Solving square polynomial systems: a practical method using Bezout matrices

Jean-Paul Cardinal
cardinal@math.univ-paris13.fr

July 29, 2018

Abstract

Let f be a polynomial system consisting of n polynomials $f_1, \cdots, f_n$ in n variables $x_1, \cdots, x_n$, with coefficients in $\mathbb{Q}$ and let $\langle f \rangle$ be the ideal generated by f. Such a polynomial system, which has as many equations as variables is called a square system. It may be zero-dimensional, i.e the system of equations $f = 0$ has finitely many complex solutions, or equivalently the dimension of the quotient algebra $A = \mathbb{Q}[x]/\langle f \rangle$ is finite. In this case, the companion matrices of f are defined as the matrices of the endomorphisms of A, called multiplication maps, $\left| \begin{array}{c} h \mapsto x_j h \end{array} \right.$, written in some basis of A. We present a practical and efficient method to compute the companion matrices of f in the case when the system is zero-dimensional. When it is not zero-dimensional, then the method works as well and still produces matrices having properties similar to the zero-dimensional case. The whole method consists in matrix calculations. An experiment illustrates the method's effectiveness.

Contents

1 Univariate case

We recall some well-known facts about univariate polynomials; in this section we consider a polynomial $f = a_0 x^d + \cdots + a_{d-1} x + a_d \in \mathbb{Q}[x]$ 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 the multiplicities into account. Moreover, since 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 n variables $x = x_1, \ldots, x_n$, then we simply define the companion matrices as the matrices of the multiplication maps $x_j : |h \mapsto x_j h$ 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 and let g be any other polynomial. We introduce a new variable y and define the **Bezout polynomial** $\delta(g)$, or **bezoutian**, as the polynomial in 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

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

where $\mathbf{x} = (1, x, \cdots, x^{m-1}) \in \mathbb{Q}[x]^m$ 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 $\delta(x^3) = -2x^2 - 2xy - 2y^2 + 3x^2 y + 3xy^2 - x^2 y^2$. 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

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

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

This equality holds 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; since this is true for all $\beta \in \mathbb{N}$, Equality (4) follows. $\qquad\square$

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

$$(-3 + x)x^3 = -2x^2,$$
$$(1)x^3 = -2x + 3x^2,$$
$$(0)x^3 = -2 + 3x - x^2$$

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 Bezout matrices are related to the companion matrix

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

$\delta(1)$	1	y	$\ldots$	y^{d-1}
1	a_{d-1}	$\ldots$	$\ldots$	a_0
x	a_{d-2}	$\ldots$	a_0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{d-1}	a_0	0	$\ldots$	0

$\delta(x)$	1	y	$\ldots$	y^{d-1}
1	$-a_d$	0	$\ldots$	0
x	0	a_{d-2}	$\ldots$	a_0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{d-1}	0	a_0	$\ldots$	0

$$\tag{5}$$

The corresponding arrays of coefficients are the 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 is related to the Bezout matrices $B(x), B(1)$ by the* **Barnett decomposition formula** *[2]*

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

Proof. Let us consider the two vectors of elements of A

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

and put $\hat{\mathbf{x}} = \mathbf{x}B(1)$. As $B(1)$ is invertible, $\hat{\mathbf{x}}$ is a basis of the vector space A^d, 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 $\hat{\mathbf{x}}x$ written on the monomial basis; $B(1)^{-1}B(x)$ is thus the matrix of $\hat{\mathbf{x}}x$ written on the Horner basis $\hat{\mathbf{x}}$. This means that the multiplication map $|h \mapsto xh$ is represented in the Horner basis $\hat{\mathbf{x}}$ by the matrix $B(1)^{-1}B(x)$, and is therefore represented in the monomial basis $\mathbf{x}$ by the matrix $B(1)(B(1)^{-1}B(x))B(1)^{-1} = B(x)B(1)^{-1}$. $\qquad\square$

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; matrix $g(X)$ is related to the Bezout matrices $B(1)$ et $B(g)$ by the so-called **general Barnett decomposition formula**

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

given that if the sizes of $B(1)$ and $B(g)$ differ then we must transform and resize $B(g)$ according to the following procedure, explained on an 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

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

thus

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

which is consistent with formula (8). However, if the degree m of g is strictly larger than the degree d 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 write both bezoutians modulo f:

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

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

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

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 $B(1)$ and $B(g)$ by P^{-1}; so the bezoutians 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 $\delta(x^3)$, which is $-2 + 3x - x^2$, is zero modulo f; we recognize the simple fact $-f = 0$. Thus,

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

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

So 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, that satisfy the general Barnett decomposition formula (8):

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

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; we use Householder matrices in the practical implementation given in [9].

2 Multivariate case

For a univariate polynomial f, the structure of the quotient algebra A consists of the monomial basis and the companion matrix; this matrix is obtained either by a direct reading of the coefficients of f, or by taking the ratio of the two matrices $B(1), B(x)$; in contrast, for a multivariate polynomial system, neither a basis nor the companion matrices (matrices of the multiplication maps $x_j : \mid h \mapsto x_j h$ in the basis) can be read from the coefficients of the given polynomials. It is easy, however, 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 $X_1, \cdots, X_n$.

Given n polynomials $f_1, \cdots, f_n$ in n 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 ideal generated by $f = f_1, \cdots, f_n$,

- $A = \mathbb{Q}[x]/\langle f \rangle$ the quotient algebra.

From now on we assume that $\langle f \rangle$ is **zero-dimensional**, that is to say the vector space A is finite dimensional [8, p. 234]. This is always the case when $n = 1$.

2.1 Construction of Bezout polynomials and Bezout matrices

2.1.1 Definition of Bezout polynomials and Bezout matrices in 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 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. With the $\delta_{i,j}$'s we form a finite difference matrix

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

whose determinant $\delta(x^\gamma) = det(\Delta(x^\gamma))$, a polynomial in both variables x, y, is called the **Bezout polynomial**, or **bezoutian**, of x^γ. This definition can be extended by linearity to a more general polynomial $g = \sum_\gamma g_\gamma x^\gamma \in \mathbb{Q}[x]$ by the formula

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

Writting $\delta(g)$ as a sum of monomials,

$$\delta(g) = \sum_{0 \leq \alpha, \beta} b_{\alpha\beta} x^\alpha y^\beta \tag{13}$$

we define $B(g)$, the **Bezout matrix** of g, as the matrix of the coefficients $B(g) = [b_{\alpha\beta}]$. If we denote by $\mathbf{x}$ and $\mathbf{y}$ the sets of all the monomials x^α et y^β that appear in (13), then we have the relation, similar to (3)

$$\delta(g) = \mathbf{x}B(g)\mathbf{y}^T \tag{14}$$

The following example (see [4] p.58) illustrates the previous definition

Example 3. We take $n = 2$, $f_1 = x_1^2 + x_1 x_2^2 - 1$, $f_2 = x_1^2 x_2 + x_1$ and want to calculate 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, we compute the finite difference matrices, as defined in (12)

$$\Delta(1) = \begin{pmatrix} x_1 + x_2^2 + y_1 & x_2 y_1 + y_1 y_2 \\ 1 + x_1 x_2 + x_2 y_1 & y_1^2 \end{pmatrix}$$

$$\Delta(x_1) = \begin{pmatrix} 1 + x_1 y_1 & x_2 y_1 + y_1 y_2 \\ 1 + x_1 x_2 + x_2 y_1 & y_1^2 \end{pmatrix}$$

$$\Delta(x_2) = \begin{pmatrix} x_1 + x_2^2 + y_1 & 1 - y_1^2 + x_2 y_1 y_2 \\ 1 + x_1 x_2 + x_2 y_1 & -y_1 \end{pmatrix}$$

whose determinants are the bezoutians

$\delta(1) = -x_2 y_1 - x_1 x_2^2 y_1 + x_1 y_1^2 + y_1^3 - y_1 y_2 - x_1 x_2 y_1 y_2 - x_2 y_1^2 y_2$

$\delta(x_1) = y_1^2 - x_1 x_2^2 y_1^2 + x_1 y_1^3 - x_1 x_2 y_1^2 y_2$

$\delta(x_2) = -1 - x_1 x_2 - x_1 y_1 - x_2 y_1 - x_2^2 y_1 + x_1 x_2 y_1^2 + x_2 y_1^3 - x_2 y_1 y_2 - x_1 x_2^2 y_1 y_2 - x_2^2 y_1^2 y_2$

The monomial appearing in these polynomials are $\mathbf{x} = (1, x_2, x_2^2, x_1, x_1 x_2, x_1 x_2^2)$ and $\mathbf{y} = (1, y_1, y_1 y_2, y_1^2, y_1^2 y_2, y_1^3)$; the Bezout matrices $B(1), B(x_1), B(x_2)$ ap-

8

pear when we write these bezoutians as double-entry arrays indexed by $\mathbf{x}, \mathbf{y}$

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

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

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

Remark 8. Contrasting with the univariate case, $\mathbf{x}$ and $\mathbf{y}$ are not bases of the vector space A. But, as we will see, they are generating sets and we will examine how to build bases from them.

2.1.2 Practical computation of the Bezout matrices

From now on, we will denote $\delta^{(k)}$ the Bezout polynomial $\delta(x_k)$ and $B^{(k)}$ the Bezout matrix $B(x_k)$; we will also adopt the convention that $x_0 = 1$. Accordingly, $\Delta(1), \Delta(x_1), \Delta(x_2)$ will be denoted by $\Delta^{(0)}, \Delta^{(1)}, \Delta^{(2)}$.

   In the previous example, the matrices $\Delta^{(0)}, \Delta^{(1)}, \Delta^{(2)}$ have size 2 and their entries are polynomials in x_1, x_2; calculating their determinant is easy. But, if either the number n of variables or the degree of the input polynomials f_i increase, then this calculation becomes impractical because one cannot apply the Gauss pivot algorithm to a polynomial entries matrix. We can however overcome this difficulty by applying the following evaluation-interpolation process :

1. Estimate a priori the set of monomials $x^\alpha y^\beta$ appearing in the Bezout polynomial $\delta(x_k)$ (for fixed k, $k = 0, \cdots, n,$). Specifically, if we suppose that the polynomial system f has multi-degree $(d_1, \cdots, d_n)$, that is to say, the partial degree in the variable x_j, for every $f_i, i = 1 \cdots n$, is smaller or equal to d_j, then the polynomial $\delta(x_k)$ has multi-degree $(d_1, 2d_2, \cdots, nd_n)$ in variable x and multi-degree $(nd_1, (n-1)d_2, \cdots, d_n)$ in variable y. Note that this estimate does not depend on $k = 0 \cdots n$.

2. Evaluate the polynomial-entries matrix $\Delta(x_k)$ on all pairs (u, v), $u = (u_1, \cdots, u_n) \in U$, $v = (v_1, \cdots, v_n) \in V$, where U, V are suitable sets of Fourier multi-points. We can choose, for instance, $U = \prod_{j=1..n} U_j$ where U_j is the set of complex roots of the polynomial $X^{jd_j} - 1$, and $V = \prod_{j=1..n} V_j$ in such a way that U_j et V_j are disjoint sets, so that the denominator of the ratio (11) never vanishes. This is obtained, for example, if V_j is the set of complex roots of the polynomial $X^{(n-j+1)d_j} - \theta_j$ with $\theta_j = e^{i\pi/j}$ (here i means the square root of -1).

3. For each $(u, v) \in U \times V$, apply the Gauss pivot method to the numerical matrix $\Delta(x_k)(u, v)$, of size $n \times n$, in order to compute its determinant $\delta(x_k)(u, v)$. This produces a matrix $C^{(k)}$, indexed by $(u, v) \in U \times V$, and defined by $C^{(k)}_{u,v} = \delta(x_k)(u, v)$.

4. Interpolate the set of numerical values $\delta(x_k)(u, v), (u, v) \in U \times V$ in order to obtain the Bezout polynomial $\delta^{(k)}$ and the Bezout matrix $B^{(k)}$. The matrices $B^{(k)}$ and $C^{(k)}$ are closely related. Indeed, $B^{(k)} = \left[b^{(k)}_{\alpha\beta} \right]$ 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 $C^{(k)}$

$$C^{(k)} = F_u B^{(k)} F_v^T \tag{15}$$

Since U and V consist of Fourier points, F_u et F_v are (unitary) Fourier matrices and $B^{(k)}$ appears to be the 2D DFT of $C^{(k)}$

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

The computation of the Bezout matrices, as described above, have been implemented in Numpy and can be found at [9].

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 of the multiplication maps by $x_1, \cdots, x_n$). For this purpose, we will adapt the process described in Section 1.4 but, before, we will specify a number of algebraic properties about polynomial δ^0 and matrices $B^{(k)}$.

2.2.1 Algebraic properties of Bezout polynomials and Bezout matrices

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

$$\hat{\mathbf{x}}_k = \mathbf{x}B^{(k)}, \quad k = 0\cdots n \tag{17}$$

where $\mathbf{x}$ is the set of all the monomials x^α appearing in the Bezout polynomials $\delta^{(k)}, \quad k = 0\cdots n$. We also write $\hat{\mathbf{x}}_0 = \hat{\mathbf{x}}$.

Example 4. Following Example 3 we have

$$\begin{array}{rcl}
\hat{\mathbf{x}} & = & (0, -x_2 - x_1 x_2^2, -1 - x_1 x_2, x_1, -x_2, 1) \\
\hat{\mathbf{x}}_1 & = & (0, 0, 0, -1 - x_2^2, -x_1 x_2, x_1) \\
\hat{\mathbf{x}}_2 & = & (-1 - x_1 x_2, -x_2 - x_2^2 - x_1, -x_2 - x_1 x_2^2, x_1 x_2, -x_2^2, x_2)
\end{array} \tag{18}$$

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

$$\hat{\mathbf{x}}x_k = \hat{\mathbf{x}}_k \tag{19}$$

This proposition can be easily checked on Example 3.

So far, the univariate and multivariate cases are very similar, except on one point : in the multivariate case the families $\mathbf{x}$ and $\hat{\mathbf{x}}$ are not, in general, bases of the vector space A. We have, however, the weaker result (see [4]) :

Proposition 5. *Both $\mathbf{x}$ and $\hat{\mathbf{x}}$ are generating families of the vector space A.*

2.2.2 Reduction process

Following the matrix handlings described in Section 1.4, we will show how to compute, starting from the generating families $\mathbf{x}, \hat{\mathbf{x}}$ and the Bezout matrices $B^{(k)}$, a basis of A and the companion matrices X_k. Let us illustrate this process on Example 3.

The first column is zero in $B^{(0)}$ but not in $B^{(2)}$; this gives, modulo $\langle f \rangle$, the relation $1 + x_1 x_2 = 0$. Then, we right-multiply $\mathbf{x}$ by the Gauss matrix P whose fifth column is $(1, 0, 0, 0, 1, 0)^T$ and left-multiply the Bezout matrices by P^{-1}; the Bezout polynomials write

$\delta^{(0)}$	1	y_1	$y_1 y_2$	y_1^2	$y_1^2 y_2$	y_1^3
1						1
x_2		-1		-1		
x_2^2						
x_1			1			
$1 + x_1 x_2$			-1			
$x_1 x_2^2$		-1				

$\delta^{(1)}$	1	y_1	$y_1 y_2$	y_1^2	$y_1^2 y_2$	y_1^3
1				1	1	
x_2						
x_2^2						
x_1						1
$1 + x_1 x_2$					-1	
$x_1 x_2^2$				-1		

$\delta^{(2)}$	1	y_1	$y_1 y_2$	y_1^2	$y_1^2 y_2$	y_1^3
1				-1		
x_2		-1	-1			1
x_2^2		-1			-1	
x_1		-1				
$1 + x_1 x_2$	-1				1	
$x_1 x_2^2$			-1			

Since $1 + x_1 x_2 = 0$, modulo $\langle f \rangle$, we can remove the fifth row in the Bezout matrices; the Bezout polynomials write

$\delta^{(0)}$	y_1	$y_1 y_2$	y_1^2	$y_1^2 y_2$	y_1^3
1					1
x_2	-1			-1	
x_2^2					
x_1			1		
$x_1 x_2^2$	-1				

$\delta^{(1)}$	y_1	$y_1 y_2$	y_1^2	$y_1^2 y_2$	y_1^3
1			1	1	
x_2					
x_2^2					
x_1				1	
$x_1 x_2^2$			-1		

$\delta^{(2)}$	y_1	$y_1 y_2$	y_1^2	$y_1^2 y_2$	y_1^3
1			-1		
x_2	-1	-1			1
x_2^2	-1			-1	
x_1	-1				
$x_1 x_2^2$		-1			

Now, the second column is zero in $B^{(0)}$ but not in $B^{(2)}$. We have $x_2 + x_1 x_2^2 = 0$, modulo $\langle f \rangle$. We repeat the previous step with the Gauss matrix P whose fifth column is $(0, 1, 0, 0, 1)^T$; the Bezout polynomials write

$\delta^{(0)}$	y_1	$y_1 y_2$	y_1^2	$y_1^2 y_2$	y_1^3
1					1
x_2				-1	
x_2^2					
x_1			1		
$x_2 + x_1 x_2^2$	-1				

$\delta^{(1)}$	y_1	$y_1 y_2$	y_1^2	$y_1^2 y_2$	y_1^3
1			1	1	
x_2		1			
x_2^2					
x_1					1
$x_2 + x_1 x_2^2$			-1		

$\delta^{(2)}$	y_1	$y_1 y_2$	y_1^2	$y_1^2 y_2$	y_1^3
1			-1		
x_2	-1				1
x_2^2	-1			-1	
x_1	-1				
$x_2 + x_1 x_2^2$		-1			

Since $x_2 + x_1 x_2^2 = 0$, modulo $\langle f \rangle$, we can remove the fifth row in each Bezout matrix; the Bezout polynomials write

$\delta^{(0)}$	y_1	y_1^2	$y_1^2 y_2$	y_1^3
1				1
x_2			-1	
x_2^2				
x_1		1		

$\delta^{(1)}$	y_1	y_1^2	$y_1^2 y_2$	y_1^3
1		1	1	
x_2	1			
x_2^2				
x_1				1

$\delta^{(2)}$	y_1	y_1^2	$y_1^2 y_2$	y_1^3
1		-1		
x_2	-1			1
x_2^2	-1		-1	
x_1	-1			

Now, the first column is zero in $B^{(0)}$ but not in $B^{(2)}$. We have $x_2 + x_2^2 + x_1 = 0$, modulo $\langle f \rangle$. We use the Gauss matrix P whose fourth column is $(0, 1, 1, 1)^T$;

the Bezout polynomials write

$\delta^{(0)}$	y_1	y_1^2	$y_1^2 y_2$	y_1^3
1				1
x_2	-1	-1		
x_2^2	-1			
$x_2 + x_2^2 + x_1$	1			

$\delta^{(1)}$	y_1	y_1^2	$y_1^2 y_2$	y_1^3		$\delta^{(2)}$	y_1	y_1^2	$y_1^2 y_2$	y_1^3
1	1	1				1	-1			
x_2						x_2				1
x_2^2			-1			x_2^2			-1	
$x_2 + x_2^2 + x_1$			1			$x_2 + x_2^2 + x_1$	-1			

Since $x_2 + x_2^2 + x_1 = 0$, modulo $\langle f \rangle$, we can remove the fourth row in each Bezout matrix; the Bezout polynomials write

$\delta^{(0)}$	y_1^2	$y_1^2 y_2$	y_1^3		$\delta^{(1)}$	y_1^2	$y_1^2 y_2$	y_1^3		$\delta^{(2)}$	y_1^2	$y_1^2 y_2$	y_1^3
1			1		1	1	1			1	-1		
x_2	-1	-1			x_2	1		-1		x_2			1
x_2^2	-1				x_2^2			-1		x_2^2		-1	

Matrix $B^{(0)}$ is now invertible; the reduction process is completed. The dimension of A is 3. We observe that $\mathbf{x} = (1, x_2, x_2^2)$ and $\mathbf{y} = (y_1, y_1^2, y_1^3)$ 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, [6], [7])

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

Remark 9. Proposition 6 is guaranted when the ideal is zero-dimensional; in this case, to complete the reduction process we just have to use zero-columns of $B^{(0)}$ or, more generally, linear combinations of columns that vanish, i.e elements of the right kernel of $B^{(0)}$.

2.2.3 Bezout matrices are related to the companion matrices

As in Proposition 3, we consider the matrices X_1, X_2 defined by

$$X_1 = B^{(1)} B^{(0)-1} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix}, \quad X_2 = B^{(2)} B^{(0)-1} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix} \tag{20}$$

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

13

Proposition 7. *After the reduction process has been completed, the companion matrices X_j and the Bezout matrices are related by the* **Barnett formulas**

$$X_j = B^{(j)} B^{(0)^{-1}}, j = 1, \cdots, n \tag{21}$$

Remark 10. As in the univariate case, we have, for all $j = 1, \cdots, n$,
$B^{(j)} B^{(0)^{-1}}$ is the multiplication matrix by x_j in the basis $\mathbf{x}$
$B^{(j)^T} B^{(0)^{-T}}$ is the multiplication matrix by y_j in the basis $\mathbf{y}$
$B^{(0)^{-1}} B^{(j)}$ is the multiplication matrix by x_j in the basis $\hat{\mathbf{x}}$
$B^{(0)^{-T}} B^{(j)^T}$ 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]). The eigenvalues of matrices X_1, X_2 above 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 this example, it is easy to check that the couples (x_1, x_2) are numerical approximations of the roots of the polynomial system $f_1 = x_1^2 + x_1 x_2^2 - 1, f_2 = x_1^2 x_2 + x_1$ defined in Example 3.

2.3   Numerical experiment

We provide a reproducible computational experiment that illustrates the effectiveness of the method presented in this article. To perform the computations, follow the steps mentioned in the README.md file of the git repository [9]. In this experiment, we solve the polynomial system $f = [$
$-3x_0^2 x_1^2 x_2^2 x_3 - 2x_0^2 x_1 x_2^2 x_3 - 3x_0^2 x_1 x_2 x_3^2 + 2x_0 x_1^2 x_3 - 2x_2^2 x_3^2 + 3x_0 x_1 + 3x_1^2,$
$2x_0^2 x_1^2 x_2 x_3^2 + x_0^2 x_1 x_2^2 x_3^2 - 2x_0 x_1^2 x_2^2 x_3^2 + 2x_0^2 x_1^2 x_2 x_3 + 2x_0^2 x_1^2 x_2 + x_0^2 x_1 x_2^2 - x_0 x_1 x_2^2 x_3 - 2x_1 x_2^2 x_3^2 + x_0 x_1^2 x_2 - x_0 x_1 x_2^2 - 3x_0 x_1^2 x_3 - x_0^2 x_2 x_3 + 3x_1 x_2 x_3^2 - x_1 x_2^2 - x_1^2 x_3 - x_0 x_2 x_3 + 3x_0 x_3^2 + 3x_0^2 + 2x_0 x_2,$
$2x_0^2 x_1^2 x_2^2 x_3 + 2x_0^2 x_2^2 x_3^2 + x_0 x_1 x_2^2 x_3^2 + x_0 x_1^2 x_2 x_3 - 4x_0^2 x_1 x_3^2 + 3x_0 x_1^2 x_3^2 + x_0^2 x_2 x_3^2 - 2x_0 x_1 x_2^2 - 4x_0^2 x_1 x_3 - 2x_0^2 x_2 x_3 - 3x_0 x_1 x_2 x_3 + x_1 x_2^2 x_3 - 3x_0 x_1 x_3^2 + 2,$
$3x_0^2 x_1^2 x_2^2 x_3 + 3x_0^2 x_1^2 x_2 x_3 + 3x_0^2 x_1 x_2^2 x_3 + 3x_1^2 x_2^2 x_3 + x_0^2 x_1^2 x_2 - 3x_0 x_1 x_2^2 x_3 + x_1^2 x_2^2 x_3 - 3x_0^2 x_2^2 - 2x_0 x_2^2 x_3 - 2x_0 x_2 x_3^2 - x_0 x_1 x_2 - 2x_0 x_2^2 + 2x_2^2 x_3 - x_1 x_3 + x_3^2]$
The size of the Bezout matrix $B^{(0)}$ is 384; this is the maximum number of solutions that a system of degree $[2, 2, 2, 2]$ can have. To initiate the reduction process, the rank of $B^{(0)}$ is needed; as the matrix has integer coefficients, we use the Sage function matrix.kernel() to calculate its rank. After the reduction

process has been completed, we find that the dimension of the quotient A is 324. The Bezout matrices have rational entries as well as the companion matrices $X_j = B^{(j)}B^{(0)^{-1}}$; computing the numerical eigenvalues of the X_j's give approximations of the roots of the polynomial system f.

2.3.1   Testing the result

We test the result of the calculations by two means.

1. We can compare the dimension of the quotient A, as provided by the Groebner basis, to the dimension provided by the Bezout process, i.e the size of the reduced Bezout matrices. If, as it is the case in this experiment, the polynomial system is in complete intersection, then the two dimensions coincide. If, on the contrary, the polynomial system is not in complete intersection, then the dimension provided by the Groebner basis is infinite whereas the dimension provided by the Bezout process is always finite.

2. We can check that $f(X) = 0$, i.e $f_i(X_1, \cdots, X_n) = 0$ for every $i = 1 \cdots, n$ and where the X_j's are the companion matrices. When the polynomial system is in complete intersection, this means that every f_i belongs to the ideal $\langle f \rangle$. But, in practice, since the companion matrices have rational entries, the calculation of each $f_i(X_1, \cdots, X_n)$ may be too heavy. Rather, we choose to make the computations in a finite field F, typically $\mathbb{Z}/p\mathbb{Z}$ with, in this experiment, $p = 2003$. Moreover, to lighten even more the computation, we choose a random vector $v \in F$ and compute only the product of each $f_i(X_1, \cdots, X_n)$ by v. Thus, we only need to do matrix-vector products. If at the end of the process, and for all $i = 1 \cdots, n$, we have $f_i(X_1, \cdots, X_n)v = 0$ then we can conclude that, with a very high probability, $f(X) = 0$. In this experiment we know that the test was succesful because in the sage variable test_XX, a Python list, all the entries equal True.

3. We can check the quality of the numerical roots α by computing the errors $Df(\alpha)^{-1}f(\alpha)$. An histogram of these errors is shown in Table 1 (the errors are calculated only when the jacobian $Df(\alpha)$ is invertible).

2.3.2   Timings

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

2.3.3   Size of output

Table 3 shows the disk space used by the Bezout matrices (after the reduction process; only the non-zero entries are taken in account) as compared to the disk space used by the Groebner basis.

2.4 What happens when the polynomial system is not zero-dimensional ?

It may happen that the square polynomial system is not zero-dimensional ; the dimension of A is then infinite and this is indicated by the Groebner calculation. The Bezout process, no matter the dimension of A is finite or infinite, always produces matrices $B^{(j)}$ of finite size ; in the case when the dimension is infinite, we still call the matrices $X_j = B^{(j)}B^{(0)^{-1}}$ companion matrices. What we observe in our experiments is that, whenever the dimension of A is infinite, we still have $f(X) = 0$ (item 2 of test 2.3.1) and the eigenvalues of the companion matrices are still roots of f (item 3 of test 2.3.1).

References

[1] W. Auzinger, H. J. Stetter, An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations, Numerical mathematics, Singapore 1998, ISNM vol. 86, Birkhauser, 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, Dualité et algorithmes itératifs pour la résolution de systèmes polynomiaux, Thèse présentée devant l'université de Rennes I, 1993.

[5] J.P. Cardinal, Résolution des systèmes polynomiaux: un solveur basé sur les matrices de Bezout, arXiv:1609.09792

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

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

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

[9] `https://github.com/jpcp13/bezout`

log10 of errors	nb of roots
$[-13.9, -13.5]$	18
$[-13.5, -13.0]$	83
$[-13.0, -12.5]$	92
$[-12.5, -12.0]$	65
$[-12.0, -11.5]$	26
$[-11.5, -11.0]$	18
$[-11.0, -10.6]$	13
$[-10.6, -10.1]$	5
$[-8.1, -7.7]$	1
$[-6.7, -6.2]$	2
$[-6.2, -5.7]$	1

Table 1: histogram of errors

Method	Computation	Software	Arithmetic	Timing
	Bezout matrices	NumPy	floating point	2.7176 s
Bezout	rank of $B(1)$ via rref()	Sage	integer	0.1046 s
	reduction process	Sage	integer	16.1806 s
	eigenvalues	SciPy	floating point	2.6073 s
Groebner	Groebner basis	Sage	integer	2647.6130 s

Table 2: timings

Method	Output	Disk space usage
Bezout	reduced Bezout matrices (non-zero entries only)	1.8355 Mo
Groebner	Groebner basis	49.1156 Mo

Table 3: sizes