# THE CANONICAL DECOMPOSITION OF $\mathcal{C}_d^n$ AND NUMERICAL GRÖBNER AND BORDER BASES *

KIM BATSELIER†, PHILIPPE DREESEN†, AND BART DE MOOR †

**Abstract.** This article introduces the canonical decomposition of the vector space of multivariate polynomials for a given monomial ordering. Its importance lies in solving multivariate polynomial systems, computing Gröbner bases and solving the ideal membership problem. An SVD-based algorithm is presented that numerically computes the canonical decomposition. It is then shown how by introducing the notion of divisibility into this algorithm a numerical Gröbner basis can also be computed. In addition, we demonstrate how the canonical decomposition can be used to decide whether the affine solution set of a multivariate polynomial system is zero-dimensional and to solve the ideal membership problem numerically. The SVD-based canonical decomposition algorithm is also extended to numerically compute border bases. A tolerance for each of the algorithms is derived using perturbation theory of principal angles. This derivation shows that the condition number of computing the canonical decomposition and numerical Gröbner basis is essentially the condition number of the Macaulay matrix. Numerical experiments with both exact and noisy coefficients are presented and discussed.

**Key words.** singular value decomposition, principal angles, Macaulay matrix, multivariate polynomials, Gröbner basis, border basis

**AMS subject classifications.** 15A03,15B05,15A18,15A23

**1. Introduction.** Multivariate polynomials appear in a myriad of applications [10, 12, 15, 43]. Often in these applications, the problem that needs to be solved is equivalent with finding the roots of a system of multivariate polynomials. With the advent of the Gröbner basis and Buchberger's Algorithm [11], symbolic methods became an important tool for solving polynomial systems. These are studied in a branch of mathematics called computational algebraic geometry [14, 15]. Other methods to solve multivariate polynomial systems use resultants [18, 25, 49] or homotopy continuation [2, 38, 53]. Computational algebraic geometry however lacks a strong focus towards numerical methods and symbolic methods have inherent difficulties to deal with noisy data. Hence, there is a need for numerically stable algorithms to cope with these issues. The domain of numerical linear algebra has this focus and numerical stable methods have been developed in this framework to solve problems involving univariate polynomials. For example, computing approximate GCDs of two polynomials has been extensively studied with different approaches [6, 13, 19, 57]. An interesting observation is that the matrices involved are in most cases structured and sparse. Some research therefore focuses on how methods can exploit this structure [5, 8, 40, 44].

*Kim Batselier is a research assistant at the Katholieke Universiteit Leuven, Belgium. Philippe Dreesen is supported by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). Bart De Moor is a full professor at the Katholieke Universiteit Leuven, Belgium. Research supported by Research Council KUL: GOA/10/09 MaNet, PFV/10/002(OPTEC), several PhD/postdoc & fellow grants; Flemish Government: IOF:IOF/KP/SCORES4CHEM, FWO: PhD/postdoc grants, projects: G.0588.09 (Brain-machine), G.0377.09(Mechatronics MPC), G.0377.12(Structured systems), IWT: PhD Grants, projects: SBO LeCoPro, SBO Climaqs, SBO POM, EUROSTARS SMART, iMinds 2012, Belgian Federal Science Policy Office: IUAP P7/(DYSCO, Dynamical systems, control and optimization, 2012-2017), EU: ERNSI, FP7-EMBOCON(ICT-248940), FP7-SADCO(MC ITN-264735), ERC ST HIGHWIND (259 166), ERC AdG A-DATADRIVE-B,COST: Action ICO806: IntelliCIS; The scientific responsibility is assumed by its authors.

†Department of Electrical Engineering ESAT-SCD, KU Leuven / iMinds – KU Leuven Future Health Department, 3001 Leuven, Belgium

Contrary to the univariate case, the use of numerical linear algebra methods for problems involving multivariate polynomials is not so widespread [9, 25, 55, 56]. It is the goal of this article to bridge this gap by introducing concepts from algebraic geometry in the setting of numerical linear algebra. The main contribution of this article is the introduction of this canonical decomposition, together with an SVD-based algorithm to compute this decomposition numerically. Furthermore, we show in this article how the canonical decomposition is central in solving the ideal membership problem, the numerical computation of a Gröbner order border basis and the determination of the number of affine solutions of a multivariate polynomial system. Finally, we derive the condition number for computing the canonical decomposition and show that it is basically the condition number of the Macaulay matrix. All algorithms are illustrated with numerical examples. To our knowledge, no SVD-based method to compute a Gröbner basis has been proposed yet. The canonical decomposition, Gröbner basis and border bases are the result of 2 consecutive SVD's and are hence computed in a numerically backward stable manner. The effect of noise on the coefficients of the polynomials is also considered in these examples. All algorithms were implemented as a Matlab[45]/Octave [17] Polynomial Numerical Linear Algebra (PNLA) package and are freely available from `https://github.com/kbatseli/PNLA_MATLAB_OCTAVE`. All numerical experiments were performed on a 2.66 GHz quad-core desktop computer with 8 GB RAM using Octave and took around 3 seconds or less to complete.

The outline of this article is as follows. First, some necessary notation is introduced in Section 2. In Section 3, the Macaulay matrix is defined. An interpretation of its row space is given that naturally leads to the ideal membership problem. The rank of the Macaulay matrix results in the canonical decomposition described in Section 4. An algorithm is described to compute this decomposition and numerical experiments are given. Both cases of exact and inexact coefficients are investigated. The notion of divisibility is introduced into the canonical decomposition in Section 5. This leads to some important applications: a condition for the zero-dimensionality of the solution set of a monomial system and the total number of affine roots can be computed. Another important application is the computation of a numerical Gröbner basis, described in Section 6. This problem has already received some attention for the cases of both exact and inexact coefficients [29, 41, 42, 46, 47, 48, 52]. Exact coefficients refer to the case that they are known with infinite precision. The results for monomial systems are then extended to general polynomial systems. In Section 7, the ideal membership problem is solved by applying the insights of the previous sections. Numerical Gröbner bases suffer from the representation singularity. This is addressed in Section 8, where we introduce border prebases and an algorithm to numerically compute them. Finally, some conclusions are given.

**2. Vector Space of Multivariate Polynomials.** In this section we define some notation. The vector space of all multivariate polynomials over $n$ variables up to degree $d$ over $\mathbb{C}$ will be denoted by $\mathcal{C}_d^n$. Consequently the polynomial ring is denoted by $\mathcal{C}^n$. A canonical basis for this vector space consists of all monomials from degree 0 up to $d$. A monomial $x^a = x_1^{a_1} \ \ldots \ x_n^{a_n}$ has a multidegree $(a_1, \ldots, a_n) \in \mathbb{N}_0^n$ and (total) degree $|a| = \sum_{i=1}^n a_i$. The degree of a polynomial $p$, $\deg(p)$, then corresponds to the highest degree of all monomials of $p$. It is possible to order the terms of multivariate polynomials in different ways and the computed canonical decomposition or Gröbner basis will depend on which ordering is used. For example, it is well-known that a Gröbner basis with respect to the lexicographic monomial ordering is typically more complex (more terms and of higher degree) then with respect to

the reverse lexicographic ordering [15, p.114]. It is therefore important to specify which ordering is used. For a formal definition of monomial orderings together with a detailed description of some relevant orderings in computational algebraic geometry see [14, 15]. The monomial ordering used in this article is the graded xel ordering [3, p.3], which is sometimes also called the degree negative lexicographic monomial ordering. This ordering is graded because it first compares the degrees of the two monomials $a, b$ and applies the xel ordering when there is a tie. The ordering is also multiplicative, which means that if $a < b$ this implies that $ac < bc$ for all $c \in \mathbb{N}_0^n$. The multiplicative property will have an important consequence for the determination of a numerical Gröbner basis as explained in Section 6. A monomial ordering also allows for a multivariate polynomial $f$ to be represented by its coefficient vector. One simply orders the coefficients in a row vector, graded xel ordered, in ascending degree. By convention a coefficient vector will always be a row vector. Depending on the context we will use the label $f$ for both a polynomial and its coefficient vector. $(.)^T$ will denote the transpose of the matrix or vector $(.)$.

**3. Macaulay Matrix.** In this section we introduce the main object of this article, the Macaulay matrix. Its row space is linked to the concept of an ideal in algebraic geometry and this leads to the ideal membership problem.

DEFINITION 3.1. *Given a set of polynomials $f_1, \ldots, f_s \in \mathcal{C}^n$, of degree $d_1, \ldots, d_s$ respectively. The Macaulay matrix of degree $d \geq \max(d_1, \ldots, d_s)$ is the matrix containing the coefficients of*

$$M(d) = \begin{pmatrix} f_1^T & x_1 f_1^T & \ldots & x_n^{d-d_1} f_1^T & f_2^T & x_1 f_2^T & \ldots & x_n^{d-d_s} f_s^T \end{pmatrix}^T \tag{3.1}$$

*as its rows, where each polynomial $f_i$ is multiplied with all monomials from degree 0 up to $d - d_i$ for all $i = 1, \ldots, s$.*

When constructing the Macaulay matrix, it is more practical to start with the coefficient vectors of the original polynomial system $f_1, \ldots, f_s$ after which all the rows corresponding to multiplied polynomials $x^a f_i$ up to a degree $\max(d_1, \ldots, d_s)$ are added. Then one can add the coefficient vectors of all polynomials $x^a f_i$ of one degree higher and so forth until the desired degree $d$ is obtained. This is illustrated in the following example.

EXAMPLE 3.1. *For the following polynomial system in $\mathcal{C}_2^2$*

$$\begin{cases} f_1: & x_1x_2 - 2x_2 & = & 0 \\ f_2: & x_2 - 3 & = & 0 \end{cases}$$

*we have that* $\max(d_1, d_2) = 2$. *The Macaulay matrix $M(3)$ is then*

$$M(3) = \begin{array}{c} \\ f_1 \\ f_2 \\ x_1 f_2 \\ x_2 f_2 \\ x_1 f_1 \\ x_2 f_1 \\ x_1^2 f_2 \\ x_1 x_2 f_2 \\ x_2^2 f_2 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 & x_1^3 & x_1^2x_2 & x_1x_2^2 & x_2^3 \end{array} \\ \begin{pmatrix} 0 & 0 & -2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -3 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

*The first two rows correspond with the coefficient vectors of $f_1, f_2$. Since* $\max(d_1, d_2) = 2$ *and $d_2 = 1$ the next two rows correspond to the coefficient vectors of $x_1 f_2$ and $x_2 f_2$*

*of degree two. Notice that these first four rows make up $M(2)$ when the columns are limited to all monomials of degree zero up to two. The next rows that are added are the coefficient vectors of $x_1f_1, x_2f_1$ and $x_1^2f_2, x_1x_2f_2, x_2^2f_2$ which are all polynomials of degree three.*

The Macaulay matrix depends explicitly on the degree $d$ for which it is defined, hence the notation $M(d)$. It was Macaulay who introduced this matrix, drawing from earlier work by Sylvester [51], in his work on elimination theory, resultants and solving multivariate polynomial systems [35, 36]. For a degree $d$, the number of rows $p(d)$ of $M(d)$ is given by the polynomial

$$p(d) \;=\; \sum_{i=1}^{s} \binom{d-d_i+n}{n} \;=\; \frac{s}{n!}\,d^n + O(d^{n-1}) \tag{3.2}$$

and the number of columns $q(d)$ by

$$q(d) \;=\; \binom{d+n}{n} \;=\; \frac{1}{n!}\,d^n + O(d^{n-1}). \tag{3.3}$$

From these two expressions it is clear that the number of rows will grow faster than the number of columns as soon as $s > 1$. Since the total number of monomials in $n$ variables from degree 0 up to degree $d$ is given by $q(d)$, it also follows that $\dim(\mathcal{C}_d^n) = q(d)$. We denote the rank of $M(d)$ by $r(d)$ and the dimension of its right null space by $c(d)$.

**3.1. Row space of the Macaulay Matrix.** Before defining the canonical decomposition, we first need to interpret the row space of $M(d)$. The row space of $M(d)$, denoted by $\mathcal{M}_d$, describes all $n$-variate polynomials

$$\mathcal{M}_d = \left\{ \sum_{i=1}^{s} h_i\, f_i \;:\; h_i \in \mathcal{C}_{d-d_i}^n \;(i = 1, \ldots, s) \right\}. \tag{3.4}$$

This is closely related to the following concept of algebraic geometry.

DEFINITION 3.2. *Let $f_1, \ldots, f_s \in \mathcal{C}_d^n$. Then we set*

$$\langle f_1, \ldots, f_s \rangle \;=\; \left\{ \sum_{i=1}^{s} h_i f_i : h_1, \ldots, h_s \in \mathcal{C}^n \right\} \tag{3.5}$$

*and call it the ideal generated by $f_1, \ldots, f_s$.*

The ideal hence contains all polynomial combinations (3.4) without any constraints on the degrees of $h_1, \ldots, h_s$. In addition, an ideal is called zero-dimensional when the solution set of $f_1, \ldots, f_s$ is finite. We will denote the set of all polynomials of the ideal $\langle f_1, \ldots, f_s \rangle$ with a degree from 0 up to $d$ by $\langle f_1, \ldots, f_s \rangle_d$. It is now tempting to interpret $\mathcal{M}_d$ as $\langle f_1, \ldots, f_s \rangle_d$ but this is not necessarily the case. $\mathcal{M}_d$ does not in general contain all polynomials of degree $d$ that can be written as a polynomial combination (3.4).

EXAMPLE 3.2. *Consider the following polynomial system in $\mathcal{C}_4^3$*

$$\begin{cases} -9 - x_2^2 - x_3^2 - 3\,x_2^2x_3^2 + 8\,x_2x_3 = 0 \\ -9 - x_3^2 - x_1^2 - 3\,x_1^2x_3^2 + 8\,x_1x_3 = 0 \\ -9 - x_1^2 - x_2^2 - 3\,x_1^2x_2^2 + 8\,x_1x_2 = 0 \end{cases}$$

*The polynomial* $p = 867\,x_1^5 - 1560\,x_3\,x_2\,x_1 - 2312\,x_2^2\,x_1 + 1560\,x_3\,x_1^2 + 2104\,x_2\,x_1^2 - 1526\,x_1^3 + 4896\,x_2 - 2295\,x_1$ *of degree five is not an element of* $\mathcal{M}_5$*. This can easily be verified by a rank test: append the coefficient vector of* $p$ *to* $M(5)$ *and the rank increases by one which means that* $p$ *does not lie in* $\mathcal{M}_5$*. However,* $p \in \mathcal{M}_{11}$*, which implies that a polynomial combination of degree eleven is necessary in order to construct* $p$*. In doing so, all terms of degrees six up to eleven cancel one another.*

Hence, the reason that not all polynomials of degree $d$ lie in $\mathcal{M}_d$ is that it is possible that a polynomial combination of a degree higher than $d$ is required. This is due to the polynomial system having roots at infinity. The problem of determining whether a given multivariate polynomial $p$ lies in the ideal $\langle f_1, \ldots, f_s \rangle$ generated by given polynomials $f_1, \ldots, f_s$ is called the ideal membership problem in algebraic geometry.

PROBLEM 3.1. *Let* $p, f_1, \ldots, f_s \in \mathcal{C}^n$*, then decide whether* $p \in \langle f_1, \ldots, f_s \rangle$*.*

Example 3.2 indicates that Problem 3.1 could be solved using numerical linear algebra: one could append the coefficient vector of $p$ as an extra row to the Macaulay matrix $M(d)$ and do a rank test for increasing degrees $d$. The two most common numerical methods for rank determination are the SVD and the rank-revealing QR decomposition. The SVD is the most robust way of determining the numerical rank of a matrix and is therefore the method of choice in this article. As Example 3.2 also has shown, the algorithm requires a stop condition on the degree $d$ for which $M(d)$ should be constructed. We can therefore restate Problem 3.1 in the following way.

PROBLEM 3.2. *Find the degree* $d_I$ *such that the ideal membership problem can be decided by checking whether*

$$rank(\begin{pmatrix} M(d_I) \\ p \end{pmatrix}) \;=\; rank(\; M(d_I)\;)$$

*holds.*

Problem 3.2 is related to finding the ideal membership degree bound. The ideal membership degree bound $I$ is the least value such that for all polynomials $f_1, \ldots, f_s$ whenever $p \in \langle f_1, \ldots, f_s \rangle$ then

$$p = \sum_{i=1}^{s} h_i\, f_i \quad h_i \in \mathcal{C}^n, \deg(h_i\, f_i) \leq I + \deg(p).$$

Upper bounds are available on the ideal membership degree bound $I$. They are for the general case tight and doubly exponential [33, 37, 54], which renders them useless for most practical purposes. In Section 7 it will be shown how Problem 3.1 can be solved numerically for zero-dimensional ideals without the need of constructing $M(d)$ for the doubly exponential upper bound on $I$.

There is a different interpretation of the row space of $M(d)$ such that all polynomials of degree $d$ are contained in it. This requires homogeneous polynomials. A polynomial of degree $d$ is homogeneous when every term is of degree $d$. A non-homogeneous polynomial can easily be made homogeneous by introducing an extra variable $x_0$.

DEFINITION 3.3. *([15, p. 373]) Let* $f \in \mathcal{C}_d^n$ *of degree* $d$*, then its homogenization* $f^h \in \mathcal{C}_d^{n+1}$ *is the polynomial obtained by multiplying each term of* $f$ *with a power of* $x_0$ *such that its degree becomes* $d$*.*

EXAMPLE 3.3. *Let* $f = x_1^2 + 9x_3 - 5 \in \mathcal{C}_2^3$*. Then its homogenization is* $f^h = x_1^2 + 9x_0x_3 - 5x_0^2$*, where each term is now of degree 2.*

The ring of all homogeneous polynomials in $n+1$ variables will be denoted $\mathcal{P}^n$ and likewise the vector space of all homogeneous polynomials in $n+1$ variables of degree $d$ by $\mathcal{P}_d^n$. This vector space is spanned by all monomials in $n+1$ variables of degree $d$ and hence $\dim(\mathcal{P}_d^n) = \binom{d+n}{n}$, which equals the number of columns of $M(d)$. This is no coincidence, given a set of non-homogeneous polynomials $f_1, \ldots, f_s$ we can also interpret $\mathcal{M}_d$ as the vector space

$$\mathcal{M}_d \;=\; \left\{ \sum_{i=1}^{s} h_i \, f_i^h \;:\; h_i \in \mathcal{P}_{d-d_i}^n \; (i = 1, \ldots, s) \right\}, \tag{3.6}$$

where the $f_i^h$'s are $f_1, \ldots, f_s$ homogenized and the $h_i$'s are also homogeneous. The corresponding homogeneous ideal is denoted by $\langle f_1^h, \ldots, f_s^h \rangle$. The homogeneity ensures that the effect of higher order terms cancelling one another as in Example 3.2 does not occur. This guarantees that all homogeneous polynomials of degree $d$ are contained in $\mathcal{M}_d$. Or in other words, $\mathcal{M}_d \;=\; \langle f_1^h, \ldots, f_s^h \rangle_d$, where $\langle f_1^h, \ldots, f_s^h \rangle_d$ is the set of all homogeneous polynomials of degree $d$ contained in the homogeneous ideal $\langle f_1^h, \ldots, f_s^h \rangle$. The homogenization of $f_1, \ldots, f_s$ typically introduces extra roots that satisfy $x_0 = 0$ and at least one $x_i \neq 0 \, (i = 1, \ldots, s)$. They are called roots at infinity. Affine roots can then be defined as the roots for which $x_0 = 1$. All nontrivial roots of $f_1^h, \ldots, f_s^h$ are called projective roots. We revisit Example 3.1 to illustrate this point.

EXAMPLE 3.4. *The homogenization of the polynomial system in Example 3.1 is*

$$\begin{cases} f_1^h: & x_1 x_2 - 2 x_2 x_0 &=& 0 \\ f_2^h: & x_2 - 3 x_0 &=& 0. \end{cases}$$

*All homogeneous polynomials* $\sum_{i=1}^{2} h_i \, f_i^h$ *of degree three belong to the row space of* $M(3)$ *from Example 3.1. The non-homogeneous polynomial system had only 1 root* $= \{(2, 3)\}$. *After homogenization, the resulting polynomial system* $f_1^h, f_2^h$ *has 2 nontrivial roots* $= \{(1, 2, 3), (0, 1, 0)\}$.

The homogeneous interpretation is in effect nothing but a relabelling of the columns and rows of $M(d)$. The fact that all homogeneous polynomials of degree $d$ are contained in $\mathcal{M}_d$ simplifies the ideal membership problem for a homogeneous polynomial to a single rank test.

THEOREM 3.4. *Let* $f_1, \ldots, f_s \in \mathcal{C}^n$ *and* $p \in \mathcal{P}_d^n$. *Then* $p \in \langle f_1^h, \ldots, f_s^h \rangle$ *if and only if*

$$\operatorname{rank}\left( \begin{pmatrix} M(d) \\ p \end{pmatrix} \right) \;=\; \operatorname{rank}\left( M(d) \right). \tag{3.7}$$

**4. The Canonical Decomposition of $\mathcal{C}_d^n$.** First, the canonical decomposition is defined and illustrated with an example. Then, the SVD-based algorithm to numerically compute the canonical decomposition is presented. This is followed by a detailed discussion on numerical aspects, which are illustrated by worked-out examples.

**4.1. Definition.** The interpretation of the row space immediately results in a similar interpretation for the rank $r(d)$ of $M(d)$. Evidently, the rank $r(d)$ counts the number of linearly independent polynomials lying in $\mathcal{M}_d$. More interestingly, the rank also counts the number of linearly independent leading monomials of $\mathcal{M}_d$. This is easily seen from bringing the Macaulay matrix $M(d)$ into a reduced row echelon form $R(d)$. In order for the linearly independent monomials to be leading monomials

a column permutation $Q$ is required which flips all columns from left to right. Then the Gauss-Jordan elimination algorithm can be run, working from left to right. The reduced row echelon form then ensures that each pivot element corresponds with a linearly independent leading monomial. We illustrate this procedure in the following example.

EXAMPLE 4.1. *Consider the polynomial system*

$$\left\{ \begin{array}{lrcl} f_1: & x_1\,x_2 - 2x_2 & = & 0 \\ f_2: & x_2 - 3 & = & 0 \end{array} \right.$$

*and fix the degree to 3. First, the left-to-right column permutation $Q$ is applied to $M(3)$. Bringing $M(3)\,Q$ into reduced row echelon form results in*

$$R(3) \;=\; \begin{array}{c} \begin{array}{cccccccccc} x_2^3 & x_1x_2^2 & x_1^2x_2 & x_1^3 & x_2^2 & x_1x_2 & x_1^2 & x_2 & x_1 & 1 \end{array} \\ \left( \begin{array}{cccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -27 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -18 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -9 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \end{array}.$$

*From the reduced row echelon form one can see that the rank of $M(3)$ is 8. Notice how the left-to-right permutation ensured that the 8 pivot elements, corresponding with the monomials $\{x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^2x_2, x_1x_2^2, x_2^3\}$ are leading monomials with respect to the monomial ordering. The Gauss-Jordan algorithm returns a set of 8 polynomials, that all together span $\mathcal{M}_3$. In addition, for each of these polynomials, its leading monomial corresponds with a particular pivot element of $R(3)$.*

The $r(d)$ polynomials that can be read off from $R(d)$ span $\mathcal{M}_d$ and we will show how for a particular degree a subset of these polynomials corresponds with a reduced Gröbner basis. Interpreting the rank $r(d)$ in terms of linearly independent leading monomials naturally leads to a canonical decomposition of $\mathcal{C}_d^n$. The vector space spanned by the $r(d)$ leading monomials of $R(d)$ will be denoted $\mathcal{A}_d$. Its complement spanned by the remaining monomials will be denoted $\mathcal{B}_d$. We will call these monomials that span $\mathcal{B}_d$ the normal set or standard monomials. This leads to the following definition.

DEFINITION 4.1. *Let $f_1, \ldots, f_s$ be a multivariate polynomial system with a given monomial ordering. Then we define the canonical decomposition as the decomposition of the monomial basis of $\mathcal{C}_d^n$ into a set of linearly independent leading monomials $A(d)$ and standard monomials $B(d)$.*

Naturally, $\mathcal{C}_d^n = \mathcal{A}_d \oplus \mathcal{B}_d$ and $\dim \mathcal{A}_d = r(d), \dim \mathcal{B}_d = c(d)$. Observe that the monomial bases for $\mathcal{A}_d$ and $\mathcal{B}_d$ also have a homogeneous interpretation.

EXAMPLE 4.2. *For the polynomial system of Example 4.1 and degree 3 the canonical decomposition is $A(3) = \{x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^2x_2, x_1x_2^2, x_2^3\}$, and $B(3) = \{1, x_1^3\}$. If $e_i$ denotes the $i$th canonical basis column vector, then these monomial bases are in matrix form*

$$A(3) \;=\; \begin{pmatrix} e_2 & e_3 & e_4 & e_5 & e_6 & e_8 & e_9 & e_{10} \end{pmatrix}^T$$

*and*

$$B(3) \;=\; \begin{pmatrix} e_1 & e_7 \end{pmatrix}^T.$$

For the sake of readability the notation for $A(d)$ and $B(d)$ is used for both the set of monomials and the matrices, as in Example 4.2. The dependence of the canonical decomposition on the monomial ordering is easily understood from Example 4.1. A different admissible monomial ordering would correspond with a different column permutation $Q$ and this would result in different monomial bases $A(3)$ and $B(3)$.

The importance of this canonical decomposition is twofold. As will be shown in Section 6, the linearly independent monomials $A(d)$ play an important role in the computation of a Gröbner basis of $f_1, \ldots, f_s$. The normal set $B(d)$ is intimately linked with the problem of finding the roots of the polynomial system $f_1, \ldots, f_s$. Indeed, it is well-known that for a polynomial system $f_1^h, \ldots, f_s^h$ with a finite number of projective roots, the quotient space $\mathcal{P}_d^n / \langle f_1^h, \ldots, f_s^h \rangle_d$ is a finite-dimensional vector space [14, 15]. The dimension of this vector space equals the total number of projective roots of $f_1^h, \ldots, f_s^h$, counting multiplicities, for a large enough degree $d$. From the rank-nullity theorem, it then follows that

$$\begin{aligned} c(d) &= q(d) - \operatorname{rank}\left( M(d) \right), \\ &= \dim \mathcal{P}_d^n - \dim \langle f_1^h, \ldots, f_s^h \rangle_d, \\ &= \dim \mathcal{P}_d^n / \langle f_1^h, \ldots, f_s^h \rangle_d, \\ &= \dim \mathcal{B}_d. \end{aligned}$$

This function $c(d)$ that counts the number of homogeneous standard monomials of degree $d$ is called the Hilbert function. This leads to the following theorem.

THEOREM 4.2. *For a zero-dimensional ideal $\langle f_1^h, \ldots, f_s^h \rangle$ with $m$ projective roots (counting multiplicities) there exists a degree $d_c$ such that $c(d) = m \; (\forall \, d \geq d_c)$.*

Furthermore, $m = d_1 \cdots d_s$ according to Bézout's Theorem [14, p.97] when $s = n$. This effectively links the degrees of the polynomials $f_1, \ldots, f_s$ to the nullity of the Macaulay matrix. The roots can be retrieved from a generalized eigenvalue problem as discussed in [1, 49, 50]. The monomials $A(d)$ also tell us how large the degree $d$ of $M(d)$ then should be to construct these eigenvalue problems, as demonstrated in Section 8. Another interesting result is that if the nullity $c(d)$ never converges to a fixed number $m$, then it will grow polynomially. The degree of this polynomial $c(d)$ then equals the dimension of the projective solution set [15, p.463].

It is commonly known that bringing a matrix into a reduced row echelon form is numerically not the most reliable way of determining the rank of a matrix. In the next section a more robust SVD-based method for computing the canonical decomposition of $\mathcal{C}_d^n$ and finding the polynomial basis $R(d)$ is presented.

**4.2. Numerical Computation of the Canonical Decomposition.** As mentioned in the previous section, the rank determination of $M(d)$ is the first essential step in computing the canonical decomposition of $\mathcal{C}_d^n$. Bringing the matrix into reduced row echelon form by means of a Gauss-Jordan elimination is not a robust method for determining the rank. In addition, since the monomial ordering is fixed no column pivoting is allowed, which potentially results in numerical instabilities. We therefore propose to use the SVD for which numerical backward stable algorithms exist [22]. In

addition, an orthogonal basis $V_1$ for $\mathcal{M}_d$ can also be retrieved from the right singular vectors. The next step is to find $A(d), B(d)$ and the $r(d)$ polynomials of $R(d)$. The key idea here is that each of these $r(d)$ polynomials is spanned by the standard monomials and one leading monomial of $A(d)$. Suppose a subset $A \subseteq A(d)$ and $B \subseteq B(d)$, both ordered in ascending order, are available. It is then possible to test whether the next monomial larger than the largest monomial of $A(d)$ is a linearly independent leading monomial. We will illustrate the principle by the following example.

EXAMPLE 4.3. *Suppose that the following subsets* $A = \{x_1, x_2\}, B = \{1\}$ *of* $A(3) = \{x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^2x_2, x_1x_2^2, x_2^3\}$, $B(3) = \{1, x_1^3\}$ *from Example 4.2 are available. The next monomial according to the monomial ordering is* $x_1^2$. *The next possible polynomial from* $R(3)$ *is then spanned by* $\{1, x_1^2\}$. *If such a polynomial lies in* $\mathcal{M}_3$ *then* $x_1^2$ *is a linearly independent leading monomial and can be added to* $A$. *If not,* $x_1^2$ *should be added to* $B$. *This procedure can be repeated until all monomials up to degree three have been tested. For the case of* $x_1^2$ *there is indeed such a polynomial present in* $R(3)$ *as can be seen from Example 4.2:* $x_1^2 - 4$. *This polynomial therefore lies in both the vector spaces* $\mathcal{M}_3$ *and span*$(\{1, x_1^2\})$. *Computing a basis for the intersection between* $\mathcal{M}_3$ *and span*$(\{1, x_1^2\})$ *will therefore reveal whether* $x_1^2 \in A(3)$.

Given the subsets $A$ and $B$, testing whether a monomial $x^a \in A(d)$ corresponds with computing the intersection between $\mathcal{M}_d$ and $\text{span}(\{B, x^a\})$. Let

$$M(d) \;=\; U\,\Sigma\,V^T$$

be the SVD of $M(d)$, then $V$ can be partitioned into $(V_1\ V_2)$ where the columns of $V_1$ are an orthogonal basis for $\mathcal{M}_d$ and the columns of $V_2$ are an orthogonal basis for the kernel of $M(d)$. If $E$ denotes the matrix for which the rows are a canonical basis for $\text{span}(\{B, x^a\})$, then one way of computing the intersection would be to solve the following overdetermined linear system

$$\begin{pmatrix} E^T & V_1 \end{pmatrix} x \;=\; 0. \tag{4.1}$$

If there is a non-empty intersection, then (4.1) has a non-trivial solution $x$. The size of the matrix $\begin{pmatrix} E^T & V_1 \end{pmatrix}$ can grow rather large, $q(d) \times (r(d) + k)$, where $k$ is the cardinality of $\{B, x^a\}$. Using principal angles to determine the intersection involves a smaller $c(d) \times k$ matrix and is therefore preferred. An intersection implies a principal angle of zero between the two vector spaces. The cosine of the principal angles can be retrieved from the following theorem.

THEOREM 4.3. *([7, p. 582]) Assume that the columns of* $V_1$ *and* $E^T$ *form orthogonal bases for two subspaces of* $\mathcal{C}_d^n$. *Let*

$$E\,V_1, \;=\; Y\,C\,Z^T, \quad C \;=\; \text{diag}(\sigma_1, \ldots, \sigma_k), \tag{4.2}$$

*be the SVD of* $E\,V_1$ *where* $Y^TY = I_k, Z^TZ = I_r$. *If we assume that* $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k$, *then the cosines of the principal angles between this pair of subspaces are given by*

$$\cos(\theta_i) = \sigma_i(E\,V_1).$$

Computing principal angles smaller than $10^{-8}$ in double precision is impossible using Theorem 4.3. This is easily seen from the second order approximation of the cosine of its Maclaurin series: $\cos(x) \approx 1 - x^2/2$. If $x < 10^{-8}$ then the $x^2/2$ term will be smaller than the machine precision $\epsilon \approx 2 \times 10^{-16}$ and hence $\cos(x)$ will be exactly

1. For small principal angles it is numerically better to compute the sines using the following Theorem.

THEOREM 4.4. *([7, p. 582-583] and [28, p. 6]) The singular values $\mu_1, \ldots, \mu_k$ of the matrix $E^T - V_1 V_1^T E^T$ are given by $\mu_i = \sqrt{1-\sigma_i^2}$ where the $\sigma_i$ are defined in* (4.2). *Moreover, the principal angles satisfy the equalities $\theta_i = \arcsin(\mu_i)$. The right principal vectors can be computed as $v_i = E^T z_i$, $(i = 1, \ldots, k)$, where $z_i$ are the corresponding right singular vectors of $E^T - V_1 V_1^T E^T$.*

Testing for a non-empty intersection between the row spaces of $U$ and $E$ is hence equivalent with inspecting the smallest singular value $\mu_m$ of $E^T - U^T U E^T$. Notice however that

$$\begin{aligned} E^T - V_1 V_1^T E^T &= (I - V_1 V_1^T) E^T, \\ &= V_2 V_2^T E^T. \end{aligned}$$

This implies that if there is a non-empty intersection, then the reduced polynomial $r$ can be retrieved as the right singular vector $v_k$ of the $c(d) \times k$ matrix $V_2^T E^T$ corresponding with $\mu_k$. The whole algorithm is summarized in pseudo-code in Algorithm 4.1 and is implemented in the PNLA package as candecomp.m. The algorithm iterates over all $n$-variate monomials from degree 0 up do $d$, in ascending order. The set containing all these monomials is denoted by $\mathcal{T}_d^n$. The computational complexity is dominated by the SVD of $M(d)$ for determining the rank and computing the orthogonal basis $V_2$. A full SVD is not required, only the diagonal matrix containing the singular vales and right singular vectors need to be computed. This takes approximately $4p(d)q(d)^2 + 8q(d)^3$ flops. Substitution of the polynomial expressions (3.2) and (3.3) for $p(d)$ and $q(d)$ in this flop counts leads to a computational complexity of approximately $4(s+2)d^{3n}/(n!)^3$. All subsequent SVDs of $V_2^T E^T$ in Algorithm 4.1 have a total computational complexity of $O(c(d)^3)$, which simplifies to $O(m^3)$ from some degree and for the case there are a finite number of projective roots. The combinatorial growth of the dimensions of the Macaulay matrix quickly prevents the computation of its SVD. We have therefore developed a recursive orthogonalization algorithm for the Macaulay matrix that exploits both its structure and sparsity [4]. This recursive algorithm uses the orthogonal $V_2$ of $M(d)$ and updates it to the orthogonal basis $V_2$ for $M(d+1)$. In this way the computational complexity of the orthogonalization step is reduced to approximately $4(s+2)d^{3n-3}/(n-1)!^3$ flops. A full description of our recursive orthogonalization is however out of the scope of this article.

ALGORITHM 4.1. *Computation of the Canonical Decomposition of $\mathcal{C}_d^n$*

***Input**: orthogonal basis $V_2$, monomial ordering*
***Output**: $A(d), B(d)$ and polynomials $R(d)$*

```
A(d), B(d), R(d) ← ∅
for all x^a ∈ T_d^n in ascending monomial order do
   construct E from B(d) and x^a
   [W S Z] ← SVD(V_2^T E^T)
   τ ← tolerance (4.4)
   if arcsin(μ_k) < τ then
      append x^a to A(d) and append v_k^T to R(d)
   else
      append x^a to B(d)
   end if
end for
```

The determination of the rank of $M(d)$ is the first crucial step in the algorithm. If a wrong rank is estimated from the SVD, the subsequent canonical decomposition will also be wrong. The default tolerance used in the SVD-based rank determination is $\tau_r = k \max(p(d), q(d)) \operatorname{eps}(\sigma_1)$ where $\operatorname{eps}(\sigma_1)$ returns the distance from the largest singular value of $M(d)$ to the next larger in magnitude double precision floating point number. The numerical rank $r(d)$ is chosen such that $\sigma_{r(d)} > \tau_r > \sigma_{r(d)+1}$. The approxi-rank gap $\sigma_{r(d)}/\sigma_{r(d)+1}$ [34, p. 920] then determines the difficulty of revealing the numerical rank. In practice, a rather well-conditioning of determining the numerical rank of the Macaulay matrix for nonzero-dimensional ideals is observed. Approxi-rank gaps are typically around $10^{10}$. This is demonstrated in the polysys_ collection.m file in the Matlab/Octave PNLA package, which contains over a 100 multivariate polynomial systems together with their approxi-rank gap. Small approxi-rank gaps around unity indicate inherent 'difficult' polynomial systems. We now determine the tolerance $\tau$ for deciding when a computed principal angle is numerically zero using a perturbation result. It is shown in [7] that the condition number of the principal angle $\theta$ between the row spaces of $M(d)$ and $E(d)$ is essentially $\max(\kappa(M), \kappa(E))$, where $\kappa$ denotes the condition number of a matrix. More specifically, let $\Delta M, \Delta E$ be the perturbations of $M(d), E(d)$ respectively with

$$\frac{||\Delta M||_2}{||M||_2} \leq \epsilon_M, \quad \frac{||\Delta E||_2}{||E||_2} \leq \epsilon_E.$$

Then the following relationship [7, p. 585]

$$|\theta - \tilde{\theta}| \leq \sqrt{2}\,(\epsilon_M\,\kappa(M) + \epsilon_E\,\kappa(E)) + \text{higher order terms} \tag{4.3}$$

holds where $\tilde{\theta}$ is the principal angle between the perturbed vector spaces. $E(d)$ is exact and unperturbed so we can therefore set $\epsilon_E = 0$. Also, when there is a nontrivial intersection then $\theta = 0$. This allows us to simplify (4.3) to

$$|\tilde{\theta}| \leq \sqrt{2}\,\epsilon_M\,\kappa(M),$$

which shows that the condition number of the principal angle is the condition number of $M(d)$. Hence, the condition number of checking whether $x^a$ is a linearly independent monomial of $\mathcal{M}_d$ is basically $\kappa(M)$. The condition number of the Macaulay matrix is defined as $\kappa(M) = \sigma_1/\sigma_{r(d)}$. Furthermore, it is shown in [7, p. 587] that when the perturbations are due to numerical computations and the orthogonal basis is computed using Householder transformations then

$$|\theta - \tilde{\theta}| \leq 12.5\,\sqrt{2}\,(p\,\kappa(M) + k\,\kappa(E))\,2^{-53} + \text{higher order terms.}$$

where $p$ is the number of rows of $M(d)$ and $k$ is the cardinality of $\{B, x^a\}$. The factor $2^{-53}$ is due to the fact that we work in double precision. This allows us to set

$$\tau = 12.5\,\sqrt{2}\,(p\,\kappa(M) + k)\,2^{-53}. \tag{4.4}$$

The singular values $\sigma_i$ of $M(d)$ are available from the SVD computation to determine $V_2$. In addition to the inspection of the approxi-rank gap $\sigma_{r(d)}/\sigma_{r(d)+1}$ one can also compare the number of computed elements in $A(d)$ with the rank of $M(d)$ to get an indication whether the algorithm has successfully computed the correct canonical decomposition. Typically, polynomial systems for which the numerical rank is ill-defined (small approxi-rank gap) also have a large condition number, making it difficult to determine the correct $A(d), B(d)$ monomials. In this case symbolical multiprecision Gaussian Elimination is needed to determine a canonical decomposition.

**4.3. Numerical Experiment - Exact Coefficients.** We first consider the case of polynomials with exact coefficients, i.e. coefficients that are known with infinite precision, and illustrate the algorithm with the following numerical example.

EXAMPLE 4.4. *Consider the following polynomial system in* $\mathcal{C}_4^3$

$$\begin{cases} x_1^2 + x_1\,x_3 - 2\,x_2 + 5 &= 0, \\ 2\,x_1^3\,x_2 + 7\,x_2\,x_3^2 - 4\,x_1\,x_2\,x_3 + 3\,x_1 - 2 &= 0, \\ x_2^4 + 2\,x_2\,x_3 + 5\,x_1^2 - 5 &= 0, \end{cases}$$

*with degrees* $d_1 = 2, d_2 = 4, d_3 = 4$. *The canonical decomposition is computed for* $d = 10$ *with Algorithm 4.1. Each polynomial is normalized such that it has a unit 2-norm and the* $333 \times 286$ *Macaulay matrix* $M(10)$ *is constructed. From its SVD the tolerance is set to* $\tau_r = 1.47 \times 10^{-13}$ *and the numerical rank is determined as 254 with an approxi-rank gap of* $\approx 4 \times 10^{13}$. *This implies that* $A(10)$ *and* $B(10)$ *will have 254 and 32 monomials respectively. Algorithm 4.1 indeed returns this number of monomials and corresponding polynomials* $R(10)$. *The principal angles corresponding with the leading monomials* $A(10)$ *are all around* $10^{-15}$. *The smallest principal angle for a monomial of the normal set* $B(10)$ *is* $2.17 \times 10^{-9}$. *Note that the rank estimated from the reduced row echelon form of* $M(10)$ *is 259, which is a strong indication that the reduced row echelon form is not well-suited to compute* $A(10), B(10)$ *and* $R(10)$, *even when column pivoting is used. Over a 100 polynomial systems can be found in the file polysys_collection.m from the PNLA package for which the canonical decomposition is correctly determined by Algorithm 4.1.*

**4.4. Numerical Experiment - Perturbed Coefficients.** Perturbing the coefficients of the polynomial system $f_1, \ldots, f_s$ will change the corresponding canonical decomposition $A(d), B(d)$. This is easily observed from Example 4.2. Suppose we perturb all coefficients of $f_2$ with noise, uniformly distributed over the interval $[0, 10^{-1}]$, to obtain $\tilde{f}_2 = -2.9056 + 0.0456\,x_1 + 1.0789\,x_2$. Then the reduced row echelon form of $\tilde{M}(3)Q$ is

$$\tilde{R}(3) = \begin{array}{c} \begin{array}{cccccccccc} x_2^3 & x_1x_2^2 & x_1^2x_2 & x_1^3 & x_2^2 & x_1x_2 & x_1^2 & x_2 & x_1 & 1 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2876 & -18.3252 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2205 & -14.0501 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0.1691 & -10.7723 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -4191.63 & 8375.27 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.1102 & -7.0250 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0.0845 & -5.3862 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -65.7197 & 127.4393 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0.0423 & -2.6931 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}.$$

The corresponding canonical decomposition hence becomes $\tilde{A}(3) = \{x_2^3, x_1x_2^2, x_1^2x_2, x_1^3, x_2^2, x_1x_2, x_1^2, x_2\}$ and $\tilde{B}(3) = \{1, x_1\}$. A small continuous change of coefficients has therefore led to a 'jump' of the canonical decomposition. This implies that the computation of the canonical decomposition for a given polynomial system under perturbations of its coefficients is an ill-posed problem. This ill-posedness is called a representation singularity [50, p. 325] and is due to the insistence that the monomials of $A(d)$ need to be leading monomials with respect to the monomial ordering. This condition is sufficient to make the representation singularity unavoidable and has implications for the numerical determination of Gröbner bases. An alternative approach

to avoid this representation singularity is the use of border bases [26, 27, 39]. We will discuss our SVD-based algorithm to compute these in Section 8, after introducing the reduced canonical decomposition.

**5. The Reduced Canonical Decomposition of $\mathcal{C}_d^n$.** In this section we introduce the notion of divisibility into the canonical decomposition. This naturally leads to the concept of a reduced canonical decomposition. First, some new notation and concepts are introduced after which Algorithm 4.1 is adjusted such that it produces the reduced decomposition. A numerical example is then worked out.

**5.1. The Reduced Monomials $A^\star(d), B^\star(d)$ and Polynomials $G(d)$.** The polynomial basis $R(d)$ will grow unbounded with the rank $r(d)$ for increasing degrees $d$. It is possible however to reduce this basis to a finite subset that generates the whole ideal $\langle f_1, \ldots, f_s \rangle$. It will be shown in Section 6 that for a sufficiently large degree, this reduced polynomial basis is a reduced Gröbner basis. First the reduced leading monomials $A^\star(d)$ are defined.

DEFINITION 5.1. *Given a set of linearly independent leading monomials $A(d)$, then the set of reduced leading monomials $A^\star(d)$ is defined as the smallest subset of $A(d)$ for which each element of $A(d)$ is divisible by an element of $A^\star(d)$.*

Since there is a one-to-one mapping between leading monomials in $A(d)$ and polynomials of $R(d)$, each element of $A^\star(d)$ will also correspond with a polynomial.

DEFINITION 5.2. *For a given canonical decomposition $A(d), B(d), R(d)$ the reduced polynomials $G(d)$ are defined as the polynomials of $R(d)$ corresponding to the reduced monomial system $A^\star(d)$:*

$$G(d) \;=\; \{r \in R(d) : \forall a \in A^\star(d),\; LM(r) = a\}.$$

The set of reduced leading monomials $A^\star(d)$ can be interpreted as a polynomial system for which the canonical decomposition can also be determined. This allows us to define the reduced normal set $B^\star(d)$.

DEFINITION 5.3. *Let $A(d), B(d)$ be a canonical decomposition implied by $f_1, \ldots, f_s$ and a given monomial ordering. Then the reduced normal set $B^\star(d)$ is the normal set obtained from the canonical decomposition implied by $A^\star(d)$ and the same monomial ordering.*

Typically $B^\star(d) \subseteq B(d)$. This is because monomial multiples of $A^\star(d)$ will fill up zero columns that would have otherwise been associated with monomials of $B(d)$. Furthermore, the presence of a monomial $x_i^\alpha$ for each variable $x_1, \ldots, x_n$ is a necessary condition for the finiteness of $B^\star(d)$. This is easily understood by an example: suppose $A^\star(d)$ does not contain any power of $x_1$, then all powers of $x_1$ will be in $B^\star(d)$ and this set will grow linearly. A monomial system $A^\star(d)$ has a projective solution set because it is already homogeneous. It is shown in [15, p.452] that the dimension of this projective solution set is always one less than its affine solution set. Hence, if the monomial ideal has a finite number of affine roots, then it will have no projective roots whatsoever. We can now state the following theorem on the zero-dimensionality of a monomial system in terms of $A^\star(d)$ and $B^\star(d)$.

THEOREM 5.4. *([15, p. 234]) A monomial system $A^\star(d)$ has $m$ affine roots, counting multiplicities, if and only if for each $1 \leq i \leq n$, there is some $\alpha_i \geq 0$, such that $x_i^{\alpha_i} \in A^\star(d)$. It then also holds that:* $\dim \mathcal{B}_d^\star = m$.

From Theorem 4.2 we know that for a polynomial system with a finite number of projective roots, the nullity of $M(d)$, $c(d)$, will equal the total number of projective roots. Theorem 5.4 will allows us to separate the affine roots from the ones at infinity

for a general polynomial system. In order to do this, the notion of a Gröbner basis will first need to be introduced.

**5.2. Numerical Computation of $A^\star(d), B^\star(d)$ and $G(d)$.** The definition of $A^\star(d)$ uses the complete set of linearly independent leading monomials $A(d)$. A straightforward way to find $A^\star(d)$ would hence be to compute $A(d)$ using Algorithm 4.1, find $A^\star(d)$ from $A(d)$ and select the corresponding polynomials of $R(d)$ to obtain $G(d)$. This is however not efficient since the whole canonical decomposition is computed while only subsets are required. By using the defining property of $A^\star(d)$ it is possible to adjust Algorithm 4.1 such that it directly computes $A^\star(d), B^\star(d)$ and $G(d)$. The algorithm iterates over a set of monomials $\mathcal{X}$ which is initially all monomials of degree 0 up to $d$. The key idea is that each monomial of $A(d)$ is a monomial multiple of a monomial of $A^\star(d)$. So as soon as a linearly independent leading monomial $x^a$ is found, all its monomial multiples do not need to be checked anymore and can be removed from $\mathcal{X}$. When the monomial $x^a$ is not linearly independent it is also removed from $\mathcal{X}$ and added to $B^\star(d)$. When $\mathcal{X}$ is empty the algorithm terminates. Removing monomial multiples of $x^a$ from $\mathcal{X}$ reduces the number of iterations significantly and also guarantees that the computed $B^\star$ is correct. The whole procedure is summarized in pseudo-code in Algorithm 5.1 and is implemented in the PNLA package as rcandecomp.m. Again, the computationally most expensive step is the orthogonalization of $M(d)$. The same arguments on the computational complexity and choosing the tolerance $\tau$ apply as for Algorithm 4.1.

ALGORITHM 5.1. *Computation of $A^\star(d), B^\star(d)$ and $G(d)$*
***Input**: orthogonal basis $V_2$, monomial ordering*
***Output**: $A^\star(d), B^\star(d)$ and polynomials $G(d)$*

$A^\star(d), B^\star(d), G(d) \leftarrow \varnothing$
$\mathcal{X} \leftarrow \mathcal{T}_d^n$
***while*** $\mathcal{X} \neq \varnothing$ ***do***
  $x^a \leftarrow$ *smallest monomial in* $\mathcal{X}$ *according to monomial ordering*
  *construct* $E$ *from* $B^\star(d)$ *and* $x^a$
  $[W\ S\ Z] \leftarrow SVD(V_2^T\ E^T)$
  $\tau \leftarrow$ *tolerance* (4.4)
  ***if*** $\arcsin(\mu_m) < \tau$ ***then***
    *append* $x^a$ *to* $A^\star(d)$ *and append* $v_m^T$ *to* $G(d)$
    *remove* $x^a$ *and all its monomial multiples from* $\mathcal{X}$
  ***else***
    *append* $x^a$ *to* $B^\star(d)$ *and remove it from* $\mathcal{X}$
  ***end if***
***end while***

**5.3. Numerical Experiments.** We revisit the polynomial system of Example 4.4 and illustrate Algorithm 5.1 when the coefficients are exact.

EXAMPLE 5.1. *$A(10)$ of Example 4.4 consists of 254 monomials. Running Algorithm 5.1 on the polynomial system results in the following reduced canonical decomposition:*

$$\begin{aligned} A^\star(10) &= \{x_1\,x_3, x_1^3\,x_2, x_2^4, x_3\,x_2^3, x_3^3\,x_2, x_1^5, x_3^5\} \\ B^\star(10) &= \{1, x_1, x_2, x_3, x_1^2, x_2\,x_1, x_2^2, x_2\,x_3, x_3^2, x_1^3, x_2\,x_1^2, x_2^2\,x_1, x_2^3, x_3\,x_2^2, x_2\,x_3^2, x_3^3, \\ &\quad x_1^4, x_2^2\,x_1^2, x_2^3\,x_1, x_3^2\,x_2^2, x_3^4, x_2^3\,x_1^2\}. \end{aligned}$$

*$A^\star(10)$ consists of 7 monomials and the normal set $B(10)$ is reduced from 32 to 22 monomials.*

**6. Gröbner basis.** In this section, the link is made between the reduced polynomials $G(d)$ and a Gröbner basis of the ideal $\langle f_1, \ldots, f_s \rangle$. This will lead to some insights on the separation of the roots of a polynomial system into an affine part and roots at infinity for the zero-dimensional case. A condition will be derived for this case to determine the affine part of the normal set. We first give the definition of a Gröbner basis.

DEFINITION 6.1. *([15, p. 77]) Given a set of multivariate polynomials $f_1, \ldots, f_s$ and a monomial ordering, then a finite set of polynomials $G = \{g_1, \ldots, g_k\}$ is a Gröbner basis of $\langle f_1, \ldots, f_s \rangle$ if*

$$\forall\, p \in \langle f_1, \ldots, f_s \rangle, \exists\, g \in G \text{ such that } LM(g) \mid LM(p).$$

In addition, we will call a Gröbner basis reduced if no monomial in any element of the basis is divisible by the leading monomials of the other elements of the basis. Note also from the definition that a Gröbner basis depends on the monomial ordering. One can think of a Gröbner basis as another set of generators of the ideal $\langle f_1, \ldots, f_s \rangle$, hence the name 'basis'. It is a classical result that for each ideal $\langle f_1, \ldots, f_s \rangle$, there exists such a finite set of polynomials $G$ [14, 15]. The finiteness of $G$ relies on Hilbert's Basis Theorem [24]. This implies that there exists a particular degree $d$ for which $G \in \mathcal{M}_d$, which leads to the following problem.

PROBLEM 6.1. *Find for a multivariate polynomial system $f_1, \ldots, f_s$ the degree $d_G$ such that for all $d \geq d_G : G \in \mathcal{M}_d$.*

The degree $d_G$ is for the general case related to $d_I$ and hence also has doubly exponential upper bounds [33, 37, 54]. Lazard proved the following useful theorem for a more practical case.

THEOREM 6.2. *([32, p.154-155]) Let $f_1, \ldots, f_s$ be multivariate polynomials of degrees $d_1, \ldots, d_s$ such that $d_1 \geq d_2 \ldots \geq d_s$. Suppose that a multiplicative monomial ordering is used and that the homogenized polynomial system $f_1^h, \ldots, f_s^h$ has a finite number of nontrivial projective roots. Then the polynomials of the reduced Gröbner basis have degrees at most $d_1 + \ldots + d_{n+1} - n + 1$ with $d_{n+1} = 1$ if $s = n$.*

This theorem provides a nice linear bound on the maximal degrees of the Gröbner basis. Unfortunately, this does not imply that $d_G \leq d_1 + \ldots + d_{n+1} - n + 1$ since $\mathcal{M}_d$ does not necessarily contain all polynomials of $\langle f_1, \ldots, f_s \rangle$ of degree $d$. The worst case we have encountered is for Example 4.4, for which $d_G = d_1 + \ldots + d_{n+1} - n + 1 + 2$. In order to determine whether a set of polynomials is a Gröbner basis one needs the notion of an S-polynomial.

DEFINITION 6.3. *([15, p. 83]) Let $f_1, f_2$ be nonzero multivariate polynomials and $x^\gamma$ the least common multiple of their leading monomials. The S-polynomial of $f_1, f_2$ is the combination*

$$S(f_1, f_2) \;=\; \frac{x^\gamma}{LT(f_1)}\, f_1 - \frac{x^\gamma}{LT(f_2)}\, f_2$$

*where $LT(f_1), LT(f_2)$ are the leading terms of $f_1, f_2$ with respect to a monomial ordering.*

It is clear from this definition that an S-polynomial is designed to produce cancellation of the leading terms and that it has a degree of at most $\deg(x^\gamma)$. A key component of Buchberger's Algorithm is constructing S-polynomials and computing

their remainder on division by a set of polynomials. It was Lazard [32] who had the insight that computing this remainder is equivalent with bringing a submatrix of the Macaulay matrix into triangular form. This led to Faugere's F4 and F5 algorithms [20, 21] which have become the golden standard to compute an exact Gröbner basis symbolically. The reduced polynomials $G(d)$ computed from Algorithm 5.1 ensure by definition that

$$\forall\, p \in \mathcal{M}_d \;\exists\, g \in G(d) \text{ such that } \mathrm{LM}(g) \mid \mathrm{LM}(p).$$

This implies that $G(d)$ is a Gröbner basis when $d \geq d_G$. Furthermore, it will be a reduced Gröbner basis. A criterion is needed to be able to decide whether $G(d)$ is a Gröbner basis. This is given by Buchberger's criterion, which we formulate in terms of the Macaulay matrix $M(d)$ and the reduced monomial system $A^\star(d)$.

THEOREM 6.4 (Buchberger's Criterion). *Let $f_1, \ldots, f_s$ be a multivariate polynomial system with reduced monomial system $A^\star(d)$ and reduced polynomials $G(d)$ for a given degree $d$. Then $G(d)$ is a Gröbner basis for $\langle f_1, \ldots, f_s \rangle$ if $M(d^\star)$ has the same reduced leading monomials $A^\star(d)$ for a degree $d^\star$ such that all S-polynomials of $G(d)$ lie in $\mathcal{M}_{d^\star}$ .*

*Proof.* Saying that $M(d^\star)$ has the same reduced leading monomials $A^\star(d)$ is equivalent with saying that all S-polynomials have a zero remainder on division by $G(d)$. This is exactly the stop-criterion for Buchberger's Algorithm [15, p.85]. ∎

Note that it follows from Buchberger's Criterion that for all degrees $d \geq d_G$, $G(d)$ and $A^\star(d)$ lead to the same reduced canonical decomposition. This implies that for all degrees $d \geq d_G$ the reduced canonical decomposition will not change anymore and results in the following useful corollary.

COROLLARY 6.5. *Let $f_1, \ldots, f_s$ be a multivariate polynomial system with a finite number of affine roots. Then $\forall d \geq d_G$ its reduced monomial set $A^\star(d)$ will contain for each variable $x_i\,(1 \leq i \leq n)$ a pure power. Furthermore, $B^\star(d)$ is then the affine normal set.*

*Proof.* This follows from Theorem 5.4 and Buchberger's Criterion that $\forall\, d \geq d_G$ both $M_G(d)$ and $M_{A^\star}(d)$ have the same reduced monomial decomposition. ∎

If it is known that the affine solution set of a polynomial ideal is zero-dimensional, then detecting pure powers in $A^\star(d)$ allows to determine the degree $d_G$. This means that by simply computing the reduced canonical decomposition it is possible to know $d_G$ without explicitly computing a Gröbner basis or any S-polynomials. Once $d_G$ is known, it then becomes possible to numerically compute all affine roots by solving an eigenvalue problem. This can be done in fact without the need of specifying a particular normal set as we will illustrate in the section on border bases.

EXAMPLE 6.1. *Again, we revisit the polynomial system in $\mathcal{C}_4^3$ from Example 4.4. Note that for this polynomial system $s = n$ and therefore $d_1 + d_2 + d_3 - n + 1 = 8$. We assume the polynomial system has a zero-dimensional solution set and start to compute the reduced canonical decomposition from $d = 4$. Algorithm 5.1 returns $A^\star(4) = \{x_1\, x_3, x_1^3\, x_2, x_2^4\}$, which already contains 1 pure power: $x_2^4$. The next pure power, $x_1^5$, is retrieved for $d = 7$ in $A^\star(7) = \{x_1\, x_3, x_1^3\, x_2, x_2^4, x_2\, x_3^3, x_1^5\}$. The last pure power, $x_3^5$, is found for $d = d_G = 10$. The Gröbner basis is therefore $G(10)$ as given in Example 5.1. Indeed, computing an exact Gröbner basis in Maple and normalizing each polynomial results in $G(10)$.*

The ill-posedness of the canonical decomposition under the influence of noise directly affects the computation of a Gröbner basis. As shown by Nagasaka in [42], it is impossible to define an approximate Gröbner basis in the same sense as an

approximate GCD or approximate factorization of multivariate polynomials. In [50], a normal set is computed from a numerical implementation of Buchberger's algorithm, e.g. Gaussian elimination, and the author states that "it appears impossible to derive strict and realistic thresholds" for the elimination pivots. This notion of a threshold for the elimination pivots is replaced in Algorithms 4.1 and 5.1 by a tolerance for both the rank test of $M(d)$ and the principal angles. Gröbner basis polynomials also typically have large integer coefficients. It is even possible that these coefficients fall out of the range of the double precision standard. In this case, it would be necessary to perform the computations in higher precision.

**7. Solving the Ideal Membership problem.** Solving the ideal membership problem for a non-homogeneous polynomial $p$ is a rank test of the Macaulay matrix as in (3.7) for a sufficiently large degree. We can now express $d_I$ in terms of $d_G$ in the following way.

THEOREM 7.1. *Consider the ideal membership problem as described in Problem 3.1. Let $G = \{g_1, \ldots, g_k\}$ be a Gröbner basis of $\langle f_1, \ldots, f_s \rangle$ and*

$$G_p \;=\; \{g \in G \;:\; LM(g) \mid LM(p)\} \text{ and } d_0 = \max_{g \in G_p} deg(g).$$

*Then*

$$d_I \;=\; d_G + deg(p) - d_0. \tag{7.1}$$

*Proof.* Since $G$ is a Gröbner basis $\exists\ g \in G : \mathrm{LM}(g) \mid \mathrm{LM}(p)$ and $G_p$ is therefore never empty. Determining whether $p \in \langle f_1, \ldots, f_s \rangle$ is equivalent with checking whether the remainder of $p$ on division by $G$ is zero. Determining this remainder is equivalent with the reduction of the matrix $\begin{pmatrix} M(d)^T & p^T \end{pmatrix} Q$ to triangular form for a sufficiently large $d$ with $Q$ the column permutation as described in Section 4. Suppose that $g \in G_p$ and $\deg(g) = d_0$. The degree $d_I$ as in (7.1) is then such that it guarantees that $\frac{\mathrm{LM}(p)}{\mathrm{LM(g)}}\ g \in \mathcal{M}_{d_I}$. In the first division of the multivariate division algorithm to compute the remainder, $p$ will be updated to $p \leftarrow p - g\,\mathrm{LM}(p)/\mathrm{LM(g)}$. The multivariate division algorithm guarantees that the new $p$ will have a smaller multidegree (according to the monomial ordering) [15, p.65]. In the next division step, another $g \in G$ such that $\mathrm{LT}(g)$ divides $\mathrm{LT}(p)$ is required. Since $p$ has a smaller multidegree, the new $g$ is also guaranteed to lie in $\mathcal{M}_{d_I}$. Therefore, all remaining steps of the division algorithm can be performed within $\mathcal{M}_{d_I}$ and the ideal membership problem can be solved. ∎

Theorem 7.1 means that in practice one can recursively compute the reduced canonical decomposition of $M(d)$ using Algorithm 5.1, do the rank test for the ideal membership problem and increase the degree as long as the rank test fails. At some point $d_G$ can be determined and the iterations can stop as soon as $d = d_G + \deg(p) - d_0$. The complexity of solving the ideal membership problem is $O(p(d)q(d)^2)$, since it is a rank-test of the Macaulay matrix $M(d)$. The worst case complexity is achieved when the degree of $p$ is larger than $d_G$, since $d_G$ could be doubly exponential. This would make the ideal membership problem infeasible. In practice, when the polynomial system $f_1, \ldots, f_s$ has a finite number of roots, and $\deg(p) \leq d_G$, then (7.1) acts as an upper bound on the degree for which the ideal membership problem can be solved. An obvious example is $p = f_1$.

EXAMPLE 7.1. *As already mentioned in Example 3.2 the given polynomial* $p = 867\,x_1^5 - 1560\,x_3\,x_2\,x_1 - 2312\,x_2^2\,x_1 + 1560\,x_3\,x_1^2 + 2104\,x_2\,x_1^2 - 1526\,x_1^3 + 4896\,x_2 - 2295\,x_1$

*lies in $\mathcal{M}_{11}$. At $d = 11$ all pure powers are found in $A^\star(11)$ which implies that the polynomial system has a finite affine solution set and $d_G = 11$. The rank test also succeeds, the numerical rank for both matrices at $d = d_I$ in (3.7) is 300.*

**8. Border Bases.** As mentioned earlier, insisting that the monomials of $A(d)$ are leading monomials with respect to a monomial ordering unavoidably leads to the representation singularity. The concept of border bases resolves the representation singularity for polynomial systems with a finite number of affine roots. More information on their properties and computation can be found in [26, 27, 30], with applications in [23, 31]. The key feature of border bases that solves the representation singularity is that they vary continuously under perturbations of the coefficients of the polynomial system. Before demonstrating this continuous change, we first define the border of a given reduced normal set $B^\star(d)$ and its corresponding pre-border basis.

DEFINITION 8.1. *([30, p. 422]) For a given reduced normal set $B^\star(d) = \{b_1, \ldots, b_m\}$, its border is*

$$\partial B^\star(d) \;=\; \{x_i\, b \mid 1 \leq i \leq n,\, b \in B^\star(d)\} \setminus B^\star(d)$$

*and its border prebasis $BB(d)$ ([30, p. 424]) is the set of polynomials*

$$BB(d) \;=\; \{bb_j \mid bb_j = t_j - \sum_{i=1}^{m} \alpha_i\, b_i \;,\; 1 \leq j \leq \mu\} \tag{8.1}$$

*with $\partial B^\star(d) = \{t_1, \ldots, t_\mu\}$.*

The polynomials of (8.1) are then a border basis for $B^\star(d)$ when they generate the polynomial ideal $\langle f_1, \ldots, f_s \rangle$ and the residue classes of the monomials in $B^\star(d)$ are a basis for the finite dimensional vector space $\mathcal{C}_d^n / I$. This can be summarized by a Buchberger's criterion for border bases [26, 27], or alternatively as commutation relations between multiplication matrices [39]. By the same argument as for the Gröbner basis it then holds that (8.1) is a border basis for a large enough degree $d$. Algorithm 4.1 can be adapted in a straightforward manner to numerically compute the $B^\star(d)$-border prebasis $BB(d)$ for a given reduced normal set $B^\star(d)$. Since each polynomial of $BB(d)$ lies in $\text{span}(B^\star(d), t)$, with $t \in \partial B^\star(d)$, one simply needs to replace the monomial $x^a$ by a monomial $t \in \partial B^\star(d)$. The whole algorithm is summarized in pseudo-code in Algorithm 8.1 and implemented in the PNLA package as getBB.m. Notice that, in contrast to the algorithms to compute the canonical and reduced canonical decomposition, each iteration of the for loop can run independently. Numerical algorithms to compute border bases such as in [23] rely on Gaussian elimination, which might be problematic for accurate rank determination as demonstrated in Example 4.4.

ALGORITHM 8.1. *Computation of a $B^\star(d)$-border prebasis $BB(d)$*
***Input**: orthogonal basis $V_2$, normal set $B^\star(d)$*
***Output**: $B^\star(d)$-border prebasis $BB(d)$*

```
BB(d) ← ∅
∂B*(d) ← border monomials of B*(d)
for all t ∈ ∂B*(d) do
    construct E from B*(d) and t
    [W S Z] ← SVD(V_2^T E^T)
    τ ← tolerance (4.4)
    if arcsin(μ_m) < τ then
        append v_m^T to BB(d)
```

***end if***
***end for***

Using Algorithm 8.1, we can now demonstrate that border bases avoid the representation singularity under perturbations of the coefficients of $f_1, \ldots, f_s$.

EXAMPLE 8.1. *Consider the polynomial system ([30, p. 430])*

$$F = \begin{cases} f_1 &= \frac{1}{4}\,x_1^2 + x_2^2 - 1, \\ f_2 &= x_1^2 + \frac{1}{4}\,x_2^2 - 1, \end{cases}$$

*and its slightly perturbed version*

$$\tilde{F} = \begin{cases} \tilde{f}_1 &= \frac{1}{4}\,x_1^2 + 10^{-5}x_1x_2 + x_2^2 - 1, \\ \tilde{f}_2 &= x_1^2 + 10^{-5}x_1x_2 + \frac{1}{4}\,x_2^2 - 1. \end{cases}$$

*Computing the reduced normal set $B^\star(3)$ for $F$ using Algorithm 5.1 results in $B^\star(3) = \{1, x_1, x_2, x_1x_2\}$. Applying Algorithm 8.1 and scaling the bb polynomials such that each leading term is monic results in the prebasis*

$$BB(3) = \begin{cases} bb_1 &= -0.8000 + x_1^2, \\ bb_2 &= -0.8000 + x_2^2, \\ bb_3 &= -0.8000x_2 + x_1^2x_2, \\ bb_4 &= -0.8000x_1 + x_1\,x_2^2, \end{cases}$$

*which is also a border basis for $\langle f_1, f_2\rangle$. Now, applying Algorithm 8.1 for the perturbed polynomial system $\tilde{F}$, using the same reduced normal set $B^\star(3)$, returns the following prebasis*

$$\tilde{BB}(3) = \begin{cases} \tilde{bb}_1 &= -0.8 + 8\times 10^{-06}x_1x_2 + x_1^2, \\ \tilde{bb}_2 &= -0.8 + 8\times 10^{-06}x_1x_2 + x_2^2, \\ \tilde{bb}_3 &= 6.4\times 10^{-6}x_1 - 0.800\,x_2 + x_1^2x_2, \\ \tilde{bb}_4 &= -0.800\,x_1 + 6.4\times 10^{-6}\,x_2 + x_1\,x_2^2. \end{cases}$$

*Note that the $-0.800$ terms in $\tilde{bb}_3$ and $\tilde{bb}_4$ have more nonzero digits, which is indicated by the trailing zeros. One can now see that the introduction of the noisy $x_1x_2$ term did not lead to any discontinuous jump from $BB(3)$ to $\tilde{BB}(3)$. The continuous change of the prebasis can be demonstrated by using symbolical computations. Replacing the coefficient of $x_1x_2$ in $\tilde{F}$ by $\epsilon$ and computing the prebases symbolically with respect to the degree negative lex ordering results in a prebasis*

$$BB(3) = \begin{cases} bb_1 &= -\frac{4}{5} + x_1^2, \\ bb_2 &= -\frac{4}{5} + x_2^2, \\ bb_3 &= -\frac{4}{5}x_2 + x_1^2x_2, \\ bb_4 &= -\frac{4}{5}x_1 + x_1x_2^2, \end{cases}$$

*for $F$ and*

$$\tilde{BB}(3) = \begin{cases} bb_1 &= -\frac{4}{5} + \frac{4}{5}\epsilon x_1x_2 + x_1^2, \\ bb_2 &= -\frac{4}{5} + \frac{4}{5}\epsilon x_1x_2 + x_2^2, \\ bb_3 &= \frac{16\epsilon}{16\epsilon^2-25}x_1 - \frac{20}{16\epsilon^2-25}x_2 + x_1^2x_2, \\ bb_4 &= -\frac{20}{16\epsilon^2-25}x_1 + \frac{16\epsilon}{16\epsilon^2-25}x_2 + x_1x_2^2, \end{cases}$$

*for $\tilde{F}$. From these symbolic expressions it is seen that $\tilde{BB}(3)$ changes continuously into $BB(3)$ when $\epsilon$ goes to zero. Note that setting $\epsilon = 10^{-5}$ in these symbolic expressions results in the numerical prebases computed by Algorithm 8.1.*

We have seen that the computation of a canonical decomposition is ill-posed under perturbations of the coefficients of $f_1, \ldots, f_s$. Nonetheless, the reduced canonical decomposition still allows to compute all affine roots of a polynomial system. Although it is guaranteed that the monomial sets $A(d), B(d)$ will change under perturbations of the coefficients, their cardinality however will not. This is due to the continuity of polynomial zeros [50, p. 304]. In other words, the total number of monomials in $B^\star(d)$ still represents the total number of affine roots for $d \geq d_G$. The Stetter matrix approach to compute all affine roots [1, 50] is to determine the normal set, write out multiplication matrices and find the components of the roots as the eigenvalues of these multiplication matrices. We will now demonstrate, using Example 8.1, how the affine roots can be computed without explicitly choosing a normal set. This is only an illustration and not the main scope of this article. More details and a formal description of our root-finding approach can be found in [16].

EXAMPLE 8.2. *The polynomial system $F$ from Example 8.1 has the following 4 affine roots*

$$(\frac{2\sqrt{5}}{5}, \pm\frac{2\sqrt{5}}{5}), (-\frac{2\sqrt{5}}{5}, \pm\frac{2\sqrt{5}}{5}).$$

*We will now compute the affine roots of the perturbed system $\tilde{F}$ without an explicit selection of a reduced normal set. Computing the reduced canonical decomposition for $\tilde{F}$ results in $A^\star(3) = \{x_1^3, x_1x_2, x_2^2\}$ and $B^\star(3) = \{1, x_1, x_2, x_1^2\}$. Since $A^\star(3)$ contains the pure powers $x_1^3, x_2^2$, we know that $d_G = 3$. We also have that $c(3) = 4$ and therefore the kernel $K(3)$ of $M(3)$ is spanned by 4 linear functionals that evaluate the rows of $M(3)$ in each of the 4 affine roots. Hence, each column of $K(3)$ will have a multivariate Vandermonde structure*

$$k = \begin{pmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 & x_1^3 & x_1^2x_2 & x_1x_2^2 & x_2^3 \end{pmatrix}^T.$$

*Observe that the Vandermonde structure allows us to write the following relation*

$$\begin{pmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 & \end{pmatrix}^T x_1 = \begin{pmatrix} x_1 & x_1^2 & x_1x_2 & x_1^3 & x_1^2x_2 & x_1x_2^2 \end{pmatrix}^T, \tag{8.2}$$

*which can be rewritten as*

$$S_1\, k\, x_1 = S_{x_1}\, k,$$

*where $S_1$ selects the 6 upper rows of $k$ and $S_{x_1}$ selects the rows of $k$ corresponding with the right-hand side of (8.2). This 'shift property' can be written for the complete kernel $K(3)$ as*

$$S_1\, K(3)\, D = S_{x_1}\, K(3), \tag{8.3}$$

*where $D$ is now a diagonal matrix containing the $x_1$ component of each corresponding affine root. $K(3)$ is not known. But it is possible to compute a numerical basis $N$ for the kernel from the SVD of $M(3)$. $N$ will obviously be related to $K(3)$ by a nonsingular $T$ as*

$$K(3) = N\, T. \tag{8.4}$$

*Substituting* (8.4) *into* (8.3) *results in*

$$S_1\ N\ T\ D\ =\ S_{x_1}\ N\ T.$$

*This can be written as the standard eigenvalue problem*

$$T\ D\ T^{-1}\ =\ (S_1\ N)^{\dagger}\ S_{x_1}\ N.$$

*Once the eigenvectors* $T$ *are computed, the kernel* $K(3)$ *can be computed from* (8.4)*, from which the affine roots can be read off after proper scaling. Note that in this approach no normal set was used to construct multiplication matrices and hence the representation singularity was of no concern. The reduced canonical decomposition was used only to determine* $d_G$*. Furthermore, in the case that there also roots at infinity, then the number of monomials in* $B^{\star}(d)$ *will inform us of the total number of affine roots. Applying this root-finding routine on* $\tilde{F}$ *returns the 4 affine roots*

$$(-0.8944, \pm 0.8944),\ (0.8944, \pm 0.8944),$$

*Each of these roots have relative forward errors of* $10^{-6}$*, which indicates that they are well-conditioned.*

**9. Conclusions.** This article introduced the canonical decomposition of the vector space $\mathcal{C}_d^n$. An SVD-based algorithm was presented which computes both the canonical and reduced decomposition reliably. It was also shown how under the presence of noise the problem of finding the canonical decomposition is ill-posed. This was resolved by introducing border prebases and an SVD-based algorithm to compute them. Furthermore, the link between the polynomials $G(d)$ and a Gröbner basis was made. This resulted in a new condition to determine the zero-dimensionality of the affine solution set. It was also shown how the ideal membership problem can be solved by means of a rank test and how the affine roots can be computed without an explicit choice of a normal set.

**Acknowledgments.** The authors would like to thank the Associate Editor Jörg Liesen and the anonymous referees for the many constructive comments.

REFERENCES

[1] W. Auzinger and H. J. Stetter, *An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations*, in Int. Conf. on Numerical Mathematics, Singapore 1988, Birkhuser ISNM 86, 1988, pp. 11–30.

[2] D.J. Bates, J.D. Hauenstein, A.J. Sommese, and C.W. Wampler, *Numerically Solving Polynomial Systems with Bertini*, SIAM, 2013.

[3] K. Batselier, P. Dreesen, and B. De Moor, *The Geometry of Multivariate Polynomial Division and Elimination*, SIAM Journal on Matrix Analysis and Applications, 34 (2013), pp. 102–125.

[4] ———, *A fast iterative orthogonalization scheme for the Macaulay matrix*, Journal of Computational and Applied Mathematics, 267 (2014), pp. 20–32.

[5] D. Bini and V. Y. Pan, *Polynomial and Matrix Computations (Vol. 1): Fundamental Algorithms*, Birkhauser Verlag, Basel, Switzerland, Switzerland, 1994.

[6] D. A. Bini and P. Boito, *Structured matrix-based methods for polynomial ϵ-gcd: analysis and comparisons*, in Proceedings of the 2007 International Symposium on Symbolic and Algebraic Computation, ISSAC '07, New York, NY, USA, 2007, ACM, pp. 9–16.

[7] Å. Björck and G. H. Golub, *Numerical Methods for Computing Angles Between Linear Subspaces*, Mathematics of Computation, 27 (1973), pp. pp. 579–594.

[8] P. Boito, *Structured Matrix Based Methods for Approximate Polynomial GCD*, Edizioni della Normale, 2011.

[9] D. Bondyfalat, B. Mourrain, and V.Y. Pan, *Computation of a specified root of a polynomial system of equations using eigenvectors*, Linear Algebra and its Applications, 319 (2000), pp. 193–209. Annual International Symposium on Symbolic and Algebraic Computation (ISSAC 98), Rostock, Germany.
[10] N. K. Bose, *Applied Multidimensional Systems Theory*, Van Nostrand Reinhold, 1982.
[11] B. Buchberger, *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*, PhD thesis, Mathematical Institute, University of Innsbruck, Austria, 1965.
[12] B. Buchberger, *Gröbner Bases and Systems Theory*, Multidimensional Systems and Signal Processing, 12 (2001), pp. 223–251.
[13] R. M. Corless, P. M. Gianni, B. M. Trager, and S. M. Watt, *The Singular Value Decomposition for Polynomial Systems*, in ACM International Symposium on Symbolic and Algebraic Computation, 1995, pp. 195–207.
[14] D. A. Cox, J. B. Little, and D. O'Shea, *Using Algebraic Geometry*, Graduate Texts in Mathematics, Springer-Verlag, Berlin-Heidelberg-New York, March 2005.
[15] ———, *Ideals, Varieties and Algorithms*, Springer-Verlag, third ed., 2007.
[16] P. Dreesen, *Back to the Roots: Polynomial System Solving Using Linear Algebra*, PhD thesis, Faculty of Engineering, KU Leuven (Leuven, Belgium), 2013.
[17] J. W. Eaton, D. Bateman, and S. Hauberg, *GNU Octave Manual Version 3*, Network Theory Ltd., 2008.
[18] I. Z. Emiris, *A General Solver Based on Sparse Resultants*, CoRR, abs/1201.5810 (2012).
[19] I. Z. Emiris, A. Galligo, and H. Lombardi, *Certified approximate univariate GCDs*, Journal of Pure and Applied Algebra, 117 118 (1997), pp. 229 – 251.
[20] J.-C. Faugère, *A new efficient algorithm for computing Gröbner bases (F4)*., Journal of Pure and Applied Algebra, 139 (1999), pp. 61–88.
[21] ———, *A new efficient algorithm for computing Gröbner bases without reduction to zero (F5)*, in Proceedings of the 2002 international symposium on Symbolic and algebraic computation, ISSAC '02, New York, NY, USA, 2002, ACM, pp. 75–83.
[22] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 3rd ed., Oct. 1996.
[23] D. Heldt, M. Kreuzer, S. Pokutta, and H. Poulisse, *Approximate Computation of Zero-dimensional Polynomial Ideals*, J. Symb. Comput., 44 (2009), pp. 1566–1591.
[24] D. Hilbert, *Ueber die Theorie der algebraischen Formen*, Springer, 1890.
[25] G. F. Jónsson and S. A. Vavasis, *Accurate solution of polynomial equations using Macaulay resultant matrices*, Math. Comput., 74 (2004), pp. 221–262.
[26] A. Kehrein and M. Kreuzer, *Characterizations of border bases*, Journal of Pure and Applied Algebra, 196 (2005), pp. 251 – 270.
[27] ———, *Computing border bases*, Journal of Pure and Applied Algebra, 205 (2006), pp. 279 – 295.
[28] A.V. Knyazev and M.E. Argentati, *Principal Angles between Subspaces in an A-based Scalar Product: Algorithms and Perturbation Estimates*, SIAM Journal on Scientific Computing, 23 (2002), pp. 2008–2040.
[29] A. Kondratyev, *Numerical Computation of Groebner Bases*, PhD thesis, RISC, Johannes Kepler University Linz, 2003.
[30] M. Kreuzer and L. Robbiano, *Computational Commutative Algebra 2*, no. v. 1 in Computational commutative algebra, Springer, 2005.
[31] J-B Lasserre, M. Laurent, B. Mourrain, P. Rostalski, and P. Trébuchet, *Moment matrices, border bases and real radical computation*, J. Symb. Comput., 51 (2013), pp. 63–85.
[32] D. Lazard, *Gröbner-Bases, Gaussian elimination and resolution of systems of algebraic equations*, in EUROCAL, 1983, pp. 146–156.
[33] ———, *A note on upper bounds for ideal-theoretic problems*, Journal of Symbolic Computation, 13 (1992), pp. 231 – 233.
[34] T. Y. Li and Z. Zeng, *A rank-revealing method with updating, downdating, and applications*, SIAM Journal on Matrix Analysis and Applications, 26 (2005), pp. 918–946.
[35] F. S. Macaulay, *On some formulae in elimination*, Proc. London Math. Soc., 35 (1902), pp. 3–27.
[36] ———, *The algebraic theory of modular systems*, Cambridge University Press, 1916.
[37] E. W. Mayr and A. R. Meyer, *The complexity of the word problems for commutative semigroups and polynomial ideals*, Advances in Mathematics, 46 (1982), pp. 305 – 329.
[38] A. Morgan, *Solving polynomial systems using continuation for engineering and scientific problems*, Prentice-Hall, Englewood Cliffs, N.J., 1987.

[39] B. Mourrain, *A new criterion for normal form algorithms*, in Proc. AAECC, volume 1719 of LNCS, Springer, 1999, pp. 430–443.
[40] B. Mourrain and V. Y. Pan, *Multivariate polynomials, duality, and structured matrices*, Journal of Complexity, 16 (2000), pp. 110 – 180.
[41] K. Nagasaka, *A Study on Grobner Basis with Inexact Input*, in Computer Algebra in Scientific Computing, Proceedings, Gerdt, VP and Mayr, EW and Vorozhtsov, EV, ed., vol. 5743 of Lecture Notes in Computer Science, 2009, pp. 247–258.
[42] ———, *Backward Error Analysis of Approximate Gröbner Basis*. Preprint, 2012.
[43] L. Pachter and B. Sturmfels, eds., *Algebraic Statistics for Computational Biology*, Cambridge University Press, August 2005.
[44] V. Y. Pan, *Structured matrices and polynomials: unified superfast algorithms*, Springer-Verlag New York, Inc., New York, NY, USA, 2001.
[45] MATLAB R2012a, *The Mathworks Inc.*, 2012. Natick, Massachusetts.
[46] T. Sasaki, *A Theory and an Algorithm of Approximate Gröbner Bases*, in 2011 13th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), Dongming Wang, V. Negru, T. Ida, T. Jebelean, D. Petcu, S. Watt, and D. Zaharie, eds., IEEE Comput. Soc., 2011, pp. 23–30.
[47] T. Sasaki and F. Kako, *Term Cancellations in Computing Floating-Point Gröbner Bases*, in Computer Algebra in Scientific Computing, Gerdt, VP and Koepf, W and Mayr, EW and Vorozhtsov, EV, ed., vol. 6244 of Lecture Notes in Computer Science, 2010, pp. 220–231.
[48] K Shirayanagi, *An Algorithm to compute Floating-Point Gröbner Bases*, in Mathematical Computation with MAPLE V: Ideas and Applications, Lee, T, ed., 1993, pp. 95–106.
[49] H. J. Stetter, *Matrix eigenproblems are at the heart of polynomial system solving*, SIGSAM Bulletin, 30 (1996), pp. 22–5.
[50] ———, *Numerical Polynomial Algebra*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2004.
[51] J. J. Sylvester, *On a theory of syzygetic relations of two rational integral functions, comprising an application to the theory of Sturms function and that of the greatest algebraical common measure*, Trans. Roy. Soc. Lond., (1853).
[52] C. Traverso and A. Zanoni, *Numerical stability and stabilization of Groebner basis computation*, in ISSAC, 2002, pp. 262–269.
[53] J. Verschelde, *Algorithm 795: PHCpack: a general-purpose solver for polynomial systems by homotopy continuation*, ACM Trans. Math. Softw., 25 (1999), pp. 251–276.
[54] C. K. Yap, *A New Lower Bound Construction for Commutative Thue Systems with Applications*, Journal Of Symbolic Computation, 12 (1991), pp. 1–27.
[55] Z. Zeng, *A numerical elimination method for polynomial computations*, Theor. Comput. Sci., 409 (2008), pp. 318–331.
[56] Z. Zeng, *The closedness subspace method for computing the multiplicity structure of a polynomial system*, in Interactions of Classical and Numerical Algebraic, 2009.
[57] Z. Zeng and B. H. Dayton, *The approximate GCD of inexact polynomials*, in Proceedings of the 2004 international symposium on Symbolic and algebraic computation, ISSAC '04, New York, NY, USA, 2004, ACM, pp. 320–327.