & & & & & 1 & & 0 & & & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & 1 & 1 & & & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & & 1 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & 3 & & 1 & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & 0 & & & 2 & 3 & 1 & & & & & & & & & & \\
 & & & & & & & & & & & & & & 0 & & & 0 & 0 & & 1 & 2 & 3 & & & & & & & \\
 & & & & & & & & & & & & & & 1 & & & 1 & 0 & & & 0 & 0 & 1 & & 2 & 3 & & & \\
 & & & & & & & & & & & & & & & & & 1 & 1 & & & 1 & 0 & & 1 & 0 & 0 & & & \\
 & & & & & & & & & & & & & & & & & & & & & 1 & 1 & & & 1 & 0 & 1 & & \\
 & & & & & & & & & & & & & & & & & & & & & & & & & 1 & 1 & & &
\end{array}\right]$$

Here, blue denotes pivotal columns, red denotes basic non-pivotal columns, brown denotes periodic non-pivotal columns, and gray denotes the augmented columns.

3. *Construct matrix $P \in \mathbb{K}[s]^{n\times n}$, unimodular, of minimal-degree that satisfies* (5.5).

   (a) $P \longleftarrow \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

   (b) $P \longleftarrow \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & s^3 \\ 0 & 0 & s^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

   (c) $P \longleftarrow \begin{bmatrix} -1-3s & 1+4s & 1-s^2 & 4-7s-3s^2-s^3 \\ 3-s & -4+s & s & 12+s+2s^2+s^3 \\ -1+4s & 2-5s & -1-s+s^2 & -14+5s \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Proposition 126. *Let* $\mathbf{A} \in \mathbb{K}[s]^{m\times n}$ *be unimodular with* $\deg(\mathbf{A}) = d$. *Then for every minimal-degree unimodular solution* P *to* $\mathbf{A}P = [I_m, 0_{n-m}]$*, we have* $\lceil \frac{d}{n-1} \rceil \leq \deg(P) \leq md$*, and these degree bounds are sharp. By sharp, we mean that for all* $n > m \geq 1$ *and* $d > 0$*, there exists a unimodular* $\mathbf{A} \in \mathbb{K}[s]^{m\times n}$ *with* $\deg(\mathbf{A}) = d$ *such that, for every minimal-degree unimodular solution* P *to* $\mathbf{A}P = [I_m, 0_{n-m}]$*, we have* $\deg(P) = \lceil \frac{d}{n-1} \rceil$*. Likewise, for all* $n > m \geq 1$ *and* $d > 0$*, there exists a unimodular* $\mathbf{A} \in \mathbb{K}[s]^{m\times n}$ *with* $\deg(\mathbf{A}) = d$ *such that, for every minimal-degree unimodular solution* P *to* $\mathbf{A}P = [I_m, 0_{n-m}]$*, we have* $\deg(P) = md$.

Proof.

1. (lower bound): Let P be a minimal-degree unimodular solution to $\mathbf{A}P = [I_m, 0_{n-m}]$. Then for $i = 1, \ldots, m$ the i-th row of $\mathbf{A}$ satisfies

   $$\mathbf{A}_{i*}P = e_i$$

   and so from Cramer's rule:

   $$\mathbf{A}_{ij} = \frac{(-1)^{j+1}}{|P|} \left|P_{j,i}\right| \quad j = 1, \ldots n,$$

   where $P_{j,i}$ denotes the submatrix of P obtained by removing the i-th column and the j-th row. We remind the reader that $|P|$ is a nonzero constant. Assume for the sake of contradiction that $\deg(P) < \lceil \frac{d}{n-1} \rceil$. Then $\deg(P) < \frac{d}{n-1}$. Since $\left|P_{j,i}\right|$ is the determinant of an $(n-1)\times(n-1)$ submatrix of P, we have $\deg\left(\mathbf{A}_{ij}\right) = \deg\left(\left|P_{j,i}\right|\right) < (n-1)\frac{d}{n-1} = d$ for all $i = 1\ldots, m$ and $j = 1, \ldots, n$. This contradicts the assumption that $\deg(\mathbf{A}) = d$. Thus, $\deg(P) \geq \lceil \frac{d}{n-1} \rceil$.

   We will prove that the lower bound $\lceil \frac{d}{n-1} \rceil$ is sharp by showing that, for all $n > m \geq 1$ and $d > 0$, the following matrix

   $$P = \left[\begin{array}{ccc|ccccc} 1 & & & -s^{d-k\lceil \frac{d}{n-1} \rceil} & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ \hline & & & 1 & -s^{\lceil \frac{d}{n-1} \rceil} & & & \\ & & & & 1 & -s^{\lceil \frac{d}{n-1} \rceil} & & \\ & & & & & \ddots & \ddots & \\ & & & & & & \ddots & -s^{\lceil \frac{d}{n-1} \rceil} \\ & & & & & & & 1 \end{array}\right]$$

   has degree $\lceil \frac{d}{n-1} \rceil$ and is a minimal-degree unimodular solution to $\mathbf{A}P = [I_m, 0_{n-m}]$ for the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & & & & s^{d-k\cdot\lceil\frac{d}{n-1}\rceil} & s^{d-(k-1)\cdot\lceil\frac{d}{n-1}\rceil} & \dots & s^{d-2\cdot\lceil\frac{d}{n-1}\rceil} & s^{d-1\cdot\lceil\frac{d}{n-1}\rceil} & s^{d-0\cdot\lceil\frac{d}{n-1}\rceil} \\ & 1 & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & 1 & & & & & & \\ & & & & 1 & s^{d-k\cdot\lceil\frac{d}{n-1}\rceil} & \dots & \dots & s^{d-2\cdot\lceil\frac{d}{n-1}\rceil} & s^{d-1\cdot\lceil\frac{d}{n-1}\rceil} \\ & & & & & \ddots & \ddots & \dots & \dots & \vdots \\ & & & & & & 1 & s^{d-k\cdot\lceil\frac{d}{n-1}\rceil} & \dots & s^{d-(m-1)\cdot\lceil\frac{d}{n-1}\rceil} \end{bmatrix}$$

Here,

- $k \in \mathbb{N}$ is the maximal such that $d > k\left\lceil\frac{d}{n-1}\right\rceil$ (explicitly $k = \left\lceil\frac{d}{\lceil\frac{d}{n-1}\rceil}\right\rceil - 1$)
- the number of zeros in the first row of **A** is $n-k-2$
- the number of rows of **A** containing just a single 1 is $\min\{n-k-2, m-1\}$, and there are an appropriate number of remaining rows (so that the total number of rows is m), explicitly $\max\{m-(n-k-2)-1, 0\}$
- the upper-left block of P is of the size $(n-k-1)\times(n-k-1)$, the lower-right block is of the size $(k+1)\times(k+1)$, and the other two blocks are of the appropriate sizes.

The proof now proceeds as in the first part of Proposition 91 to show that such P and **A** always exist, P is a minimal-degree unimodular solution to $\mathbf{A}P = [I_m, 0_{n-m}]$, and $\deg(P) = \left\lceil\frac{d}{n-1}\right\rceil$.

2. (upper bound): Clearly, md is an upper bound on matrix P. We will prove that the upper bound md is sharp by showing that, for all $n > m \geq 1$ and $d > 0$, the following matrix

$$P = \begin{bmatrix} 1 & -s^d & s^{2d} & \cdots & (-1)^{m-1}s^{(m-1)d} & 0 & \cdots & 0 & (-1)^m s^{md} \\ & 1 & -s^d & \cdots & (-1)^{m-2}s^{(m-2)d} & 0 & \cdots & 0 & (-1)^{m-1}s^{(m-1)d} \\ & & \ddots & \ddots & \vdots & & & & \vdots \\ & & & \ddots & \vdots & & & & \vdots \\ & & & & 1 & & & & -s^d \\ & & & & & & & & 1 \\ & & & & & 1 & & & \\ & & & & & & \ddots & & \\ & & & & & & & 1 & \end{bmatrix}$$

has degree md and is a minimal-degree unimodular solution to $\mathbf{A}P = [I_m, 0_{n-m}]$ for the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & s^d & & & \\ & \ddots & \ddots & & \\ & & 1 & s^d & \end{bmatrix}.$$

Clearly, $\deg(P) = md$, $|P| = \pm 1$, and $\mathbf{A}P = [I_m, 0_{n-m}]$. Moreover, the first m columns form a minimal-degree solution to $\mathbf{A}B = I_m$. Lastly, it is easy to see that the last $n-m$ columns of P have linearly independent leading vectors as well, so they form a minimal basis for ker($\mathbf{A}$). Thus, P is a minimal-degree unimodular solution to $\mathbf{A}P = [I_m, 0_{n-m}]$.

□

Experiments on random inputs indicate that for generic unimodular $\mathbf{A} \in \mathbb{K}^{m\times n}$ with deg($\mathbf{A}$) $= d$, and for all minimal-degree unimodular solutions P to $\mathbf{A}P = [I_m, 0_{n-m}]$, $\deg(P) = \lceil \frac{md}{n-m} \rceil$.

BIBLIOGRAPHY

[Ant05] Antoniou, E. N. et al. “Numerical computation of minimal polynomial bases: a generalized resultant approach”. *Linear Algebra Appl.* **405** (2005), pp. 264–278.

[AL10] Antritter, F. & Lévine, J. “Flatness characterization: two approaches”. *Advances in the theory of control, signals and systems with physical modeling*. Vol. 407. Lect. Notes Control Inf. Sci. Springer, Berlin, 2010, pp. 127–139.

[BL00] Beckermann, B. & Labahn, G. “Fraction-free computation of matrix rational interpolants and matrix GCDs”. *SIAM J. Matrix Anal. Appl.* **22**.1 (2000), pp. 114–144.

[Bec99] Beckermann, B. et al. “Shifted normal forms of polynomial matrices”. *Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation (Vancouver, BC)*. ACM, New York, 1999, pp. 189–196.

[Bec06] Beckermann, B. et al. “Normal forms for general polynomial matrices”. *J. Symbolic Comput.* **41**.6 (2006), pp. 708–737.

[Bee87] Beelen, T. G. J. *New algorithms for computing the Kronecker structure of a pencil with applications to systems and control theory*. Dissertation, Technische Hogeschool Eindhoven, Eindhoven, 1987, With a Dutch summary. Technische Hogeschool Eindhoven, Department of Mathematics, Eindhoven, 1987, pp. viii+135.

[Ber15] Bernardin, L. et al. *Maple Programming Guide*. Maplesoft. 2015.

[Buc65] Buchberger, B. “Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal [An Algorithm for Finding the Basis Elements in the Residue Class Ring Modulo a Zero Dimensional Polynomial Ideal] (Trans. in *Journal of Symbolic Comp., Special Issue on Logic, Math., and Comp. Science: Interactions*, 41(3-4):475–511, 2006)”. PhD thesis. 1965.

[Can93] Caniglia, L. et al. “Algorithmic aspects of Suslin’s proof of Serre’s conjecture”. *Comput. Complexity* **3**.1 (1993), pp. 31–55.

[Car35] Cartan, Élie. *La méthode du repère mobile, la théorie des groupes continus, et les espaces généralisés*. Vol. 5. Exposés de Géométrie. Paris: Hermann, 1935.

[CW02] Chen, F. & Wang, W. “The μ-basis of a planar rational curve-properties and computation”. *Graphical Models* **64**.6 (2002), pp. 368–381.

[Che05] Chen, F. et al. “The μ-basis and implicitization of a rational parametric surface”. *J. Symbolic Comput.* **39**.6 (2005), pp. 689–706.

[CL07] Cheng, H. & Labahn, G. “Output-sensitive modular algorithms for polynomial matrix normal forms”. *Journal of Symbolic Computation* **42** (2007), 733Ű–750.

[Cle17] Clelland, J. N. *From Frenet to Cartan: the method of moving frames*. Vol. 178. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2017, pp. xvi+414.

[Cox98a] Cox, D. et al. *Using algebraic geometry*. Vol. 185. Graduate Texts in Mathematics. Springer-Verlag, New York, 1998, pp. xii+499.

[CI15] Cox, D. A. & Iarrobino, A. A. “Strata of rational space curves”. *Comput. Aided Geom. Design* **32** (2015), pp. 50–68.

[Cox98b] Cox, D. A. et al. “The moving line ideal basis of planar rational curves”. *Comput. Aided Geom. Design* **15**.8 (1998), pp. 803–827.

[D’A04] D’Andrea, C. “On the structure of μ-classes”. *Comm. Algebra* **32**.1 (2004), pp. 159–165.

[Elk12] Elkadi, M. et al. “Approximate GCD of several univariate polynomials with small degree perturbations”. *Journal of Symbolic Computation* **47** (2012), pp. 410–421.

[FQ07] Fabiańska, A. & Quadrat, A. “Applications of the Quillen-Suslin theorem to multidimensional systems theory”. *Gröbner bases in control theory and signal processing*. Vol. 3. Radon Ser. Comput. Appl. Math. Walter de Gruyter, Berlin, 2007, pp. 23–106.

[Far16] Farouki, R. T. “Rational rotation-minimizing frames—recent advances and open problems”. *Appl. Math. Comput.* **272**.part 1 (2016), pp. 80–91.

[Far14] Farouki, R. T. et al. “Rotation-minimizing osculating frames”. *Comput. Aided Geom. Design* **31**.1 (2014), pp. 27–42.

[FO99] Fels, M. & Olver, P. J. “Moving Coframes. II. Regularization and Theoretical Foundations”. *Acta Appl. Math.* **55** (1999), pp. 127–208.

[FG90] Fitchas, N. & Galligo, A. “Nullstellensatz effectif et conjecture de Serre (théorème de Quillen-Suslin) pour le calcul formel”. *Math. Nachr.* **149** (1990), pp. 231–253.

[Gre78] Green, M. L. “The moving frame, differential invariants and rigidity theorems for curves in homogeneous spaces”. *Duke Math. Journal* **45** (1978), pp. 735–779.

[Gri74] Griffiths, P. A. “On Cartan’s method of Lie groups as applied to uniqueness and existence questions in differential geometry”. *Duke Math. Journal* **41** (1974), pp. 775–814.

[Gug89] Guggenheimer, H. “Computing frames along a trajectory”. *Comput. Aided Geom. Design* **6**.1 (1989), pp. 77–78.

[Gug63] Guggenheimer, H. W. *Differential Geometry*. McGraw-Hill, New York, 1963.

[Hil90] Hilbert, D. “Ueber die Theorie der algebraischen Formen”. *Math. Ann.* **36**.4 (1890), pp. 473–534.

[Hon17] Hong, H. et al. “Algorithm for computing μ-bases of univariate polynomials”. *J. Symbolic Comput.* **80**.3 (2017), pp. 844–874.

[Ima15] Imae, J. et al. “Practical Computation of Flat Outputs for Nonlinear Control Systems”. *Proceedings of the 3rd International Conference on Manufacturing, Optimization, Industrial and Material Engineering (MOIME)*. IOP Conference Series-Materials Science and Engineering. IOP publishing, Bristol, 2015.

[IL16] Ivey, T. A. & Landsberg, J. M. *Cartan for beginners: Differential geometry via moving frames and exterior differential systems*. Vol. 175. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2016, pp. xviii + 453.

[JG09] Jia, X. & Goldman, R. “μ-bases and singularities of rational planar curves”. *Comput. Aided Geom. Design* **26**.9 (2009), pp. 970–988.

[LS92] Logar, A. & Sturmfels, B. “Algorithms for the Quillen-Suslin theorem”. *J. Algebra* **145**.1 (1992), pp. 231–239.

[LY05] Lombardi, H. & Yengui, I. “Suslin's algorithms for reduction of unimodular rows”. *J. Symbolic Comput.* **39**.6 (2005), pp. 707–717.

[Mar01] Martin, P. et al. “Flat systems: open problems, infinite dimensional extension, symmetries and catalog”. *Advances in the control of nonlinear systems (Murcia, 2000)*. Vol. 264. Lect. Notes Control Inf. Sci. Springer, London, 2001, pp. 33–57.

[Olv15] Olver, P. “Modern developments in the theory and applications of moving frames”. *London Math. Soc. Impact150 Stories*. Vol. 1. London Math. Soc, 2015, pp. 14–50.

[PW95] Park, H. & Woodburn, C. “An algorithmic proof of Suslin's stability theorem for polynomial rings”. *J. Algebra* **178**.1 (1995), pp. 277–298.

[PW98] Polderman, J. W. & Willems, J. C. *Introduction to the Mathematical Theory of Systems and Control*. Springer, New York, 1998.

[Pop70] Popov, V. M. “Some properties of the control systems with irreducible matrix-transfer functions” (1970), 169–180. Lecture Notes in Math., Vol. 144.

[Pop72] Popov, V. M. “Invariant description of linear, time-invariant controllable systems”. *SIAM J. Control* **10** (1972), pp. 252–264.

[SC95] Sederberg, T. & Chen, F. "Implicitization using moving curves and surfaces". *Computer Graphics Proceedings, Annual Conference Series* **2** (1995), pp. 301–308.

[Shi12] Shi, X. et al. "Using μ-bases to implicitize rational surfaces with a pair of orthogonal directrices". *Comput. Aided Geom. Design* **29**.7 (2012), pp. 541–554.

[SG09] Song, N. & Goldman, R. "μ-bases for polynomial systems in one variable". *Comput. Aided Geom. Design* **26**.2 (2009), pp. 217–230.

[SV05] Storjohann, A. & Villard, G. "Computing the Rank and a Small Nullspace Basis of a Polynomial Matrix". *Proceedings of the 2005 International Symposium on Symbolic and Algebraic Computation*. ISSAC '05. Beijing, China: ACM, 2005, pp. 309–316.

[TW14] Tesemma, M. & Wang, H. "Intersections of rational parametrized plane curves". *Eur. J. Pure Appl. Math.* **7**.2 (2014), pp. 191–200.

[VS78] Vardulakis, A. I. G. & Stoyle, P. N. R. "Generalized resultant theorem". *J. Inst. Math. Appl.* **22**.3 (1978), pp. 331–335.

[Wae70] Waerden, B. L. van der. *Algebra I*. Ungar, New York, 1970.

[Wan08] Wang, W. et al. "Computation of rotation minimizing frame". *ACM Trans. Graph* **27**.1 (2008), 18pp.

[ZS01] Zheng, J. & Sederberg, T. W. "A direct approach to computing the μ-basis of planar rational curves". *J. Symbolic Comput.* **31**.5 (2001), pp. 619–629.

[ZL14] Zhou, W. & Labahn, G. "Unimodular completion of polynomial matrices". *ISSAC 2014—Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*. ACM, New York, 2014, pp. 413–420.

[Zho12] Zhou, W. et al. "Computing minimal nullspace bases". *ISSAC 2012—Proceedings of the 37th International Symposium on Symbolic and Algebraic Computation*. ACM, New York, 2012, pp. 366–373.