Solving zero-dimensional polynomial systems: a practical method using Bezout matrices

Jean-Paul Cardinal

January 19, 2018

Abstract

Let $\mathbb{Q}[x]$ be the algebra of polynomials, with rational coefficients in the variables $x = x_1, \dots, x_n$. Given in $\mathbb{Q}[x]$ a zero-dimensional ideal $\langle f \rangle = \langle f_1, \dots, f_n \rangle$, we present a practical and efficient method to compute the algebraic structure of the quotient algebra $A = \mathbb{Q}[x]/\langle f \rangle$. The entire method consists in matrix computations. A set of experiments illustrate the method's effectiveness.

Contents

1 Univariate case

We recall some well-known facts about univariate polynomials; in this section we consider a polynomial $f = a_0x^d + \cdots + a_{d-1}x + a_d \in \mathbb{Q}[x]$ in the variable x, with rational coefficients; we denote by $A = \mathbb{Q}[x]/\langle f\rangle$ the quotient algebra of $\mathbb{Q}[x]$ by the ideal $\langle f\rangle$, and we denote indifferently by x the variable x, its projection on the quotient algebra A and the multiplication map $x : |h \mapsto xh$ defined on A. The special basis $\mathbf{x} = (1, x, \cdots, x^{d-1})$ of the vector space A is called the **monomial basis**.

1.1 Mutiplication maps

The multiplication map $x : |h \mapsto xh$ is an endomorphism of A which, when written in the monomial basis, has a matrix X called the **companion matrix** of f. The matrix X is Hessenberg and writes

$$X = \begin{bmatrix} 0 & \cdots & 0 & -a_d/a_0 \\ 1 & 0 & \cdots & -a_{d-1}/a_0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & -a_1/a_0 \end{bmatrix} \tag{1}$$

Proposition 1. *The characteristic polynomial of X is f.*

Remark 1. We deduce from Proposition 1 that the eigenvalues of X are the roots of f, taking account the multiplicities. Moreover, as the matrix X is Hessenberg, we can use reliable techniques, like the QR algorithm, to compute its eigenvalues. This gives a practical and fast method to compute numerical approximations of the roots of f.

Remark 2. If g_1, g_2 are two polynomials of $\mathbb{Q}[x]$ that are equal modulo f, then $g_1(X) = g_2(X)$; therefore, if $g \in A$, then the matrix $g(X)$ is defined without any ambiguity; this matrix does not depend on the choice of the particular representative of g, and it is the matrix of the map $g : |h \mapsto gh$, written in the monomial basis.

1.2 Bezout polynomials and Bezout matrices

As we have seen, the companion matrix can be used to calculate the roots of a univariate polynomial. Interestingly, this can be naturally extended to zero-dimensional multivariate systems; if we have n polynomials $f = f_1, \ldots, f_n$ in the variables $x = x_1, \ldots, x_n$, then we simply define the companion matrices as the matrices of the multiplication maps $x_j : |h \mapsto x_jh$ defined on the quotient algebra $A = \mathbb{Q}[x]/\langle f\rangle$, written in some basis of A. However, in the multivariate case, A has no canonical basis, and the companion matrices do not have an obvious form. We can nonetheless resolve this problem by using a family of matrices, the so-called Bezout matrices, that exists both in the univariate case and in the multivariate case, and that serve as intermediate matrices to construct the companion matrices. Let's introduce the Bezout matrices in the case of univariate polynomials.

Definition 1. Let $f \in \mathbb{Q}[x]$ be a fixed polynomial. Let us introduce a new variable y and let g be any other polynomial. The **Bezout polynomial** $\delta(g)$, or **bezoutian**, is defined as the polynomial in the two variables x, y

$$\delta(g) = \frac{f(x)g(y) - f(y)g(x)}{x - y}$$

This polynomial is of degree $m-1$ in both variables x, y, where m is the maximum of the degrees of f and g. If we write the bezoutian

$$\delta(g) = \sum_{\alpha,\beta=0,\cdots,m-1} b_{\alpha\beta} x^\alpha y^\beta \tag{2}$$

then the matrix of coefficients $B(g) = [b_{\alpha\beta}]$ is called the **Bezout matrix**.

Remark 3. The size of a Bezout matrix may be loosely defined; when working with several Bezout matrices, it may be desirable to pad some of them with extra columns or lines of zeros to get compatible sizes.

Remark 4. The Bezout poynomial $\delta(g)$ and the Bezout matrix $B(g)$ satisfy the following equality

$$\delta(g) = \mathbf{x} B(g) \mathbf{y}^T \tag{3}$$

with $\mathbf{x} = (1, x, \cdots, x^{m-1}) \in \mathbb{Q}[x]^m$ and $\mathbf{y} = (1, y, \cdots, y^{m-1}) \in \mathbb{Q}[y]^m$ are two vectors of monomials.

Example 1. We choose $f = x^2 - 3x + 2$ as the fixed polynomial, and we examine the two cases $g = 1$ and $g = x^3$. The Bezout polynomials are $\delta(1) = -3 + x + y$ and $\delta(x^3) = -2x^2 - 2xy - 2y^2 + 3x^2y + 3xy^2 - x^2y^2$. The Bezout matrices $B(1)$ et $B(x^3)$ appear when we write $\delta(1)$ and $\delta(x^3)$ as double-entry arrays indexed by the monomials $1, x, x^2$ and $1, y, y^2$.

$\delta(1)$	1	y	y^2
1	-3	1	0
x	1	0	0
x^2	0	0	0

$\delta(x^3)$	1	y	y^2
1	0	0	-2
x	0	-2	3
x^2	-2	3	-1

Proposition 2. *Let f be a fixed polynomial and g be another polynomial; if we denote by m the maximum of the degrees of f and g and put $\mathbf{x} = (1, x, \cdots, x^{m-1})$, then*

$$\mathbf{x} B(1) g = \mathbf{x} B(g) \tag{4}$$

where the equality must be understood componentwise in $\mathbb{Q}[x]^m$ and modulo f.

Proof. We rewrite $\delta(g)$ as

$$\delta(g) = g(x)\frac{f(x) - f(y)}{x - y} - f(x)\frac{g(x) - g(y)}{x - y}$$

$$\delta(g) = g(x)\delta(1) - f(x)\frac{g(x) - g(y)}{x - y}$$

This is an equality between elements of $\mathbb{Q}[x][y]$. If $h \in \mathbb{Q}[x][y]$ and $\beta \in \mathbb{N}$, we denote by $h_\beta \in \mathbb{Q}[x]$ the coefficient of y^β in the polynomial h; then

$$\delta(g)_\beta = g(x)\delta(1)_\beta - f(x)\left(\frac{g(x)-g(y)}{x-y}\right)_\beta$$

which holds in $\mathbb{Q}[x]$. Thus, we have $\delta(g)_\beta = g(x)\delta(1)_\beta$ modulo f; as this is true for all $\beta \in \mathbb{N}$, Equality (4) follows. $\square$

Remark 5. Each column of a Bezout matrix, when left-multiplied by $\mathbf{x}$, is a polynomial in the variable x; when this does not bring to confusion, we will think of columns of a Bezout matrix as elements of $\mathbb{Q}[x]$, and lines as elements of $\mathbb{Q}[y]$. Saying Proposition 2 differently : each column of $B(1)$, when multiplied by g, equals the column of same index of $B(g)$, modulo f.

Example 2. Returning to Example 1, Proposition 2 says that the following equalities hold, modulo f

$$\begin{aligned}(-3+x)x^3 &= -2x^2,\\ (1)x^3 &= -2x+3x^2,\\ (0)x^3 &= -2+3x-x^2\end{aligned}$$

Remark 6. If we work with lines instead of columns, then Proposition 2 says that $gB(1)\mathbf{y}^T = B(g)\mathbf{y}^T$, giving equalities in $\mathbb{Q}[y]/\langle f\rangle$.

1.3 Relation between Bezout matrices and the companion matrix

Given a fixed polynomial f, of degree d, the two bezoutians $\delta(1)$ and $\delta(x)$ write

$$\begin{array}{c|cccc} \delta(1) & 1 & y & \dots & y^{d-1} \\ \hline 1 & a_{d-1} & \dots & \dots & a_0 \\ x & a_{d-2} & \dots & a_0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{d-1} & a_0 & 0 & \dots & 0 \end{array} \qquad \begin{array}{c|cccc} \delta(x) & 1 & y & \dots & y^{d-1} \\ \hline 1 & -a_d & 0 & \dots & 0 \\ x & 0 & a_{d-2} & \dots & a_0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{d-1} & 0 & a_0 & \dots & 0 \end{array} \tag{5}$$

that make appear the two Bezout matrices $B(1)$, which is clearly invertible, and $B(x)$. These two matrices are specially important because they are related to the companion matrix:

Proposition 3. *The compagnon matrix X and the Bezout matrices $B(x), B(1)$ are related by the* **Barnett decomposition** *formula [2]*

$$X = B(x)B(1)^{-1} \tag{6}$$

Proof. Let us consider the two families of elements of the quotient algebra A

$$\begin{aligned}\mathbf{x}B(1) &= (a_{d-1}+a_{d-2}x+\dots+a_0x^{d-1},\cdots,a_1+a_0x,a_0).\\ \mathbf{x}B(x) &= (-a_d, a_{d-2}x+\dots+a_0x^{d-1},\cdots,a_0x)\end{aligned} \tag{7}$$

and put $\hat{\mathbf{x}} = \mathbf{x}B(1)$. As $B(1)$ is invertible, the family $\hat{\mathbf{x}}$ is a basis of the vector space A, called the **Horner basis**. From Proposition 2, we have $\hat{\mathbf{x}}x = \mathbf{x}B(1)$. By construction, $B(1)$ is the matrix of the Horner basis $\hat{\mathbf{x}}$ written on the monomial basis $\mathbf{x}$, and $B(x)$ is the matrix of the family $\hat{\mathbf{x}}x$ written on the monomial basis; $B(1)^{-1}B(x)$ is thus the matrix of the family $\hat{\mathbf{x}}x$ written on the Horner basis $\hat{\mathbf{x}}$. This means that the multiplication map $x : |h \mapsto xh$ is represented in the Horner basis $\hat{\mathbf{x}}$ by the matrix $B(1)^{-1}B(x)$; the multiplication map x is also represented in the monomial basis $\mathbf{x}$ by the matrix $B(1)(B(1)^{-1}B(x))B(1)^{-1} = B(x)B(1)^{-1}$. □

1.4 Barnett decomposition formula

The Barnett decomposition formula relates the companion matrix, representing the multiplication map x, to the Bezout matrices of the polynomials 1 and x; this can be naturally extended as follows. Let $g \in \mathbb{Q}[x]$ be any polynomial; the Bezout matrices $B(1)$ et $B(g)$ are related to the matrix $g(X)$ by the **general Barnett decomposition formula**

$$B(g)B(1)^{-1} = g(X) \tag{8}$$

given that if the sizes of $B(1)$ and $B(g)$ differ, then we must transform and resize $B(g)$ according to a procedure that we shall now explain on the following example.

Formula (8) is easily checked when the degree of g is smaller or equal to the degree of f, because $B(1)$ and $B(g)$ have the same size; for example, if $f = x^2 - 3x + 2$, then we have

$\delta(1)$	1	y
1	-3	1
x	1	0

$\delta(x)$	1	y
1	-2	1
x	1	0

$\delta(x^2)$	1	y
1	0	-2
x	-2	3

and

$$B(x)B(1)^{-1} = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} = X \qquad B(x^2)B(1)^{-1} = \begin{bmatrix} -3 & -6 \\ 2 & 7 \end{bmatrix} = X^2 \tag{9}$$

which is consistent with formula (8). On the other hand, if m, the degree of g is strictly larger than d, the degree of f, then the sizes of $B(g)$ and $B(1)$ differ, and the product $B(g)B(1)^{-1}$ no longer makes sense. This can be fixed by indexing the Bezout matrices by the same monomials, namely $\mathbf{x} = (1, x, \cdots, x^{m-1})$ and $\mathbf{y} = (1, y, \cdots, y^{m-1})$. For example, with f as above and $g = x^3$, we have

$\delta(1)$	1	y	y^2
1	-3	1	0
x	1	0	0
x^2	0	0	0

$\delta(x^3)$	1	y	y^2
1	0	0	-2
x	0	-2	3
x^2	-2	3	-1

In doing so, $B(1)$ is no longer invertible; the key to obtain simultaneously matrices of equal size and the invertibility of $B(1)$, is to reduce the bezoutians modulo f. Let's illustrate this process on the previous example, and write

$$\delta(x^3) = \begin{bmatrix}1 & x & x^2\end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 0 & -2 & 3 \\ -2 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ y \\ y^2 \end{bmatrix}$$

$$\delta(x^3) = \begin{bmatrix}1 & x & x^2\end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 0 & -2 & 3 \\ -2 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ y \\ y^2 \end{bmatrix}$$

$$\delta(x^3) = \begin{bmatrix}1 & x & 2-3x+x^2\end{bmatrix} \begin{bmatrix} 4 & -6 & 0 \\ -6 & 7 & 0 \\ -2 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ y \\ y^2 \end{bmatrix}$$

To sum up, we have post-multiplied the row vector $\begin{bmatrix}1 & x & x^2\end{bmatrix}$ by the Gauss transform

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

and pre-multiplied the two Bezout matrices $B(1)$ and $B(g)$ by P^{-1}. The bezoutians now write

$\delta(1)$	1	y	y^2
1	-3	1	0
x	1	0	0
$2-3x+x^2$	0	0	0

$\delta(x^3)$	1	y	y^2
1	4	-6	0
x	-6	7	0
$2-3x+x^2$	-2	3	-1

According to the relations (4) the third column of $\delta(x^3)$, $-2+3x-x^2$, is zero modulo f; we recognize the simple fact $-f=0$. Thus,

$$\delta(1) = \begin{bmatrix}1 & x\end{bmatrix} \begin{bmatrix} -3 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ y \end{bmatrix}$$

$$\delta(x^3) = \begin{bmatrix}1 & x\end{bmatrix} \begin{bmatrix} 4 & -6 \\ -6 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ y \end{bmatrix} + (2-3x+x^2)(-2+3y-y^2)$$

Now, the bezoutians, when reduced modulo f write

$\delta(1)$	1	y
1	-3	1
x	1	0

$\delta(x^3)$	1	y
1	4	-6
x	-6	7

We have obtained two Bezout matrices of equal size, with $B(1)$ invertible. If we compute the matrix ratio

$$B(x^3)B(1)^{-1} = \begin{bmatrix} -6 & -14 \\ 7 & 15 \end{bmatrix} = X^3 \tag{10}$$

then we see that it is consistent with the general Barnett decomposition formula (8).

Remark 7. Instead of a Gauss matrix, we may use any matrix that maps a given column vector to a column vector containing just one non-zero entry, such as, for example, a Householder orthogonal matrix. This is the choice made in the implementation of the practical method given in [8].

2 Multivariate case

For a univariate polynomial, the structure of the quotient algebra A is made the monomial basis and of the companion matrix; this matrix is obtained either by reading the coefficients of the given polynomial f, or by making the ration of the matrices $B(1), B(x)$; in contrast, for a multivariate polynomial system, neither a basis nor the companion matrices (matrices in the given basis of the multiplication maps $x_j : \big| \; h \mapsto x_j h \;$) are apparent on the coefficients of the given polynomials. It is, however, relatively easy to construct the Bezout matrices $B(1), B(x_1), \cdots, B(x_n)$, from which one can derive a basis of A and the related companion matrices $X_j, j = 1, \cdots, n$.

Let's start by setting the framework; given n polynomials $f_1, \cdots, f_n$ in the variables $x_1, \cdots, x_n$, with coefficients in $\mathbb{Q}$, we denote by

- $\mathbb{Q}[x]$ the ring of polynomials in the variables $x = x_1, \cdots, x_n$,
- $\langle f \rangle$ the idal generated by $f = f_1, \cdots, f_n$,
- $A = \mathbb{Q}[x]/\langle f \rangle$ l'algbre quotient

From now on we assume that the ideal $\langle f \rangle$ is **zero-dimensional**; that is, the vector space A is finite dimensional [7, p. 234]. This is always the case when $n = 1$.

2.1 Construction of Bezout polynomials and Bezout matrices

2.1.1 Extension of Definition 14 to the multivariate case

Definition 2. Let $x^\gamma = x_1^{\gamma_1} \cdots x_n^{\gamma_n} \in \mathbb{Q}[x]$ be some monomial. We introduce a new variable set $y = y_1, \cdots, y_n$ and we consider, for each couple of indices $i, j = 1 \cdots n$, the ratio

$$\delta_{i,j}(x^\gamma) = \frac{y_j^{\gamma_j} f_i(y_1, \cdots, y_{j-1}, x_j, \cdots, x_n) - x_j^{\gamma_j} f_i(y_1, \cdots, y_j, x_{j+1}, \cdots, x_n)}{x_j - y_j} \tag{11}$$

which is a polynomial in the variables x, y. We get a matrix of finite differences, something like a multivariate rate of increase

$$\Delta(x^\gamma) = (\delta_{ij}(x^\gamma))_{ij} \tag{12}$$

The **Bezout polynomial**, or **Bezoutian**, of the monomial x^γ is by definition

$$\delta(x^\gamma) = det(\Delta(x^\gamma)) \tag{13}$$

which belongs to $\mathbb{Q}[x,y]$. For a more general polynomial $g = \sum_\gamma g_\gamma x^\gamma \in \mathbb{Q}[x]$, this definition is extended by linearity

$$\delta(g) = \sum_\gamma g_\gamma \delta(x^\gamma)$$

The **Bezout matrix** $B(g) = [b_{\alpha\beta}]$ is then defined as the matrix of the coefficients of $\delta(g)$ expressed on the monomials $(x^\alpha y^\beta)_{0\leq\alpha,\beta}$ appearing in $\delta(g)$

$$\delta(g) = \sum_{0\leq\alpha,\beta} b_{\alpha\beta} x^\alpha y^\beta \tag{14}$$

If we denote by $\mathbf{x}$ and $\mathbf{y}$ the sets of all the monomials x^α et y^β that appear in 14, then we have the following relation, similar to (3)

$$\delta(g) = \mathbf{x} B(g) \mathbf{y}^T \tag{15}$$

The following example, from [4], will illustrate this construction

Example 3. We consider $n = 2$, $f_1 = x_1^2 + x_1x_2^2 - 1$, $f_2 = x_1^2x_2 + x_1$ and we are interested by the calculation of the Bezout matrices $B(1), B(x_1), B(x_2)$, which are useful for computing the companion matrices X_1, X_2. To begin with, let us compute the finite differences matrices, as defined in (12)

$$\Delta(1) = \begin{pmatrix} x_1 + x_2^2 + y_1 & x_2y_1 + y_1y_2 \\ 1 + x_1x_2 + x_2y_1 & y_1^2 \end{pmatrix}$$
$$\Delta(x_1) = \begin{pmatrix} 1 + x_1y_1 & x_2y_1 + y_1y_2 \\ 1 + x_1x_2 + x_2y_1 & y_1^2 \end{pmatrix}$$
$$\Delta(x_2) = \begin{pmatrix} x_1 + x_2^2 + y_1 & 1 - y_1^2 + x_2y_1y_2 \\ 1 + x_1x_2 + x_2y_1 & -y_1 \end{pmatrix}$$

whose determinants are the bezoutians

$$\delta(1) = -x_2y_1 - x_1x_2^2y_1 + x_1y_1^2 + y_1^3 - y_1y_2 - x_1x_2y_1y_2 - x_2y_1^2y_2$$
$$\delta(x_1) = y_1^2 - x_1x_2^2y_1^2 + x_1y_1^3 - x_1x_2y_1^2y_2$$
$$\delta(x_2) = -1 - x_1x_2 - x_1y_1 - x_2y_1 - x_2^2y_1 + x_1x_2y_1^2 + x_2y_1^3 - x_2y_1y_2 - x_1x_2^2y_1y_2 - x_2^2y_1^2y_2$$

The monomial sets appearing in these polynomials are $\mathbf{x} = (1, x_2, x_2^2, x_1, x_1x_2, x_1x_2^2)$ and $\mathbf{y} = (1, y_1, y_1y_2, y_1^2, y_1^2y_2, y_1^3)$; the Bezout matrices $B(1), B(x_1), B(x_2)$ appear when we write these bezoutians as double-entry arrays indexed by $\mathbf{x}, \mathbf{y}$

$\delta(1)$	1	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1			-1			1
x_2		-1			-1	
x_2^2						
x_1				1		
x_1x_2			-1			
$x_1x_2^2$		-1				

$\delta(x_1)$	1	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1				1		
x_2						
x_2^2						
x_1						1
x_1x_2					-1	
$x_1x_2^2$				-1		

$\delta(x_2)$	1	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1	-1					
x_2		-1	-1			1
x_2^2		-1			-1	
x_1		-1				
x_1x_2	-1			1		
$x_1x_2^2$			-1			

Remark 8. Contrasting with the univariate case, $\mathbf{x}$ and $\mathbf{y}$ are not bases of the vector space A. We will see that they are, however, generating sets and we will show how to extract bases from them.

2.1.2 Practical computation of the Bezout matrices

In the previous example, the matrices $\Delta(1), \Delta(x_1), \Delta(x_2)$ are of size 2 and their entries are polynomials in x_1, x_2; it is easy to calculate their determinant. When either the number of variables variables n, or the degree of the input polynomials f_i, increase, then this calculation becomes impractical because one cannot use the Gauss pivot algorithm to a matrix with polynomial entries. However, one can overcome this difficulty by applying the following evaluation-interpolation process

1. A priori estimation of the set of monomials $x^\alpha y^\beta$ appearing in the bezoutian $\delta(x_k)$
2. Evaluation of $\Delta(x_k)$ on an adequate set $U \times V$ of Fourier multi-points $u = (u_1, \cdots, u_n) \in U$ et $v = (v_1, \cdots, v_n) \in V$
3. For each $(u, v) \in U \times V$, numerical computation, by the Gauss pivot method, of the determinant $\Delta(x_k)(u, v)$.
4. Interpolation of the set of calculated values $\Delta(x_k)(u, v)$ by the desired polynomial $\delta(x_k)$.

Let us specify what are the monomials of $\delta(x_k)$ and the Fourier points used to evaluate de $\delta(x_k)$. Suppose the polynomial system f has multi-degree $(d_1, \cdots, d_n)$, that is to say, for all $i, j = 1..n$ the degree of f_i in the variable x_j is smaller or equal to d_j. We fix an integer k between 0 and n and we adopt the convention that $x_0 = 1$. It is easy to show that the polynomial $\delta(x_k)$ has multi-degree $(d_1, 2d_2, \cdots, nd_n)$ in the variable x and has multi-degree $(nd_1, (n-1)d_2, \cdots, d_n)$ in the variable y. To evaluate $\delta(x_k)$, we then choose the Fourier sets $U =$

$\prod_{j=1..n} U_j$ where U_j is the set of complex roots of $X^{jd_j} - 1$. We also choose $V = \prod_{j=1..n} V_j$ in such a way U_j et V_j are disjoint sets, so that the denominator of (11) never vanishes. This is realized, for example, when V_j is the set of complex roots of $X^{(n-j+1)d_j} - \theta_j$ avec $\theta_j = e^{i\pi/j}$. These considerations lead to the following algorithm providing the sets U and V.

Data: $d = (d_1, \cdots, d_n)$, multi-degree of polynomial system
Result: U, V, two sets of Fourier points
for $j = 1, \cdots, n$ **do**
  $U_j \leftarrow$ roots of $X^{jd_j} - 1$;
  $V_j \leftarrow$ roots of $X^{(n-j+1)d_j} - e^{i\pi/j}$;
end
$U \leftarrow \prod_{j=1..n} U_j$;
$V \leftarrow \prod_{j=1..n} V_j$;

Algorithm 1: Construction of U, V, two sets of Fourier points

Then, we evaluate the bezoutian $\delta(x_k)$ on the Fourier points $(u, v) \in U \times V$

Data: polynomial sytem f, index k
Result: matrix $C^{(k)}$ containing the evaluations $\delta(x_k)(u, v)$
$(d_1, \cdots, d_n) \leftarrow$ multi-degree of f;
Get U, V via Algorithm 1;
$D \leftarrow \prod_{j=1..n} jd_j$;
$C^{(k)} \leftarrow \textsc{zeros}(D, D)$;
for $(u, v) \in U \times V$ **do**
  $\Delta \leftarrow \textsc{zeros}(n, n)$;
  for $i, j = 1..n$ **do**
    $\Delta_{i,j} \leftarrow \delta_{i,j}(x_k)(u, v)$
  end
  $C^{(k)}_{u,v} \leftarrow \textsc{det}(\Delta)$
end

Algorithm 2: Evaluation of the bezoutian $\delta(x_k)$ on $U \times V$

Let us show, to conclude, how the Bezout matrix $B(x_k)$ is simply related to $C^{(k)}$. To simplify, we denote the bezoutian $\delta(x_k)$ by $\delta^{(k)}$ and the Bezout matrix $B(x_k)$ by $B^{(k)}$; recall that $C^{(k)}$ denotes the evaluation matrix of $\delta^(k)$ on $U \times V$. The matrix $B^{(k)} = \left[b^{(k)}_{\alpha\beta}\right]$ satisfies $\delta^{(k)}(x, y) = \sum_{\alpha,\beta} b^{(k)}_{\alpha\beta} x^\alpha y^\beta$, thus $C^{(k)}_{u,v} = \delta^{(k)}(u, v) = \sum_{\alpha,\beta} b^{(k)}_{\alpha\beta} u^\alpha v^\beta$; this writes as a matrix product $\left[C^{(k)}_{u,v}\right]_{u,v} = [u^\alpha]_{u,\alpha} \left[b^{(k)}_{\alpha,\beta}\right]_{\alpha,\beta} \left[v^\beta\right]^T_{v,\beta}$. If we define the Fourier matrices $F_u = [u^\alpha]_{u,\alpha}$ and $F_v = \left[v^\beta\right]_{v,\beta}$, then we get the evaluation-interpolation relation between matrices $B^{(k)}$ and $C^{(k)}$

$$C^{(k)} = F_u B^{(k)} F_v^T \tag{16}$$

Since U and V consist of Fourier points, F_u et F_v are unitary and $B^{(k)}$ writes as the matrix product

$$B^{(k)} = F_u^* C^{(k)} \overline{F_v} \tag{17}$$

The computation of the Bezout matrices, as described above, have been implemented in Numpy and can be found at [8].

2.2 Barnett decomposition formula and structure of the quotient algebra.

Since the ideal $\langle f \rangle$ is zero-dimensional, the dimension of the quotient algebra $A = \mathbb{Q}[\mathbf{x}]/\langle f \rangle$ is finite; we may look for some basis and its related companion matrices (matrices in the basis of the multiplication maps by $x_1, \cdots, x_n$). For this purpose, we shall adapt the process described in Section 1.4 but, before this, we shall specify a number of algebraic properties about the polynomial $\delta(1)$ and the Bezout matrices $B(x_k)$.

2.2.1 Algebraic properties of polynomial $\delta(1)$ and of matrix $B(1)$

The following properties are simple; for a proof, the interested reader may refer to [4]. As in Proposition 3, we define families of elements of A by forming the vector-matrix products

$$\hat{\mathbf{x}}_k = \mathbf{x} B(x_k), \quad k = 0 \cdots n \tag{18}$$

with the convention that $x_0 = 1$ and where $\mathbf{x}$ is the set of all the monomials x^α that appear in the bezoutians $\delta(1), \delta(x_1), \cdots, \delta(x_n)$. and .

Example 4. Following Example 3 we have

$$\begin{array}{rcl} \hat{\mathbf{x}}_0 & = & (0, -x_2 - x_1 x_2^2, -1 - x_1 x_2, x_1, -x_2, 1) \\ \hat{\mathbf{x}}_1 & = & (0, 0, 0, -1 - x_2^2, -x_1 x_2, x_1) \\ \hat{\mathbf{x}}_2 & = & (-1 - x_1 x_2, -x_2 - x_2^2 - x_1, -x_2 - x_1 x_2^2, x_1 x_2, -x_2^2, x_2) \end{array} \tag{19}$$

Proposition 4. *(see [4]). For all $k = 1 \cdots n$ we have*

$$\hat{\mathbf{x}}_0 x_k = \hat{\mathbf{x}}_k \tag{20}$$

These relations can be easily checked on Example 3.

So far, there has been a great similarity between the univariate case and the multivariate cases; however, there is one notable difference: in the multivariate case the families $\mathbf{x}$ and $\hat{\mathbf{x}}$ are, in general, no longer bases in the vector space A. We have, however, the weaker result (see [4]).

Proposition 5. *Both $\mathbf{x}$ and $\hat{\mathbf{x}}$ are generating families in A.*

2.2.2 Reduction process

The previous result is important because from we can construct the whole structure of the algebra A. Following the matrix handlings described in Section 1.4, we shall show how to compute a basis of A and the companion matrices X_k, from the generating families $\mathbf{x}$ and $\hat{\mathbf{x}}$ and the Bezout matrices $B(x_k), k = 0, \cdots, n$.

Let us illustrate this process on Example 3.

The first column of $B(x_1)$ is zero but that of de $B(x_2)$ is not; this gives the relation $1 + x_1x_2 = 0$, modulo I. Then, we right-multiply $\mathbf{x}$ by the Gauss matrix P whose thith column is $(1, 0, 0, 0, 1, 0)^T$ and left-multiply the Bezout matrices by P^{-1}; the bezoutians write

$B(1)$	1	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1						1
x_2		-1			-1	
x_2^2						
x_1				1		
$1+x_1x_2$			-1			
$x_1x_2^2$		-1				

$B(x_1)$	1	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1				1	1	
x_2						
x_2^2						
x_1						1
$1+x_1x_2$					-1	
$x_1x_2^2$				-1		

$B(x_2)$	1	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1				-1		
x_2		-1	-1			1
x_2^2		-1			-1	
x_1		-1				
$1+x_1x_2$	-1			1		
$x_1x_2^2$			-1			

As we have $1 + x_1x_2 = 0$ we remove the first column and the fifth row in the Bezout matrices; the bezoutians write

$B(1)$	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1					1
x_2	-1			-1	
x_2^2					
x_1			1		
$x_1x_2^2$	-1				

$B(x_1)$	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1			1	1	
x_2					
x_2^2					
x_1					1
$x_1x_2^2$			-1		

$B(x_2)$	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1			-1		
x_2	-1	-1			1
x_2^2	-1			-1	
x_1	-1				
$x_1x_2^2$		-1			

The second column of $B(1)$ is zero but that of $B(x_2)$ is not. This implies that $x_2 + x_1x_2^2 = 0$. We repeat the previous step with the Gauss matrix P whose fifth column is $(0,1,0,0,1)^T$; the bezoutians write

$B(1)$	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1					1
x_2				-1	
x_2^2					
x_1			1		
$x_2+x_1x_2^2$	-1				

$B(x_1)$	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1			1	1	
x_2			1		
x_2^2					
x_1					1
$x_2+x_1x_2^2$			-1		

$B(x_2)$	y_1	y_1y_2	y_1^2	$y_1^2y_2$	y_1^3
1			-1		
x_2	-1				1
x_2^2	-1			-1	
x_1	-1				
$x_2+x_1x_2^2$		-1			

As we heve $x_2+x_1x_2^2=0$, we can remove the second column and the fifth row in each Bezout matrix; the bezoutians write

$B(1)$	y_1	y_1^2	$y_1^2y_2$	y_1^3
1				1
x_2			-1	
x_2^2				
x_1		1		

$B(x_1)$	y_1	y_1^2	$y_1^2y_2$	y_1^3
1		1	1	
x_2		1		
x_2^2				
x_1				1

$B(x_2)$	y_1	y_1^2	$y_1^2y_2$	y_1^3
1		-1		
x_2	-1			1
x_2^2	-1		-1	
x_1	-1			

The first column of $B(1)$ is zero but that of $B(x_2)$ is not. This implies taht $x_2+x_2^2+x_1=0$. The new Gauss matrix is P whose fourth column is $(0,1,1,1)^T$; the bezoutians write

$B(1)$	y_1	y_1^2	$y_1^2y_2$	y_1^3
1				1
x_2		-1	-1	
x_2^2		-1		
$x_2+x_2^2+x_1$		1		

$B(x_1)$	y_1	y_1^2	$y_1^2y_2$	y_1^3
1		1	1	
x_2				
x_2^2				-1
$x_2+x_2^2+x_1$				1

$B(x_2)$	y_1	y_1^2	$y_1^2y_2$	y_1^3
1		-1		
x_2				1
x_2^2			-1	
$x_2+x_2^2+x_1$	-1			

As $x_2+x_2^2+x_1=0$, we remove the first column eand fourth row in each Bezout matrix; the bezoutians write

$B(1)$	y_1^2	$y_1^2y_2$	y_1^3
1			1
x_2	-1	-1	
x_2^2	-1		

$B(x_1)$	y_1^2	$y_1^2y_2$	y_1^3
1	1	1	
x_2	1		-1
x_2^2			-1

$B(x_2)$	y_1^2	$y_1^2y_2$	y_1^3
1	-1		
x_2			1
x_2^2		-1	

Matrix $B(1)$ is now invertible; the reduction process is completed. The dimension of A is 3. We observe that $\mathbf{x} = (1, x_2, x_2^2)$ et $\mathbf{y} = (y_1, y_1^2, y_1^3)$ are bases of A; the associated Horner bases are $\hat{\mathbf{x}} = (-x_2 - x_2^2, -x_2, 1)$ and $\hat{\mathbf{y}} = (y_1^3, -y_1^2 - y_1^2 y_2, -y_1^2)$. More generally we have ([4] p.57, [5], [6])

Proposition 6. *After the reduction process described above is completed, that is to say when $B(1)$ is invertible and all the matrices $B(x_k), k = 0, \cdots, n$ have the same size and are indexed by the same families $\mathbf{x}, \mathbf{y}$, then each family $\mathbf{x}, \mathbf{y}$ is a basis of A.*

Remark 9. Proposition 6 is guaranted only when the ideal is zero-dimensional; in this case, to complete the reduction process we just have to use zero-columns of $B(1)$ or, more generally, linear combinations of columns that vanish, i.e elements of the right kernel of $B(1)$. If, however, the ideal is not zero-dimensional, then our experiments show that the reduction process, using both the right-kernel and the left-kernel of $B(1)$, generally produces an interesting result.

2.2.3 Barnett formula and companion matrices

Following Example 3, we define the matrices X_1, X_2

$$X_1 = B(x_1)B(1)^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix}, \quad X_2 = B(x_2)B(1)^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix} \tag{21}$$

We see that X_1, X_2 are the multiplication matrices by the variables x_1, x_2 in the basis $\mathbf{x}$; these are the companion matrices associated to the basis $\mathbf{x}$. More generally, we have

Proposition 7. *When the reduction process has been completed and we have at our disposal Bezout matrices $B(x_j)$ and bases $\mathbf{x}, \mathbf{y}$, the companion matrices X_j, i.e the multiplication matrices by the variables x_1, x_2 in the basis $\mathbf{x}$, can be calculated by the* **Barnett formulas**

$$X_j = B(x_j)B(1)^{-1} \tag{22}$$

Remark 10. As in the univariate case, we have, for all $j = 1, \cdots, n$,
$B(x_j)^T B(1)^{-T}$ is the multiplication matrix by y_j in the basis $\mathbf{y}$
$B(1)^{-1} B(x_j)$ is the multiplication matrix by x_j in the basis $\hat{\mathbf{x}}$
$B(1)^{-T} B(x_j)^T$ is the multiplication matrix by y_j in the basis $\hat{\mathbf{y}}$

2.2.4 Numerical computation of the roots

As in the univariate case, (see Proposition 1), the roots of the polynomial system$f_1, \cdots, f_n$ are the eigenvalues of the companion matrices ([1]).

In Example 3,the eigenvalues of matrices X_1, X_2 are

x_1	x_2
-1.32472	0.75488
$0.66236 + 0.56228i$	$-0.87744 + 0.74486i$
$0.66236 - 0.56228i$	$-0.87744 - 0.74486i$

Since A is a commutative algebra, the matrices X_1, X_2 commute and have the same eigenvectors. We must be careful to sort the eigenvalues of X_1, X_2 so that they correspond to the same eigenvectors. In Example 3, it is easy to check that the couples (x_1, x_2) are numerical approximations of the roots of the polynomial system $f_1 = x_1^2 + x_1x_2^2 - 1, f_2 = x_1^2x_2 + x_1$.

2.3 Numerical experiment

In this experiment, we solve the particular polynomial system $f = [$
$x_0^3x_1x_2^3+2x_0x_1^3x_2^3+3x_0^3x_1^2x_2-4x_0^2x_1^2x_2-3x_0^2x_2^3+x_0^3x_1-x_0^3x_2-3x_0x_2^3-3x_0x_2^2+3x_0^2,$
$3x_0^2x_1^3 - 3x_0x_1^3x_2 - 2x_0x_1^2x_2^2 - 2x_0x_1^3 + 2x_0x_2^3 - 4x_1^3 + 2x_0^2x_2 - 4x_1x_2,$
$2x_0^3x_1^3x_2^2 + x_0^2x_1^3x_2^3 - 3x_0^3x_1^3x_2 + 2x_0^2x_1^3x_2^2 - 3x_0^3x_1x_2^3 + 2x_0^3x_2^3 - x_0^2x_2^3 - 4x_0^3 + 3x_1x_2$
$]$
The size of the Bezout matrix $B(1)$ is 120; this is the maximum number of solutions that a system of degree [1, 1, 1, 1, 1] can have. To initiate the reduction process, the rank of $B(1)$ is needed; as the matrix has integer coefficients, we use the Sage function matrix.kernel() to calculate its rank. This is the only computation that we do in exact arithmetic; all the subsequent computations are done in floating-point arithmetic. After the reduction process has been completed, we find that the dimension of the quotient A is 92. Since the computations have been done numerically, the Bezout matrices and the companion matrices are numerical matrices and the eigenvalues of the companion matrices $X_j = B(x_j)B(1)^{-1}$ are numerical approximations of the roots of the polynomial system f.

2.3.1 Quality of the results

To check the quality of the numerical roots α, we compute the errors $f(\alpha)$. These errors are shown in Table 1.

2.3.2 Timings

Table 2 shows the timings of the Bezout computations as compared to the timings of the Groebner computations.

References

[1] W. Auzinger, H. J. Stetter, An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equation, Numerical mathematics, Singapore 1998, ISNM vol. 86, Birkhuser, pp. 11-30

[2] S. Barnett, A note on the Bezoutian matrix, SIAM J. Appl. Math., 22:84-86, 1972

[3] G.H. Golub, C.F. van Loan, Matrix Computations, The Johns Hopkins University Press, 1989.

[4] J.P. Cardinal, Dualit et algorithmes itratifs pour la rsolution de systmes polynomiaux, Thse prsente devant l'universit de Rennes I, 1993.

[5] B. Mourrain, Bezoutian and quotient ring structure, J. of Symbolic Comput., 39:397-415, 2005.

[6] T. Mora, Solving polynomial equation systems, Cambridge University Press, 2015

[7] D. Cox, J. Little, D. O'Shea, Ideals, Varieties and Algorithms, Springer, 2006

[8] https://github.com/jpcp13/bezout

log10 of errors	nb of roots
$[-14.6, -13.8]$	25
$[-13.8, -13.0]$	33
$[-13.0, -12.2]$	17
$[-12.2, -11.4]$	9
$[-11.4, -10.6]$	5
$[-10.6, -9.8]$	1
$[-9.8, -9.0]$	2

Table 1: histogram of errors

Mthode	Computation	Software	Arithmetic	Timing
Bezout	Bezout matrices	NumPy	floating point	0.3485 s
	rank of $B(1)$ via rref()	Sage	integer	0.0258 s
	matrices reduction	Numpy	floating point	0.0371 s
	eigenvalues	SciPy	floating point	0.0634 s
Groebner	Groebner basis computation	Sage	integer	3.4288 s

Table 2: timings