Solving Systems of Algebraic Equations*

Daniel LAZARD

Departement de Mathématiques, Université de Poitiers, 86022 Poitiers Cedex, France

[Now at *LIP6, Université Paris VI, 75252 Paris Cedex 05, France*]

Communicated by M. Nivat
Received October 1979
Revised March 1980

Abstract. Let $f_1, \ldots, f_k$ be k multivariate polynomials which have a finite number of common zeros in the algebraic closure of the ground field, counting the common zeros at infinity. An algorithm is given and proved which reduces the computations of these zeros to the resolution of a single univariate equation which degree is the number of common zeros. This algorithm gives the whole algebraic and geometric structure of the set of zeros (multiplicities, conjugate zeros,...). When all the polynomials have the same degree, the complexity of this algorithm is polynomial relatively to the generic number of solutions.

1. Introduction

Solving systems of algebraic equations is a crucial problem in algorithmic algebra. We will not consider methods from analysis here (Newton's method, optimization, etc.), which are fast, but give no information on the number, or even the existence, of solutions.

Algebraic methods, on the other hand, can give all desired information on the solution space (dimension, degree, multiplicity, field of definition, etc.) but are extremely slow. Thus to improve their performance, we have been led to using the entire arsenal of symbolic manipulation techniques (cf.[12]).

Previously developed techniques for solving systems of algebraic equations all lead back to eliminating one unknown after another with the help of resultants or analogous methods. This leads to manipulating polynomials of very high degree in order to compute resultants of resultants. It follows that the complexity of the problem, which is already very large, is increased by a considerable amount. We will return to this question later.

Furthermore, it is astonishing to observe that modern techniques of algebraic geometry are not used to solve systems of algebraic equations explicitly; it is especially astonishing, because algebraic geometry is generally defined as the study of solution sets of such equations.

The method of solution developed in this article is new. It draws its origins on one hand from 'classical elimination theory', and in particular from the U-resultant method [11,8], and on the other hand from work by the author on solving systems of linear equations over the polynomial ring [6]. This method is 'geometric' in the sense that it does not destroy the problem's geometric nature. In particular, invariance under linear transformations of the variables is preserved.

This method consists essentially of writing down a certain rectangular matrix of degree one polynomials and reducing it in accordance with a method akin to that of Gauss in order to obtain a square matrix whose coefficients are again of degree one. It turns out that the determinant of this matrix is a product

*Résolution des Systèmes d'Équations Algébriques. *Theoretical Computer Science* **15** (1981), 77–110. Translated by Michael Abramson.

of linear factors which are in correspondence with the solutions, the coordinates of a solution being the coefficients of the corresponding factor.

This algorithm only works when the dimension of the solution space is zero. Since such spaces are 'trivial' in algebraic geometry, the geometric techniques employed are relatively elementary, and are limited to relations between affine and projective space, the notion of multiplicity, and Hilbert's Nullstellensatz. The algebraic techniques come down mainly to properties of graded rings and modules of dimension one. It is necessary to add to this the theorem letting us restrict the degrees to be considered (Theorem 3.3), whose proof uses the homology of the Koszul complex.

The drafting of this article has posed certain problems because it was not possible to separate the description of the algorithm in an easily programmable form from its interpretation in an abstract language, which is the only way to justify the operations carried out. We have tried to resolve this difficulty by structuring the article in the following manner.

Section 2 describes how the algorithm works on a particularly simple example.

In Section 3, we pose the problem in a precise manner and transcribe it into the algebraic language that lets us justify the algorithm. In Sections 4 and 5, we describe and justify the algorithm. This description being sufficiently abstract, we give an easily programmable form of it in Appendix A.4.

In Section 6, we explain how to terminate the computations, specifically by using the algorithm from Section 5 in a new way. Section 7 is devoted to studying the multiplicity of solutions.

The object of Section 8 is the study of our algorithm's complexity and its comparison to that of the classical method. In conclusion (Section 9), we discuss the advantages of our method.

Finally, we collect the proofs of the theorems from Section 3 in Appendices A.1, A.2 and A.3. Although the first two of these theorems are 'well-known' in principle, we did not find the statements we needed in the literature. The third is implicitly proved in [6], but we feel it useful to give a direct proof of it.

2. An Example

In this section, we show how the algorithm, which is the object of this article, works on a particularly simple example. We do not give much justification and we refer often to the rest of the article. Nevertheless, we believe that this example may help in the comprehension of later sections.

Consider the system of very simple equations

$$x^2 + xy + 2x + y - 1 = 0 \tag{1}$$
$$x^2 - y^2 + 3x + 2y - 1 = 0 \tag{2}$$

which has three affine solutions, $(0,1)$, $(1,-1)$ and $(-3,1)$, and one solution at infinity in the direction $(-1,1)$ (common asymptotic direction of both surfaces).

Let f_0 and f_1 denote the first members of (1) and (2), so

$$\begin{aligned} f_0 &= -1 + 2x + y + x^2 + xy \\ f_1 &= -1 + 3x + 2y + x^2 - y^2. \end{aligned}$$

Set

$$f_2 = U + Vx + Wy$$

where U, V, and W are indeterminates.

With the integer D introduced in Theorem 3.3 being equal to 3, we construct the following table in which dots represent zeros.

$$\begin{array}{c|ccc|ccc|cccccc|} & 1 & x & y & 1 & x & y & 1 & x & y & x^2 & xy & y^2 \\ \hline 1 & -1 & . & . & -1 & . & . & U & . & . & . & . & . \\ x & 2 & -1 & . & 3 & -1 & . & V & U & . & . & . & . \\ y & 1 & . & -1 & 2 & . & -1 & W & . & U & . & . & . \\ x^2 & 1 & 2 & . & 1 & 3 & . & . & V & . & U & . & . \\ xy & 1 & 1 & 2 & . & 2 & 3 & . & W & V & . & U & . \\ y^2 & . & . & 1 & -1 & . & 2 & . & . & W & . & . & U \\ x^3 & . & 1 & . & . & 1 & . & . & . & . & V & . & . \\ x^2y & . & 1 & 1 & . & . & 1 & . & . & . & W & V & . \\ xy^2 & . & . & 1 & . & -1 & . & . & . & . & . & W & V \\ y^3 & . & . & . & . & . & -1 & . & . & . & . & . & W \\ \hline \end{array} \tag{3}$$

Here is how the table is constructed. The external column consists of the monomials of degree less than or equal to $D = 3$. The three parts of the table correspond to the polynomials f_0, f_1, and f_2. The external row of the part corresponding to f_i consists of the monomials of degree $3 - \deg(f_i)$. If m is a monomial in the external column and n is a monomial in the external row of the part corresponding to f_i, then the coefficient that lies at the intersection of the corresponding row and column is the coefficient of m in nf_i. So 2 is the coefficient of xy in xf_1.

By carrying out Gaussian elimination on the matrix that appears on the left part of the table, this becomes

$$\begin{array}{|lcccccc|}
\hline
1\,.\,.\,.\,.\,. & . & . & . & . & . & . \\
.\,1\,.\,.\,.\,. & . & . & . & . & . & . \\
.\,.\,1\,.\,.\,. & . & . & . & . & . & . \\
.\,.\,.\,1\,.\,. & . & . & . & . & . & . \\
.\,.\,.\,.\,1\,. & . & . & . & . & . & . \\
.\,.\,.\,.\,.\,1 & . & . & . & . & . & . \\
.\,.\,.\,.\,.\,. & U+V-W & U+V-W & -U-V+W & U & -U & U \\
.\,.\,.\,.\,.\,. & -V+W & -U+V & U & U-4V+W & V & 0 \\
.\,.\,.\,.\,.\,. & -2V+3W & -2U+W & 3U+V & -3V & U+W & V \\
.\,.\,.\,.\,.\,. & -V+2W & -U-V+W & 2U+V & -U & U & W \\
\hline
\end{array} \tag{4}$$

After a Gauss reduction on the coefficient matrix of U in (4), the nontrivial part of (4) becomes

$$\begin{array}{|cccccc|}
\hline
U+V+W & 0 & 0 & -V+W & V-W & -V+W \\
-V+W & -U+2V-W & V & V-W & -V+W & V-W \\
W & -2V & U-V+W & V-W & -V+W & V-W \\
0 & 0 & 0 & -V+W & V-W & -V+W \\
\hline
\end{array} \tag{5}$$

Obvious transformations on the rows and columns make additional zeros appear.

$$\begin{array}{|cccccc|}
\hline
U+V-W & 0 & 0 & 0 & 0 & 0 \\
-V+W & -U+2V-W & V & 0 & 0 & 0 \\
W & -2V & U-V+W & 0 & 0 & 0 \\
0 & 0 & 0 & -V+W & 0 & 0 \\
\hline
\end{array} \tag{6}$$

The determinant of the left part of (6) is

$$(U+V-W)(V-W)(U-3V+W)(U+W). \tag{7}$$

The coefficients of the factors of (7) give the projective coordinates of the solutions, the coordinate at infinity being the coefficients of U, the constant term of f_2. Hence, these solutions are

$$(1,1,-1) \quad (0,1,-1) \quad (1,-3,1) \quad (1,0,1),$$

the second being the solution at infinity.

Remarks. (i) Rather than viewing table (3) globally as we have just done, the algorithm we describe below proceeds in two stages: First, reducing the numeric part and computing the matrix C from the reduction, then computing the last rows (here 4) of the product of C and the right part. In order to save room in memory, there is reason to compute the matrix C by successively writing distinct columns of the left part of (3) into the same area in memory. This is what we have done in Appendix A.4.

(ii) We interpreted the interior of the table (3) as a matrix whose rows and columns are indexed by the monomials (external row and column). We recall then the general definition of matrices from [1]. It is easy indeed to enumerate the monomials with the help of a lexicographic order (we have done this implicitly to represent the matrix), but we lose information by doing it. In particular, the manner in which one column is deduced from an adjacent one becomes obscure.

(iii) This example is so simple that the computation and factorization of the determinant of (6) is immediate. We indicate below (Section 6) the way to proceed in the general case.

3. Preliminaries

A *system of algebraic equations*

$$\begin{cases} f_0(x_1,\dots,x_n) = 0 \\ \qquad\vdots \\ f_{k-1}(x_1,\dots,x_n) = 0 \end{cases} \tag{8}$$

consists of k known polynomials $f_0,\dots,f_{k-1}$ in n variables with coefficients in a field K. We say then that the system is *defined over* K.

To system (8), which is defined over K, we can associate the ring $B = K[x_1,\dots,x_n]/(f_0,\dots,f_{k-1})$, the quotient of the polynomial ring by the ideal generated by the f_i. This ring is called the *affine ring of the system.*

If L is a field extension of K, we often consider the ring $B_L = L \otimes_K B = L[x_1,\dots,x_n]/(f_0,\dots,f_{k-1})$. We call *a solution in* L *of system* (8) any element $(\xi_1,\dots,\xi_n)$ in L^n such that

$$f_i(\xi_1,\dots,\xi_n) = 0 \qquad \text{for } i = 0,\dots,k-1.$$

Theorem 3.1. *The following conditions are equivalent.*

(i) *System* (8) *has only finitely many solutions in the algebraic closure* $\overline{K}$ *of* K.

(ii) *For all extensions* L *of* K, *system* (8) *has only finitely many solutions in* L.

(iii) *The ring* B *is a finite dimensional* K*-vector space.*

If these conditions are satisfied, the number of solutions in L *is equal to at most the number of solutions in* $\overline{K}$, *which is itself equal to at most the dimension of* B *as a* K*-vector space. In other words, if* $(\xi_1, \ldots, \xi_n)$ *is a solution, the* ξ_i *are algebraic over* K. *Finally, the solutions in* $\overline{K}$ *(and hence in all extensions of* $\overline{K}$*) are in bijection with the maximal ideals of* $B_{\overline{K}}$.

The proof of this theorem is 'well-known' in principle, and like that of Theorem 3.2 below, is given in the appendix. Theorem 3.1 will not be very useful to us elsewhere, but it seems indispensable for understanding Theorem 3.2.

Let d_i denote the total degree of the polynomial f_i (for $i = 0, \ldots, k-1$). To the polynomial f_i, we can associate the homogeneous polynomial

$$F_i(X_0, \ldots, X_n) = X_0^{d_i} f_i\left(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}\right),$$

so

$$f_i(x_1, \ldots, x_n) = F_i(1, x_1, \ldots, x_n).$$

The system of equations

$$\begin{cases} F_0(X_0, \ldots, X_n) = 0 \\ \qquad\vdots \\ F_{k-1}(X_0, \ldots, X_n) = 0 \end{cases} \tag{9}$$

is the *homogeneous system* associated to system (8). If $(\zeta_0, \ldots, \zeta_n)$ is a solution of (9), the same is true of $(\lambda\zeta_0, \ldots, \lambda\zeta_n)$ for all λ. Hence, it is natural to identify two such solutions and then view the solutions of (9) in a field L as elements of projective space $\mathbb{P}_n(L)$. This is what we will do when we consider the number of solutions of system (9).

If $(\xi_1, \ldots, \xi_n)$ is a solution of (8), $(1, \xi_1, \ldots, \xi_n)$ is a solution of (9). Conversely, if $(\zeta_0, \ldots, \zeta_n)$ is a solution of (9) such that $\zeta_0 \neq 0$, then $(\zeta_1/\zeta_0, \ldots, \zeta_n/\zeta_0)$ is a solution of (8). Therefore, solutions of (9) will be called *projective solutions* of (8). Projective solutions for which $\zeta_0 = 0$ are called *solutions at infinity* of system (8), other solutions are called *affine*.

To the polynomials F_i, we can associate the graded ring $A = K[X_0, \ldots, X_n]/(F_0, \ldots, F_{k-1})$, the quotient of the polynomial ring in $n+1$ variables $X_0, \ldots, X_n$ by the ideal generated by the F_i. If L is an extension of K, we set

$$A_L = L \otimes_K A = L[X_0, \ldots, X_n]/(F_0, \ldots, F_{k-1}).$$

It is immediate that $B \cong A/(X_0 - 1)A$ and $B_L \cong A_L/(X_0 - 1)A_L$. The ring A is called the *ring of system* (9) or the *graded ring of system* (8).

For all graded rings or modules X, (for example A, A_L, or $K[X_1, \ldots, X_n]$), let X^d denote the set of all homogeneous elements of degree d and set $X^+ = \oplus_{d>0} X^d$. Finally, let $\dim_K(V)$ denote the dimension of the K-vector space V.

Theorem 3.2. *The following conditions are equivalent.*

(i) *System* (8) *has only finitely many projective solutions in the algebraic closure* $\overline{K}$ *of* K.

(ii) *For all extensions* L *of* K, *system* (8) *has only finitely many projective solutions in* L.

(iii) *There exists an integer* D *such that*

$$\dim_K(A^d) = \dim_K(A^D) \qquad \textit{for all } d \geq D.$$

(iv) *For all (resp. for one) infinite extension* L *of* K, *there exists an integer* D' *and an element* $y \in A_L^1$ *such that multiplication by* y *is a surjection from* $A_L^{D'-1}$ *onto* $A_L^{D'}$.

If these conditions are satisfied, then

(a) *Multiplication by* y *is a bijection from* A_L^d *onto* A_L^{d+1} *for all* $d \geq \max(D, D')$.

(b) *The number of projective solutions of system* (8) *is at most* $\dim_K(A^D)$.

(c) *If L' is the algebraic closure of K in L, the solutions in $\mathbb{P}_n(L)$ actually lie in $\mathbb{P}_n(L')$.*

(d) *System* (8) *satisfies the conditions of Theorem 3.1.*

(e) *System* (9) *has no non-trivial solution in $\overline{K}$ if and only if $A^D = 0$.*

It is clear that Theorem 3.2 is the projective analog of Theorem 3.1. The relevance of passing to the homogeneous system (9) may not seem obvious; in fact, this relevance stems from being able to easily calculate the integers D and D'.

Theorem 3.3. *If d_i denotes the common degree of the polynomials F_i and f_i, and if these polynomials are sorted in decreasing order of degrees ($d_0 \geq \cdots \geq d_{k-1}$), we may take $D = D' = d_0 + \ldots + d_n - n$ in the statement of Theorem 3.2. (If $k \leq n$, we set $d_i = 1$ for $i \geq k$.)*

In the statements of Theorems 3.1 and 3.2, the equivalent conditions do not actually depend on the field K; indeed, for all K-vector spaces V, $\dim_K(V) = \dim_L(L \otimes_K V)$. We have introduced a variable field L (assertion (ii)) because of the following situation. If K is the field $\mathbb{Q}$ of rational numbers, and if the conditions of Theorems 3.1 and 3.2 are satisfied, solving system (8) in $\mathbb{C}$ or in the algebraic closure of $\mathbb{Q}$ amounts to the same thing. This assertion is clearly false if the equivalent conditions are not satisfied (take a single equation $x^2 - y^3 = 0$ which admits, as it were, the solution $x = \pi^3$, $y = \pi^2$).

4. Reduction - First Part

Suppose system (8) satisfies the equivalent conditions of Theorem 3.2, and consider an arbitrary element
$$y = u_0X_0 + \ldots + u_nX_n,$$
of degree 1 in $K[X_0, \ldots, X_n]$. If we add the equation
$$y = 0$$
to system (9), the ring A is replaced by $A' = A/yA$. Two cases can occur. If there exists a solution to system (9) that cancels y as well, then ${A'}^D \neq 0$ (Theorem 3.2(e)), which means that multiplication by y from A^{D-1} into A^D is not surjective. On the other hand, if no solution of (9) cancels y, then ${A'}^D = 0$ and multiplication by y is surjective. Thus studying this multiplication by y may enable us compute the solutions of systems (8) and (9).

Rather than making the u_i vary in K or $\overline{K}$, we proceed by introducing indeterminates $U_0, \ldots, U_n$ and setting
$$L = U_0X_0 + \ldots + U_nX_n.$$
This is an element of $R[U_0, \ldots, U_n]$ with $R = K[X_0, \ldots, X_n]$.

To simplify subsequent notation, we set once and for all
$$R = K[X_0, \ldots, X_n]$$
and
$$V_U = K[U_0, \ldots, U_n] \otimes_K V$$
for all K-vector spaces V. We identify elements v in V with their images $1 \otimes v$ in V_U. Thus
$$L \in (R^1)_U \subset R_U$$
for example, since R is graded (total degree in X). We always view the U_i as having degree zero, that is to say, notations such as R^d_U, A^d_U will mean $(R^d)_U$, $(A^d)_U$.

Having done this, consider the commutative diagram
$$\begin{tikzcd} R_U^{D-1} \arrow[r, "L"] \arrow[d, "p_U^{D-1}"'] & R_U^D \arrow[d, "p_U^D"] \\ A_U^{D-1} \arrow[r, "L"] & A_U^D \end{tikzcd}$$
where L denotes multiplication by L and vertical arrows are canonical projections. Thus p^D_U is the extension of the canonical projection $p^D : R^D \to A^D$ to R^D_U, and the kernel of p^D is the set of $G_0F_0 + \ldots + G_{k-1}F_{k-1}$ where $G_i \in R^{D-d_i}$ for all i. In other words, (p^D, A^D) is the cokernel of the linear function
$$\phi : R^{D-d_0} \times \cdots \times R^{D-d_{k-1}} \longrightarrow R^D$$
defined by $\phi(G_0, \ldots, G_{k-1}) = \sum G_iF_i$.

It is easy to write the matrix Φ of ϕ when we take the basis formed by the monomials as the basis of the K-vector spaces that appear.

Example. In the example of Section 2, the matrix Φ is the left part of table (3).

We will call the following recursive operation *Gauss reduction* of a matrix. If the matrix is zero or empty, do nothing. Otherwise, choose a non-zero element of this matrix (the pivot), add a multiple of the pivot row to all rows except that of the pivot so as to cancel the element corresponding to the column of the pivot, and do a Gauss reduction of the submatrix obtained by omitting the row and column of the pivot.

Thus a Gauss reduction on the matrix Φ would let us compute a new basis $e_1, \dots, e_s$ of R^D such that $e_1, \dots, e_r$ would be a basis of the image of ϕ and the matrix of ϕ in this new basis would have the form

$$\begin{array}{c} r\left\{\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right. \\ s-r \end{array} \left(\begin{array}{cccc|c} 1 & & & * & \\ & 1 & & & \\ & & 1 & & * \\ 0 & & & 1 & \\ \hline & & 0 & & 0 \end{array}\right) = \Phi'$$
$$\underbrace{\qquad\qquad}_{r}$$

(up to permutation of the columns and rows and multiplying the rows by the inverses of the pivots) where $*$ denotes any coefficients whatsoever. It is clear that the restriction of p^D to the subspace generated by $e_{r+1}, \dots, e_s$ is an isomorphism. In particular, $\dim_K(A_D) = s - r$. It follows that if $r = s$, the system of equations being considered has no solution.

Let C be the change of basis matrix that the Gauss reduction lets us obtain. Then $\Phi' = C\Phi$.

Remark. To program the Gauss reduction and compute the matrix C, which is itself useful in what follows, it is not necessary to write the matrix Φ all at once. It suffices to write one column after another in the same area in memory. Initializing C to an identity matrix, we proceed as follows for each column of Φ: write the column, multiply it by the old value of C, find a pivot in the resulting column and in the rows where we have not yet found a pivot, if there is a pivot the operations on the rows that cancel the rest of the column being considered correspond to left multiplication by a matrix C_1, replace C by C_1C and go to the next column. This remark is used in the explicit algorithm of [7].

Returning to $K[U_0, \dots, U_n]$-modules, it is easy to write the matrix of $L : R_U^{D-1} \to R_U^D$ in the basis of the monomials, a matrix also denoted by L. It is immediate that CL is the matrix of L in the new basis of R_U^D calculated above (every basis of R^D is also a basis of R_U^D). The matrix of $p_U^D L = L p_U^{D-1}$ in the basis of A_U^D consisting of the images of $e_{r+1}, \dots, e_s$ consists of the last $s - r$ rows of CL, and hence is easy to compute.

Example. In the example of Section 2, the matrix L is the right part of table (3) and the last $s-r$ rows of CL form the non-trivial part of the matrix (4).

At this step of the reduction we need to restate the problem in order to give it a more general form which will allow us to work recursively.

Thus we must deal with a graded A-module M generated by its elements of degree $< D$ (here $M = A$). The component M^{D-1} of M is expressed as a quotient of a K-vector space V (here $V = R^{D-1}$), and we know the matrix Λ of the composition of multiplication by L and the projection $p : V \to M^{D-1}$:

$$\begin{tikzcd} V_U \arrow[d, "p_U"'] \arrow[dr, "\Lambda"] & \\ M_U^{D-1} \arrow[r, "L"] & M_U^D \end{tikzcd}$$

Since we assume the equivalent conditions of Theorem 3.2 are satisfied for all infinite extensions K' of K, there exists y in A_K^1 such that multiplication by y is a surjection from $M_{K'}^{D-1}$ onto $M_{K'}^D$ and a bijection from $M_{K'}^d$ onto $M_{K'}^{d+1}$ for all $d \geq D$. This condition will be called *condition* (Y).

Lemma 4.1. *If $x \in M_{K'}^{D-1}$, $z \in A_U^1$, and condition* (Y) *is satisfied, then*

$$yx = 0 \quad \Longrightarrow \quad zx = 0.$$

Proof. By the commutativity of A, $yzx = zyx = 0$. Since $zx \in A_U \otimes M^D_{K'}$, multiplication by y is injective, which implies $zx = 0$. $\square$

Proposition 4.1. *If condition* (Y) *is satisfied, there exists a basis of V defined by a matrix Γ with coefficients in K, for which the matrix of $L\,p_U$ has the form*

$$\Lambda\Gamma = (\Lambda' \quad 0)$$

where Λ' is a square matrix.

Example. In Section 2, $\Lambda\Gamma$ is the matrix (6).

Proof. If K is infinite, we may take $K' = K$ and choose the basis of V in such a way that the kernel of $y\,p$ (which is a surjection) is generated by the last elements of this basis. Then Lemma 4.1 implies that the kernel of $L\,p_U$ contains the kernel of $y\,p$, which proves the proposition, bearing in mind the dimensions of V,M^D, and $\ker(y\,p)$.

If K is finite, it is necessary to use Lemma 4.1 in another way. If $x \in M^{D-1}_{K'}$, then $yx = 0$ implies $X_i x = 0$ for all i. In other words, $\ker(y\,p) \subset \bigcap \ker(X_i p)$. Equality actually holds because y is a linear combination (with coefficients in K') of the X_i. Then since $y\,p$ is surjective,

$$\dim_{K'}(V_{K'}) = \dim_{K'}\left(\bigcap \ker(X_i p)\right) + \dim_{K'}(M^D_{K'}). \quad (10)$$

Since all these vector spaces are actually defined over K, this equality is likewise true over K, and it suffices to take a basis of V whose last elements consist of a basis of $\bigcap \ker(X_i p)$. Since $L\,p_U$ is annihilated over $\bigcap \ker(X_i p)$, equation (10) implies the proposition. $\square$

Theorem 4.1. *Let r be the number of rows of Λ. Suppose condition* (Y) *is satisfied.*

(a) *The rank of Λ is equal to r.*

(b) *The ideal generated by the $r \times r$ determinants extracted from Λ is principal and is generated by a homogeneous polynomial $G(U_0, \dots, U_n)$ of degree r in $U_0, \dots, U_n$.*

(c) *If $M = A$, the polynomial G factors as a product of degree one factors over an algebraic closure $\overline{K}$ of K. The polynomial $\alpha_0 U_0 + \ldots + \alpha_n U_n$ is such a factor if and only if $(\alpha_0, \dots, \alpha_n)$ is a common zero of the F_i.*

Proof. (a) follows from surjectivity of multiplication by y: if $y = \sum y_i X_i$, the matrix of $y\,p$ is obtained by substituting y_i for U_i in the matrix of Λ, and this substitution can only reduce the rank.

(b) Multiplication by Γ (cf. Proposition 4.1) can be accomplished by a succession of operations consisting of either multiplying a column by an element of K or adding a multiple of one column to another column. These operations do not change the ideal generated by the determinants, which shows that this ideal is equal to the ideal generated by the determinant G of Λ'.

(c) To show the third assertion, we adjoin the equation $u_0 X_0 + \ldots + u_n X_n = 0$ to the equations $F_i = 0$. This determines a new ring $A' = A/(u_0 X_0 + \ldots + u_n X_n)$, and ${A'}^D = 0$ if and only if the new system has no common zero. But then ${A'}^D = A^D/(u_0 X_0 + \ldots + u_n X_n)A^{D-1}$ and ${A'}^D = 0$ implies that multiplication by $u_0 X_0 + \ldots + u_n X_n$ is surjective, that is to say $G(u_0, \dots, u_n) \neq 0$. Furthermore, if $(\alpha_{0,j}, \dots, \alpha_{n,j})$ $(j = 1, \dots, h)$ are the common zeros of the old system, the new system does not have a common zero if and only if

$$\prod_j (\alpha_{0,j} u_0 + \ldots + \alpha_{n,j} u_n) \neq 0.$$

In other words, $\prod_j (\alpha_{0,j} U_0 + \ldots + \alpha_{n,j} U_n)$ and $G(U_0, \dots, U_n)$ have the same zeros, which gives the result with the help of the Nullstellensatz ([5, Theorem 33], for example). $\square$

Remark. We will see further (Section 7) that the multiplicity of a factor of G is equal to the multiplicity of the corresponding zero. In other words, this theorem enables us not only to compute the zeros, but also to determine their multiplicity. In particular, we can distinguish multiple (singular) zeros from the others.

5. Reduction - Second Part

In the algorithm's present stage, we know neither the element y nor the matrix Γ a fortiori, and it is very onerous to compute G directly as a PGCD of a family of determinants. So we now proceed with a sequence of recursive reductions which will result in obtaining the matrix Λ' in block triangular form. The reason for this reduction, other than avoiding the PGCD computation, is to obtain G in an already partially factored form. This reduction method will lead to working over A-modules M obtained as quotients of submodules of A.

Let us return now to the diagram

$$\begin{tikzcd} V_U \arrow[d, "p_U"'] \arrow[dr, "\Lambda"] & \\ M_U^{D-1} \arrow[r, "L"] & M_U^D \end{tikzcd}$$

The coefficients of the matrix Λ are polynomials, homogeneous of degree one in U. Choose an index i such that U_i appears in Λ. For example, suppose that U_0 appears in Λ and take $U_i = U_0$.

By doing a Gauss reduction on the coefficient matrix of U_0, and by carrying out the same operations on the rows of Λ, it is easy to compute a square matrix Γ_1 with coefficients in K such that

$$\Gamma_1\Lambda = \begin{pmatrix} U_0\Lambda_0 + \Lambda_1 & \Lambda_2 \\ \Lambda_3 & \Lambda_4 \end{pmatrix}$$

where Λ_0 is a triangular invertible matrix, $\Lambda_1, \Lambda_3, \Lambda_4$ are matrices independent of U_0, and Λ_2 may depend on U_0.

Example. In Section 2, the matrix $\Gamma_1\Lambda$ is the matrix (5). However, to get simpler matrices in Section 2, we have carried out Gauss reductions on the rows as well as columns ('diagonalization'). This is why the matrix Λ_0 is diagonal and Λ_2 does not depend on U $(= U_0)$.

Two cases will be considered.

First case. The number of rows of Λ_3 and Λ_4 is zero, i.e.

$$\Gamma_1\Lambda = \begin{pmatrix} U_0\Lambda_0 + \Lambda_1 & \Lambda_2 \end{pmatrix}.$$

Let us decompose Λ_2 into $\Lambda_2 = U_0\Lambda_2' + \Lambda_2''$ where Λ_2' has coefficients in K and Λ_2'' has coefficients in $K[U_1, \dots, U_n]$. Set

$$\Gamma_2 = \begin{pmatrix} I_1 & -\Lambda_0^{-1}\Lambda_2' \\ 0 & I_2 \end{pmatrix}$$

where I_1 and I_2 are identity matrices of suitable dimensions. Then

$$\Gamma_1\Lambda\Gamma_2 = \begin{pmatrix} U_0\Lambda_0 + \Lambda_1 & \Lambda_2'' - \Lambda_1\Lambda_0^{-1}\Lambda_2' \end{pmatrix}$$

Lemma 5.1. $\Lambda_2'' - \Lambda_1\Lambda_0^{-1}\Lambda_2' = 0$.

Proof. Using the notation of Proposition 4.1,

$$\begin{aligned} \begin{pmatrix} U_0\Lambda_0 + \Lambda_1 & \Lambda_2'' - \Lambda_1\Lambda_0^{-1}\Lambda_2' \end{pmatrix} &= \Gamma_1 \begin{pmatrix} \Lambda' & 0 \end{pmatrix} \Gamma^{-1}\Gamma_2 \\ &= \begin{pmatrix} \Gamma_1\Lambda' & 0 \end{pmatrix} \Gamma^{-1}\Gamma_2 \end{aligned}$$

and there exists then two matrices Δ_1 and Δ_2 with coefficients in K such that

$$\begin{aligned} U_0\Lambda_0 + \Lambda_1 &= \Gamma_1\Lambda'\Delta_1 \quad \text{and} \\ \Lambda_2'' - \Lambda_1\Lambda_0^{-1}\Lambda_2' &= \Gamma_1\Lambda'\Delta_2. \end{aligned}$$

Substituting 1 for U_0 and 0 for the other U_i, it becomes

$$\begin{aligned} \Lambda_0 &= \Gamma_1\Lambda'(1, 0, \dots, 0)\Delta_1 \quad \text{and} \\ 0 &= \Gamma_1\Lambda'(1, 0, \dots, 0)\Delta_2. \end{aligned}$$

Thus $\Gamma_1\Lambda'(1, 0, \dots, 0)$ is invertible and $\Delta_2 = 0$, which proves the lemma. □

Corollary. *G is the determinant of the nonzero columns of $\Gamma_1\Lambda\Gamma_2$.*

Second case. The number of rows of Λ_3 and Λ_4 is $r \neq 0$.

There are two ways to treat this situation. We can show that the matrix $(\Lambda_3 \;\; \Lambda_4)$ is the matrix of a function $\Lambda : V_U \to M_U^D$ corresponding to a module M satisfying condition (Y), then compute the matrix Γ such that $(\Lambda_3 \;\; \Lambda_4)\Gamma = (\Lambda' \;\; 0)$ for some square matrix Λ'. This is possible, recursively, because the number of rows of $(\Lambda_3 \;\; \Lambda_4)$ is strictly less than that of Λ. Then

$$\Gamma_1\Lambda\Gamma = \begin{pmatrix} * & \Lambda'' \\ \Lambda' & 0 \end{pmatrix}.$$

We can show that Λ'' is the matrix of another function $\Lambda : V_U \to M_U^D$ corresponding to another module M satisfying condition (Y). Another recursion lets us then compute a matrix Γ' such that $\Lambda''\Gamma' = (\Lambda''' \ \ 0)$ where Λ''' is also a square matrix. It is clear that the desired polynomial G is the product of the determinants of Λ' and Λ'''.

The trouble with this method is that the recursion follows a binary tree and its programming is therefore more complicated than the method we now describe. It consists similarly of interpreting $(\Lambda_3 \ \ \Lambda_4)$ as the matrix of a function $\Lambda : V_U \to M_U^D$, but instead of putting this matrix in the form $(\Lambda' \ \ 0)$, we simply apply the discussion from the first case. We choose an indeterminate U_i, for example U_1, and matrices $\Gamma_1^1, \Lambda_0^1, \ldots, \Lambda_4^1$ such that

$$\Gamma_1^1(\Lambda_3 \quad \Lambda_4) = \begin{pmatrix} U_1\Lambda_0^1 + \Lambda_1^1 & \Lambda_2^1 \\ \Lambda_3^1 & \Lambda_4^1 \end{pmatrix},$$

Λ_0^1 is triangular invertible and Λ_1^1, Λ_3^1, Λ_4^1 are independent of U_1 (and also U_0 of course). If the number of rows of $(\Lambda_3^1 \ \ \Lambda_4^1)$ is not zero, we can begin again, writing

$$\Gamma_1^2(\Lambda_3^1 \quad \Lambda_4^1) = \begin{pmatrix} U_2\Lambda_0^2 + \Lambda_1^2 & \Lambda_2^2 \\ \Lambda_3^2 & \Lambda_4^2 \end{pmatrix}$$

and iterating until we obtain an integer k such that

$$\Gamma_1^k(\Lambda_3^{k-1} \quad \Lambda_4^{k-1}) = \begin{pmatrix} U_k\Lambda_0^k + \Lambda_1^k & \Lambda_2^k \end{pmatrix}.$$

If we set $\Lambda_2^k = U_k\Lambda_2' + \Lambda_2''$ with Λ_2'' independent of U_k and

$$\Gamma_2 = \begin{pmatrix} I_1 & -(\Lambda_0^k)^{-1}\Lambda_2' \\ 0 & I_2 \end{pmatrix},$$

Lemma 5.1 shows that

$$\Gamma_1^k(\Lambda_3^{k-1} \quad \Lambda_4^{k-1})\Gamma_2 = (U_k\Lambda_0^k + \Lambda_1^k \quad 0).$$

For all i, set

$$\Gamma^i = \begin{pmatrix} I & 0 \\ 0 & \Gamma_1^i \end{pmatrix}$$

where I is an identity matrix of suitable dimension. Then

$$\Gamma^k\Gamma^{k-1}\cdots\Gamma^1\Gamma_1\Lambda\Gamma_2 = \begin{pmatrix} \Lambda^2 & \Lambda^1 \\ U_k\Lambda_0^k + \Lambda_1^k & 0 \end{pmatrix}.$$

Thus the ideal generated by the determinants of Λ is the product of the determinant of $U_k\Lambda_0^k + \Lambda_1^k$ and the ideal generated by the determinants of Λ^1. By applying the above discussion to Λ^1 and iterating, we obtain the following theorem, with the stipulation that we can apply Lemma 5.1 whenever necessary.

Example. In Section 2, the matrices Λ_3^1 and Λ_4^1 have no rows (i.e. $k = 1$), the matrix $U_k\Lambda_0^k + \Lambda_1^k$ is simply $-V + W$, and the matrix Λ^1 is simply the matrix $I_0\Lambda_0 + \Lambda_1$ from the previous stage bordered by two zero columns. Its reduction is already done.

Theorem 5.1. *The algorithm described above lets us compute the polynomial G as the product of determinants of some number of square matrices of the form $U_k\Lambda_0^k + \Lambda_1^k$ where Λ_0^k is a triangular invertible matrix with coefficients in K and Λ_1^k is a square matrix with coefficients that are linear and homogeneous in $U_{k+1}, \ldots, U_n$.*

Remark. The procedure REDUC in Appendix A.4 is precisely a translation of the above algorithm, except for one detail: in REDUC, we have systematically right multiplied the intermediary matrices by matrices

$$\begin{pmatrix} (\Lambda_0^i)^{-1} & -(\Lambda_0^i)^{-1}\Lambda_2' \\ 0 & I \end{pmatrix},$$

which lengthens the computation time, but simplifies the programming slightly.

Proof. To complete the proof of Theorem 5.1, it is necessary to explain the meaning of the matrices $(\Lambda_3 \quad \Lambda_4)$ and Λ^1. To do this, let $\varepsilon_1, \ldots, \varepsilon_s$ and $e_1, \ldots, e_t$ denote the bases of V and M^D for which

$$\Gamma_1\Lambda = \begin{pmatrix} U_0\Lambda_0 + \Lambda_1 & \Lambda_2 \\ \Lambda_3 & \Lambda_4 \end{pmatrix}$$

is the matrix of $L\,p_U$.

By substituting 1 for U_0 and 0 for the other U_i, we see that the image of multiplication by X_0 is generated by $e_1, \ldots, e_r$ where r is the dimension of Λ_0. This shows that $(\Lambda_3 \ \ \Lambda_4)$ is the matrix of the composition

$$p_1 : V_U \xrightarrow{L\,p_U} M_U^D \longrightarrow (M/X_0M)_U^D.$$

Set $M' = M/X_0M$. It is easy to verify that M' satisfies condition (Y) and that p_1 is the composition $L'\,p'_U$ where L' is multiplication by L in M' and p' is the composition of p and the canonical projection $M^{D-1} \to M^{D-1}/X_0M^{D-2}$. This shows that Lemma 5.1 does indeed apply to the matrix $(\Lambda_3 \;\; \Lambda_4)$.

By iterating the above, we see that $(\Lambda_3^{k-1} \;\; \Lambda_4^{k-1})$ corresponds to the function

$$V_U \longrightarrow (M/X_0M + \ldots + X_{k-1}M)_U^D$$

and that $M^D = X_0M^{D-1} + \ldots + X_kM^{D-1}$. Let $e'_1, \ldots, e'_s$ and $\varepsilon'_1, \ldots, \varepsilon'_t$ denote the bases of V and M^D for which

$$\begin{pmatrix} \Lambda^2 & \Lambda^1 \\ U_k\Lambda_0^k + \Lambda_1^k & 0 \end{pmatrix}$$

is the matrix of $L\,p_U$, and let r denote the dimension of $U_k\Lambda_0^k + \Lambda_1^k$. The images $\varepsilon'_{t-r+1}, \ldots, \varepsilon'_t$ form a basis of $(M/X_0M + \ldots + X_{k-1}M)^D$ and the images of $e'_{s-r+1}, \ldots, e'_s$ are zero in this module.

It follows that $X_0M^{D-1} + \ldots + X_{k-1}M^{D-1}$ is of dimension $t - r$ and has $\varepsilon'_1, \ldots, \varepsilon'_{t-r}$ for a basis. Let V' denote the subspace of V generated by $e'_{r+1}, \ldots, e'_s$. The image of V'_U under $L\,p_U$ is contained in $X_0M_U^{D-1} + \ldots + X_{k-1}M_U^{D-1}$. Finally, set $M'' = A\,p(V') + X_0M + \ldots + X_{k-1}M$. Since multiplication by X_i and y commute, we see easily that M'' satisfies condition (Y), the image of the restriction p' of p to V' lies in M''^{D-1}, and the matrix of $L\,p'$ is Λ^1. The function $p' : V' \to M''^{D-1}$ is surjective: First argue in an infinite extension K' of K. Since $L\,p_U(V'_U) \subset M''^D_U$, we obtain the inclusion $y\,p_{K'}(V'_{K'}) \subset M''^D_{K'}$ by substituting the coefficients of y for U_i. By passing to the quotient, we get a surjection from $V_{K'}/V'_{K'}$ onto $M^D_{K'}/M''^D_{K'}$, which is bijective because dimensions are equal. We conclude that the function $p_{K'}$ induces an injection from $V_{K'}/V'_{K'}$ into $M^{D-1}_{K'}/M''^{D-1}_{K'}$, which implies that $p'_{K'} : V'_{K'} \to M''^{D-1}_{K'}$ is a surjection, and that the same is true for $p' : V' \to M''^{D-1}$.

The preceding shows that the method described for reducing the matrix Λ is applicable to the matrix Λ^1 and concludes the proof of Theorem 5.1. □

Remark. The validity of the other reduction method broadly outlined is proved similarly, the matrices Λ' and Λ'' corresponding then to the modules M/X_0M and $X_0M + p'(V')$. In fact, this other method turns out to be better [13].

6. Final Calculations

6.1 The Case Where the Conditions of Theorem 3.2 are not Satisfied

If the system has infinitely many projective solutions, the rank of the matrix Λ reduced in the preceding section is less than its number of rows, and the modules that appear do not all satisfy condition (Y). So during the algorithm's execution, this situation is inevitable revealed by at least one of the following two events:

(a) Appearance of a matrix $(\Lambda_3^i \;\; \Lambda_4^i)$ that is zero.

(b) Invalidity of Lemma 5.1 at the time it is used.

These two events are tested by comparing to zero certain matrices that appear during the algorithm's execution. It follows that the algorithm lets us verify if the equivalent conditions of Theorem 3.2 are satisfied.

Remark. Event (a) is automatically tested during execution. It is not the same for event (b). If we are sure of the validity of the equivalent conditions of Theorem 3.2, then we save execution time by not testing for (b).

It is possible that even if the system has infinitely many solutions, event (a) already occurred. If that were the case, the above time savings would always be possible.

6.2 Solutions Having a Zero Coordinate

The algorithm described in the preceding section gives solutions $(a_0, \ldots, a_n)$ as linear factors $a_0U_0 + \ldots + a_nU_n$ of determinants of matrices of the form $U_k\Lambda_0^k + \Lambda_1^k$ where Λ_0^k is a triangular invertible matrix with coefficients in K and Λ_1^k is a square matrix with coefficients that are linear and homogeneous relative to $U_{k+1}, \ldots, U_n$. Every linear factor of such a determinant has the form

$$a_k U_k + \ldots + a_n U_n$$

with $a_k \neq 0$. Thus the corresponding solutions all satisfy $a_0 = a_1 = \cdots = a_{k-1} = 0$ and $a_k \neq 0$. Among these solutions, some may satisfy $a_i = 0$ for some $i > k$. It is easy to separate them from the others: it suffices to change the order

$$U_k, \ldots, U_n$$

in which we consider the U_i to

$$U_i, U_k, U_{k+1}, \ldots, U_{i-1}, U_{i+1}, \ldots, U_n,$$

then to apply the algorithm described in Section 5 to the matrix $U_k \Lambda_0^k + \Lambda_1^k$.

This separation of solutions helps in reducing the size of the determinants being computed, and hence can greatly simplify the end of solving.

In particular, the algorithm separates the solutions at infinity ($a_0 = 0$) from the others. This is particularly attractive because homogenation often introduces numerous superfluous solutions at infinity.

6.3 Solutions on a Particular Hyperplane

It can happen that, for geometric reasons, it is probable that some solutions lie on a particular hyperplane of equation

$$b_0 X_0 + \ldots + b_n X_n = 0.$$

This means that $a_0 b_0 + \ldots + a_n b_n = 0$.

For example, suppose b_0 is not zero and let us make a linear change of variables

$$U_i' = U_i + \frac{b_i}{b_0} U_0 \quad \text{for } i > 0, \qquad U_0' = U_0.$$

Then the polynomial $a_0 U_0 + \ldots + a_n U_n$ becomes

$$\frac{1}{b_0} \sum_{i=0}^{n} a_i b_i U_0' + a_1 U_1' + \ldots + a_n U_n'.$$

This indicates a way to separate solutions satisfying $\sum a_i b_i = 0$ from the others: Make the above change of variables in the matrix

$$U_k \Lambda_0^k + \Lambda_1^k$$

and apply the algorithm of Section 5 to the resulting matrix. This leads to matrices of the form $U_i' \Lambda_0^i + \Lambda_1^i$. Those of these matrices for which $i \neq 0$ correspond to solutions such that $\sum a_i b_i = 0$.

6.4 Computing the Solutions

There is one case where the computation of solutions corresponding to

$$U_k \Lambda_0^k + \Lambda_1^k$$

is immediate. It is when the dimension of this matrix is 1. Then its unique coefficient

$$a_k U_n + \ldots + a_n U_n$$

is equal to its determinant and thus gives the solution

$$(a_k, \ldots, a_n)$$

without computation.

In the general case, the method consisting of computing the determinant of $H = U_k \Lambda_0^k + \Lambda_1^k$ and factoring the resulting polynomial is costly. It is faster to proceed as follows. We may assume $\Lambda_1^k \neq 0$. If not, the desired determinant is simply a power of U_k. Suppose then, for example, that U_{k+1} appears in Λ_1^k, and substitute 1 for U_{k+1} and 0 for $U_{k+2}, \ldots, U_n$. We obtain a matrix

$$\overline{H} = U_k \Lambda_0^k + \overline{\Lambda}_1^k$$

where Λ_0^k and Λ_1^k are scalar matrices. If $G(U_k, \ldots, U_n)$ is the determinant of H and $\overline{G}(U_k) = G(U_k, 1, 0, \ldots, 0)$ is that of $\overline{H}$, it is immediate that for all factors

$$a_k U_k + \ldots + a_n U_n$$

of G, $a_k \neq 0$ and $-a_{k+1}/a_k$ is a root of $\overline{G}$. Since the solutions and factors of G are defined up to multiplication by an element of K, we may assume $a_k = 1$. It follows that the coefficients a_{k+1} of the solutions are the roots of the equation $\overline{G}(-U_k) = 0$.

Suppose for a moment that this equation is solved, and let a_{k+1} be one of its roots. The corresponding solutions lie on the hyperplane of equation

$$a_{k+1} X_k - X_{k+1}.$$

With the help of a new application of the algorithm of Section 5, the argument described in Section 6.3 above lets us separate the solutions $(1, a_{k+1}, \dots)$ from the others. Furthermore, if there is only one solution of the form $(1, a_{k+1}, \dots)$, which is the most frequent case, its coefficients are obtained as rational functions of a_{k+1}. If there are several solutions of the form $(1, a_{k+1}, \dots)$, the algorithm of Section 5 may lead to matrices $U_k \Lambda_0^k + \Lambda_1^k$ of dimension greater than 1, but their dimension is less than that of the matrix that is left. By successively making $U_{k+2}, \dots, U_n$ play the role played by U_{k+1} in the preceding argument, we finish by computing all the coefficients of the solutions $(1, a_{k+1}, \dots)$.

It remains to see how to compute the roots of $\overline{G}(-U_k)$. One method consists of computing the polynomial, by interpolation for example, and calculating its roots by any of the classical methods. A second method consists of noticing that the roots of $\overline{G}(-U_k)$ are precisely the eigenvalues of $\overline{\Lambda}_1^k (\Lambda_0^k)^{-1}$ (recall that Λ_0^k is triangular invertible). Thus methods for computing eigenvalues let us solve the equation $\overline{G}(-U_k) = 0$ directly without computing the polynomial $\overline{G}$.

If we are interested in the algebraic structure of the solution set, there is still reason to compute $\overline{G}$ and factor it into irreducible and unitary factors. If $f(U_k)$ is such a factor, we can work in the field $K(a_{k+1}) = K[X]/(f(-X))$. Thus the pursuit of the solution produces the solutions in terms of a_k, most often rationally. Furthermore, conjugate solutions are grouped together.

7. Multiplicity

It is frequently the case that some solutions of the system being considered are multiple solutions. For example, the intersection of a circle and a tangent is a double point. These multiple solutions often have a geometric origin different from simple solutions and ought to be viewed as irregular. It is therefore useful to recognize them and eventually to compute the multiplicity. The algorithm we have described does this without extra computation, as asserted by the following theorem whose proof is the object of this section.

Theorem 7.1. *In the statement of Theorem 4.1, the multiplicity of a factor $a_0U_0 + \ldots + a_nU_n$ of G is equal to the multiplicity of the corresponding solution $(a_0, \dots, a_n)$.*

Proof. For what is stated to make sense, we need a precise definition of the multiplicity of a solution $(a_0, \dots, a_n)$. By definition, this multiplicity will be the multiplicity in the ring A of the prime ideal $\mathfrak{p}$ generated by the $a_iX_j - a_jX_i$ (see Proposition A.3 and [4, p. 51]), that is to say, the length of the Artinian ring $A_\mathfrak{p}$.

Let us make a linear and homogeneous change of variables in such away that the solution $(a_0, \dots, a_n)$ becomes $(0, \dots, 0, 1)$ relative to the new variables: it suffices to take $X_i'' = X_i - (a_i/a_n)X_n$ for $i < n$, $X_n' = X_n$, $U_n' = a_0U_0 + \ldots + a_nU_n$ and $U_i' = U_i$ for $i \neq n$ (we assume $a_n \neq 0$, which we can always do after a possible permutation of variables).

When we apply the algorithm of Section 5, the polynomial G corresponding to a module M factors as a product of a polynomial G' corresponding to a module $M' = M/(X_0'M + \ldots + X_{k-1}'M)$ and a polynomial G'' corresponding to a module $M'' = X_0'M + \ldots + X_{k-1}'M + p(V')$. Let M''' denote the module $X_0'M + \ldots + X_{k-1}'M$. From the exact sequence

$$0 \to M''' \to M \to M' \to 0,$$

we deduce the relation

$$\mathrm{mult}(\mathfrak{p}, M) = \mathrm{mult}(\mathfrak{p}, M''') + \mathrm{mult}(\mathfrak{p}, M')$$

between the multiplicities ($\mathrm{mult}(\mathfrak{p}, M)$ is the length of $M_\mathfrak{p}$). But since $\mathfrak{p}$ is generated by $X_0', \dots, X_{n-1}'$, the element X_n' is invertible in $A_\mathfrak{p}$. It follows that if $z \in p(V')$, the image $z/1$ of z in $M_\mathfrak{p}''$ is equal to $X_n'z/X_n'$, which shows that $M_\mathfrak{p}'' = M_\mathfrak{p}'''$ (because $L\,p(V') \subset M_U''')$ and that M'' and M''' have the same multiplicity.

It follows then from above that it is sufficient to prove the theorem when G is the determinant of a matrix of the form $U_k\Lambda_0^k + \Lambda_1^k$ corresponding to some module M, and the solution is $(0, \dots, 0, 1)$.

If $k \neq n$, G has no factor equal to X_n. Furthermore, X_k is nilpotent and X_n invertible in the

Artinian ring $A_{\mathfrak{p}}$. Since multiplication by X_k from M^{d-1} into M^d is surjective for $d \geq D$,

$$X_n^{D+l} z = X_k^l t$$

for all z in M and all l. If l is sufficiently large, $X_k^l = 0$ in $A_{\mathfrak{p}}$, which implies $M_{\mathfrak{p}} = 0$, that is to say $\operatorname{mult}(\mathfrak{p}, M) = 0$.

If $k = n$, then $\Lambda_1^n = 0$ and $G = \lambda U_n^r$ with $\lambda \in K$ and $r = \operatorname{rank}(\Lambda_0^n)$. On the other hand, there exists a series $0 = M_0 \subsetneqq M_1 \subsetneqq \cdots \subsetneqq M_l = M$ of graded submodules of M such that $M_{i+1}/M_i \cong (A/\mathfrak{p}_i)(l_i)$ for all i, where $\mathfrak{p}_i$ is a graded prime ideal of A and $(A/\mathfrak{p}_i)(l_i)$ is defined by $(A/\mathfrak{p})(l_i) = (A/\mathfrak{p})^{d+l_i}$ [4, Section I.7.4]. Since $\dim_K(M^d)$ is independent of d for d sufficiently large, the same is true for $\dim_K(M_{i+1}^d/M_i^d)$. If $\mathfrak{p}_i$ is the ideal $\mathfrak{m}$ generated by all the X_i, then $A/\mathfrak{p}_i \cong K$ and this dimension is zero. If $\mathfrak{p}_i$ is a prime ideal distinct from $\mathfrak{m}$ and $\mathfrak{p}$, there exists $z \in A_1$ such that $z \in \mathfrak{p}$ and $z \notin \mathfrak{p}_i$. Set $z = \sum_{i=0}^{n-1} z_i X_i$. By substituting the z_i for the U_i in $U_n \Lambda_0^n$, we see that multiplication by z annihilates M^{D-1}, and hence $(M_{i+1}/M_i)^d$ for $d \geq D-1$ as well. This implies $z^D M_{i+1}/M_i = 0$, which is contrary to the hypotheses because z does not divide zero in M_{i+1}/M_i. So if $\mathfrak{p}_i \neq \mathfrak{m}$, then $\mathfrak{p}_i = \mathfrak{p}$, and since $A/\mathfrak{p} \cong K[X_n]$, $\dim_K(M_{i+1}^d/M_i^d) = 1$ for d sufficiently large.

Now we have shown that $\dim_K M^D$ is precisely the number of indices i such that $\mathfrak{p}_i = \mathfrak{p}$. It is easy to verify that the Artinian ring $A_{\mathfrak{p}}$ contains the field $K(X_n)$ and that $A_{\mathfrak{p}}/\mathfrak{p}A_{\mathfrak{p}} \cong K(X_n)$. This shows that

$$\operatorname{mult}(\mathfrak{p}, M) = \dim_{K(X_n)}(M_{\mathfrak{p}}) = \sum_i \dim_{K(X_n)}(M_{i+1}/M_i)_{\mathfrak{p}}.$$

But $(M_{i+1}/M_i)_{\mathfrak{p}} \cong (A/\mathfrak{p}_i) \otimes A_{\mathfrak{p}}$. So this module is zero if $\mathfrak{p}_i = \mathfrak{m}$, and isomorphic to $K(X_n)$ if $\mathfrak{p}_i = \mathfrak{p}$. This shows then that $\operatorname{mult}(\mathfrak{p}, M) = \dim_K M^D = r$ and concludes the proof of Theorem 7.1. □

8. Complexity

The complexity of the algorithm we have described obviously depends on the field K over which we work. There are three important cases which lead to different complexity estimates.

(a) The complexity of operations in $\overline{K}$ is constant. This is the case when we work in the field of complex numbers.

(b) The complexity of operations of K is constant but not that of operations in $\overline{K}$. This is the case when K is a finite field.

(c) The complexity of operations in K is not constant. This is the case when K is the field of rational numbers.

Furthermore, the complete solution includes the solution of one (or several) equations in one unknown. This complexity also depends on the field over which we work. If, in particular, the field K is $\mathbb{Q}$ or a finite field, this solution is in fact a factorization, and the total complexity of the described algorithm depends on that of one such solution-factorization.

Finally, our algorithm's complexity depends heavily on the algebraic and geometric structure of the solution set.

On the other hand, the difficulty of the problem depends on numerous parameters, namely the number of equations k, the number of variables n and the sequence of degrees $d_0, \ldots, d_{k-1}$.

All this means that a complete complexity study would go outside the framework of this article, and that we will limit ourselves to the particular case where every equation has the same degree d.

Lemma 8.1. *If every equation has the same degree d, the number L of rows and the number C of columns of the matrix Φ in Section 4 satisfy*

$$L \leq (ed)^n \qquad \text{and} \qquad C \leq k(ed)^n.$$

Proof. Indeed $D = (n+1)d - n$,

$$L = \frac{(D+1)\cdots(D+n)}{n!} \leq \frac{((n+1)d)^n}{n!} \leq (ed)^n$$

and

$$C = k\frac{(D+1-d)\cdots(D+n-d)}{n!} \leq kL. \qquad \square$$

Proposition 8.1. *The algorithm's step described at the beginning of Section 4 requires $O\left(k(ed)^{3n}\right)$ operations in K and simultaneous storage of at most $O\left((ed)^{2n}\right)$ elements of K.*

Proof. This follows immediately from Gauss reduction which is carried out bearing in mind that we need only store the matrix C, not the matrix Φ. $\square$

Proposition 8.2. *If there are N solutions counted with their multiplicity, the step of the algorithm described in Section 5 requires simultaneously storing at most $O(ne^n d^n N)$ elements of K. The number of operations in K is $O(ne^n d^n N^3)$, but only $O(ne^n d^n N^2)$ if there are no solutions at infinity.*

Proof. The matrix Λ in Section 4 has N rows of course. Its number of columns is bounded above by $(ed)^n$ and each of its coefficients consists of $n+1$ elements of K. To write this matrix, we also need the last N rows of the matrix C which have at least $(ed)^n$ columns. It follows that writing Λ requires at most $(n+2)e^n d^n N$ memory. It is easy to verify that computing $\Lambda = CL$ requires no operations of K.

Reducing Λ consists of a series of Gauss reductions. When the number of pivots found equals the number of rows of Λ, we can extract a square submatrix from Λ. In general, there is no solution at infinity, and the step is terminated after a number of pivotings equals the number of rows N of Λ. This leads then to at most $O\left(N^2(n+1)(ed)^n\right)$ operations. If there are many solutions on the coordinate hyperplanes, there can be up to N square matrices $U_k\Lambda_0^k + \Lambda_1^k$ isolated by the algorithm of Section 5, which can lead to at most $\frac{1}{2}N(N+1)$ pivots, and hence $O(ne^n d^n N^3)$ operations. $\square$

Remark. This estimation is very crude, because the separation of a square submatrix does not completely destroy the pivoting that was already carried out on the rest of the matrix.

Theorem 8.1. *Consider system* (8)*, all of whose equations have degree at most d, and which has N simple solutions all of whose coordinates are distinct and nonzero (general position). According to the algorithm, solving this system requires at most*

- *$O(ke^{3n}d^{3n})$ operations in K*
- *$O(nN^4)$ operations in degree N extensions of K or in the algebraic closure of K.*
- *Computing the eigenvalues of an $N \times N$ matrix or solving a degree N equation in one unknown and $O(N^{3.8})$ operations in K.*

Proof. Since $N \le d^n$ (Bezout's theorem), we have $nN^2e^n d^n < ke^{3n}d^{3n}$ and together, the first two steps require $O(ke^{3n}d^{3n})$ operations. Computing first coordinates of the solutions requires either computing the eigenvalues of an $N \times N$ matrix, or computing and factoring its characteristic polynomial. Computing the polynomial by interpolation requires that of N determinants in K, so at least $O(N^{3.8})$ operations, the complexity of interpolation being negligible.

Knowing the first coordinate of a solution, we can apply the new algorithm of Section 5 to obtain all the others in $O(nN^3)$ operations. Then together, the N solutions require $O(nN^4)$ operations in extensions of K, which proves the theorem. $\square$

Corollary. *If the solutions satisfy the hypothesis of Theorem 8.1, and if $k \le 2n$ and $d \ge 3$, the algorithm is polynomial relative to the theoretic maximal number of solutions.*

Proof. This maximal number is d^n (Bezout's theorem). It suffices then to verify that operations in a degree d^n algebraic extension of K are polynomials in terms of operations of K, and that the solution of an algebraic equation is also. The first point is easy. The second depends heavily on the field K considered. If K is the field of real or complex numbers, this is indeed the case, provided that we are limited to a given precision. The same is true if K is a finite field, however the corresponding result for the field of rational numbers is still conjectural [3]. $\square$

The above estimate of the complexity of the algorithm being considered, although very partial, enables us to compare it to algorithms based on the resultant.

Proposition 8.3. *Under the hypotheses of Theorem 8.1, algorithms based on the resultant require at least $d^{2^{n-1}}$ operations in K and at least $d^{2^{n-1}}$ elements of K in memory.*

Proof. Using the resultant to eliminate a variable doubles the degrees of the polynomials. Then this method leads to computing and manipulating a polynomial of degree $d^{2^{n-1}}$. $\square$

Remarks. (1) This result shows that for large values of n, our algorithm is faster than those based on the resultant. For the sake of theoretical proofs and experimental comparisons from here on, it will be necessary to determine for which values of n our algorithm is more effective. We conjecture that this is the case from $n = 2$ on.

(2) The matrices Φ and L resemble Sylvester matrices. More precisely, the Sylvester matrix is the matrix Φ in the case where $n = k = 2$. Now we know how to carry out a Gauss reduction on the Sylvester matrix in a number of operations on the order of its dimension squared. In the general case, how do we use the structure of the matrices Φ and L to speed up reduction?

9. Conclusion

To conclude, we would like to return to the advantages and limitations of the method we have described.

The main drawback is certainly that it is only applicable when there are only finitely many solutions. It can be troublesome when this non-finiteness comes solely from solutions at infinity which may be irrelevant for the problem being considered. Further generalization of the method in the case of an infinite number of solutions doesn't seem very likely. In any case, it will require the entire arsenal of modern algebraic geometry.

We have already mentioned that our method lets us obtain all possible algebraic and geometric information of the solution set, but we feel it important to emphasize that the algorithm itself is 'geometric'. Let us say precisely what is meant by this. Every reasonable problem on polynomials, for example solving systems of algebraic equations, is in some sense invariant under linear change of variables. It is this invariance that constitutes the geometric nature of the problem. So an algorithm may be called geometric if it doesn't change under this invariance. It may be useful to note that the invariance that arises here from the linearity in the U_i, is not only preserved, but is used in an essential way to Section 6. It seems to us that the geometric nature of problems on polynomials has been insufficiently utilized in the majority of algorithms concerned, and that in the majority of cases, 'geometric' methods are bound to be much more efficient than classical methods where polynomials are viewed as polynomials in one variable with polynomial coefficients.

Aside from the first step, the complexity of our method depends on the exact number of solutions. This point of view should not be overlooked in a complete complexity study because, in general, an interesting problem has only a few solutions.

To improve our method, and compute its complexity with precision, it will be necessary to study in order of priority the following points.

- How do we use the particular structure of the matrix Φ to speed up the first step, especially when some of the given polynomials are sparse?
- How do we manage the recursion of the step in Section 5, or how do we compute the complexity of it to resolve the following paradox? The existence of solutions with zero coordinates seems to increase the complexity, whereas in practice it simplifies solving.
- Comparison with existing algorithms. We could not carry out this comparison for lack of having them at our disposal. A non-trivial aspect of this comparison is to verify that it is easy to write code for our algorithm using floating point variables in any programming language whereas earlier algorithms require first using or adapting programs for handling polynomials.

Appendix A.1. Proof of Theorem 3.1

Theorem 3.1 is easily deduced from results found in any commutative algebra treatise. We show here how they are deduced from that of Kaplansky [5].

Proposition A.1. *With the notation of Section 3, if L is an algebraically closed extension of K, the function $(b_1, \ldots, b_n) \mapsto (X_1 - b_1, \ldots, X_n - b_n)$ defines a bijection from the solution set of system* (8) *in L onto the set of maximal ideals of B_L.*

Proof. This is Hilbert's Nullstellensatz [5, Theorem 32]. □

Proposition A.2. *Assertions* (i) *and* (iii) *of Theorem 3.1 are equivalent.*

Proof. If $\dim_K B = \dim_{\overline{K}}(B_{\overline{K}})$ is finite, the length of $B_{\overline{K}}$ is finite and every prime ideal of $\overline{B}$ is maximal and minimal [5, Theorem 89]. Hence these ideals are finite in number [5,Theorem 88], and the same is true for the solutions in $\overline{K}$ of system (8) (Proposition A.1). Conversely, if $B_{\overline{K}}$ has only finitely many maximal ideals, its dimension is zero [5, §1-3, Example 4] because it is a Hilbert ring obtained as a quotient of a polynomial ring [5, Theorem 31]. Hence the ring $B_{\overline{K}}$ has finite length [5, Theorem 89], and therefore finite dimension as a $\overline{K}$-vector space: a simple $B_{\overline{K}}$-module is isomorphic to the quotient of $B_{\overline{K}}$ by a maximal ideal. Such a quotient is isomorphic to $\overline{K}$ (Proposition A.1), and is therefore a $\overline{K}$-vector space of dimension 1. Thus we have shown that successive quotients of the composition series of $B_{\overline{K}}$ are finite dimensional $\overline{K}$-vector spaces. This implies immediately that the same is true for $B_{\overline{K}}$. □

It is now easy to finish the proof of Theorem 3.2. It is clear that (ii) implies (i). If (iii) is satisfied, then

$$\dim_L(B_L) = \dim_K(B),$$

and Proposition A.2 applied to the ring B_L shows that system (8) has only finitely many solutions in the algebraic closure of L, and hence in L as well.

Let $(\xi_1, \dots, \xi_n)$ be a solution in the field L. Free to enlarge L, we may assume that L is an algebraically closed extension of $\overline{K}$. Consider the maximal ideal $\mathfrak{m} = (X_1 - \xi_1, \dots, X_n - \xi_n)$ of B_L (Proposition A.1). The intersection $\mathfrak{m} \cap B_{\overline{K}}$ is a prime ideal of $B_{\overline{K}}$ which is maximal if (iii) is satisfied (proof of Proposition A.2), and hence has the form $(X_1 - \zeta_1, \dots, X_n - \zeta_n)$ with ζ_i in $\overline{K}$ (Proposition A.1). It is easy to verify that $\xi_i = \zeta_i$ for all i, which shows that the ξ_i are algebraic over K.

With the inequalities on the number of solutions now obvious, it remains to show that the number of solutions in $\overline{K}$ is bounded above by $\dim_K B = \dim_{\overline{K}}(B_{\overline{K}})$. By showing that $\dim_{\overline{K}}(B_{\overline{K}})$ is finite, we have in fact shown that this number is the length of $B_{\overline{K}}$. If $\mathfrak{m}$ is a maximal ideal of $B_{\overline{K}}$, the simple module $B_{\overline{K}}/\mathfrak{m}$ appears of course as a quotient in a composition series of $B_{\overline{K}}$, and hence in the whole thing. This shows that the number of maximal ideals (i.e. solutions) is bounded above by the length of $B_{\overline{K}}$ (i.e. $\dim_K(B)$), which concludes the proof of Theorem 3.1

Appendix A.2. Proof of Theorem 3.2

To prove this theorem, we need another form of the Nullstellensatz.

Proposition A.3. *For every extension L of K, the function that associates to $(a_0, \dots, a_n)$ the ideal generated by $a_i X_j - a_j X_i$ such that $0 \le i, j \le n$ is an injection from the (projective) solution set in L of system* (9) *into the set of graded prime ideals of A_L, maximal among those not containing A_L^1.*

If L is algebraically closed, this function is a bijection.

Proof. If $(a_0, \dots, a_n)$ is a solution, there exists an i such that $a_i \neq 0$. Suppose $i = 0$ for example, and let I denote the ideal generated by the $a_i X_j - a_j X_i$ in $L[X_0, \dots, X_n]$. This ideal is generated by the

$$X_j - \frac{a_j}{a_0} X_0$$

because

$$a_i X_j - a_j X_i = a_i \left(X_j - \frac{a_j}{a_0} X_0 \right) - a_j \left(X_i - \frac{a_i}{a_0} X_0 \right).$$

It follows that $L[X_0, \dots, X_n]/I$ is isomorphic to $L[X_0]$, which shows that I is prime and that the only graded prime ideal containing I is the one generated by the X_i.

Since the F_i are homogeneous, the equality

$$F_i(a_0, \dots, a_n) = 0$$

implies

$$F_i \left(a_0 \frac{X_0}{a_0}, \dots, a_n \frac{X_0}{a_0} \right) = 0.$$

In other words, $F_i \equiv 0 \mod (X_0 - a_0(X_0/a_0), \dots, X_n - a_n(X_0/a_0))$, which shows that I contains every F_i, and that modulo the F_i, the ideal I is indeed a graded prime ideal of A_L and maximal among the graded prime ideals not containing A^1_L.

Let I' be the ideal generated by the $a'_iX_j - a'_jX_i$ for a different solution $(a'_0, \dots, a'_n)$. It is easy to see that $I = I'$ if and only if the a'_i are proportional to the a_i, which shows injectivity of the function.

Suppose L is algebraically closed, and consider a graded prime ideal J of A_L not containing A^1_L. The Nullstellensatz [5, Theorem 33] shows that there exist $a_0, \dots, a_n$ in L, not all zero, which annihilate the elements of J. This implies that J is contained in the ideal generated by the $a_iX_j - a_jX_i$, and that the function is surjective. □

Proposition A.4. *Let A be a graded ring such that A^0 is a field and $\mathfrak{m}$ is the ideal generated by A^1.*

(a) *The ideal $\mathfrak{m}$ is the only graded prime ideal if and only if there exists an integer D such that $A^d = 0$ for $d \geq D$.*

(b) *The graded prime ideals other than $\mathfrak{m}$ are all minimal if and only if there exists an integer D such that $\dim_K(A^d) = \dim_K(A^D)$ for all $d \geq D$.*

Proof. If we localize the ring A at the prime ideal $\mathfrak{m} = AA_1$, it is immediate that

$$\dim_K(A_\mathfrak{m}/\mathfrak{m}^n A_\mathfrak{m}) = \dim_K(A/\mathfrak{m}^n) = \sum_{i=0}^{n-1} \dim_K(A^i).$$

Therefore, if A^d is zero for d sufficiently large, the dimension of the ring A is 0 and $\mathfrak{m}$ is a minimal prime ideal [10, Chapter III, Theorem 1]. If $\dim_K(A^d)$ is constant for d sufficiently large, every prime ideal strictly contained in $\mathfrak{m}$ is minimal (ibid.) Conversely, since every minimal prime ideal of A is graded [2, Chapter IV, §3, Proposition 1], if $\mathfrak{m}$ is the only graded prime ideal, it is a minimal prime ideal and the dimension of the ring $A_\mathfrak{m}$ is zero. To show the converse of (b), consider a (graded) minimal prime ideal $\mathfrak{p}$ of A, and a homogeneous element x that does not lie in $\mathfrak{p}$. Any minimal prime ideal containing $\mathfrak{p}$ and x is graded [2, loc. cit.] and hence equal to $\mathfrak{m}$ if all the other graded prime ideals are minimal. It follows that there exist no prime ideals between $\mathfrak{p}$ and $\mathfrak{m}$ [5, Theorem 142] and we conclude with the theorem cited in [10]. □

We can now prove Theorem 3.2.

(iv) ⇒ (iii): Assertion (iv) implies that $\dim_L(A^d_L)$ is decreasing for $d \geq D'$. Then there exists an integer D such that $\dim_L(A^d_L) = \dim_L(A^D_L)$ for all $d \geq D$. Assertion (iii) follows because $\dim_K(A^d) = \dim_L(A^d_L)$ for all d.

(iii) ⇒ (ii): If (iii) is satisfied, $\dim_L(A^d_L) = \dim_K(A^D)$ for all extensions L of K and all integers $d \geq D$. If $(a_0, \dots, a_n)$ is a solution in L of system (9), then Proposition A.4 implies that the prime ideal generated by the $a_iX_j - a_jX_i$ is minimal. Hence these ideals are finite in number [5, Theorem 88], which implies assertion (ii) (Proposition A.3).

(ii) ⇒ (i): Obvious.

(i) ⇒ (iv): Show first assertion (iv) limited to infinite fields L contained in $\overline{K}$. If $(a_0, \dots, a_n)$ is a solution in $\overline{K}$, the vector subspace of $A^1_{\overline{K}}$ generated by $a_iX_j - a_jX_i$ is a hyperplane whose intersection with A^1_L is a proper subspace. Since there are only finitely many such subspaces, there exists $y = \sum y_iX_i$ in A^1_L which does not lie in any of them.

Since y is homogeneous, the prime ideals of $A_{\overline{K}}$, minimal among those containing y, are graded. Since $\mathfrak{m}A_{\overline{K}}$ is the only graded prime ideal containing y (Proposition A.3), there exists an integer d such that $\mathfrak{m}^d A_{\overline{K}} \subset A_{\overline{K}} y$. In addition, the annihilator of y in $A_{\overline{K}}$ is generated by a finite number of homogeneous elements $z_1, \dots, z_n$. Let d' be the largest degree of the z_i. Then if z is a homogeneous element of degree $\geq d + d'$ in $\mathrm{Ann}_{\overline{K}}(y)$, then

$$z \in \sum_i \mathfrak{m}^d A_{\overline{K}} z_i \subset \sum_i A_{\overline{K}} y z_i = 0.$$

On the other hand, consider the ring $B = A_L/(y-1)A_L$. It satisfies the conditions of Theorem 3.1, because there is a bijection between the projective solutions of system (9) and the system obtained by adjoining the equation $\sum y_iX_i = 1$ to system (9). Thus B is a finitely generated L-vector space and there exists an integer d'' such that every element of B is the image of an element of degree $\leq d''$ in A_L. Set $D' = \max(d'', d + d')$. Then if $t \in A^i_L$ for $i > D'$, we have

$$t = (y-1)u + v$$

with $\deg(v) \leq d'' < \deg(t)$. If u' is the highest degree homogeneous part of u, then $t = yu'$ because if $yu' = 0$, we would have $\deg(u') \geq d + d'$ and hence $u' = 0$.

Therefore, assertion (iv) is proved when $L \subset \overline{K}$, and hence (i) $\Rightarrow$ (iii). If L is not contained in $\overline{K}$ and if (i) is satisfied, then $\dim_L(A_L^d)$ is constant for d sufficiently large. We can apply the equivalence already proved substituting L for K, which proves (iv) completely.

(a) Follows immediately from (iii) and (iv).

(b) Resume the notation of the proof of (i) $\Rightarrow$ (iv). The projective solutions of system (8) are in bijection with the solutions of the system obtained by adjoining $y = 1$ to system (9). Then their number is bounded above by $\dim_L(B)$. We show that if $d > D'$, the surjection from A_L onto B induces a bijection from A_L^d onto B. Let $z \in A_L^d$. If $z = (y-1)t$, the highest degree term of t will be cancelled by y (because z is homogeneous). This shows that the function $A_L^d \to B$ is injective. Every element of B is the sum of elements that are images of homogeneous elements of A_L. Now if $z \in A_L^i$ with $i < d$, then $z \equiv y^{d-i}z \mod (y-1)$. According to (iv), if $z \in A_L^i$ with $i > d$, then $z = y^{i-d}t$, which shows that the function $A_L^d \to B$ is a bijection and $\dim_K(A^d) = \dim_L(B)$.

(c) Every solution in L of system (9) is proportional to a solution that satisfies $y = 1$. According to Theorem 3.1, this solution is algebraic over K, which proves (c).

(d) Immediate

(e) This is Proposition A.4(a).

Appendix A.3. Proof of Theorem 3.3

Theorem 3.3 follows easily from Theorem 1 of [7]. It may be useful to give a direct proof.

By Theorem 3.2, we may assume the field K is infinite. The proof is by induction on the number n of indeterminates. For this, set

$$R = K[X_0, \dots, X_n],$$

and let the ideal of R generated by $F_0, \dots, F_{k-1}$ be denoted by I. Thus the ring A we study is R/I. Consider the Koszul complex or the exterior algebra of I (for the terminology and basic concepts, one can consult any book on homological algebra, for example [9]). It is the complex

$$\Lambda : 0 \longrightarrow \Lambda_k \xrightarrow{\delta_k} \Lambda_{k-1} \xrightarrow{\delta_{k-1}} \dots \xrightarrow{\delta_2} \Lambda_1 \xrightarrow{\delta_1} \Lambda_0 \longrightarrow 0$$

in which Λ_α is a free R-module having for a basis the elements $e_{i_1} \wedge \cdots \wedge e_{i_\alpha}$ such that $0 \leq i_1 < i_2 < \cdots < i_\alpha < k$ and $\Lambda_0 = R$. The function δ_α is defined by

$$\delta_\alpha(e_{i_1} \wedge \cdots \wedge e_{i_\alpha}) = \sum_{j=1}^{\alpha} (-1)^{j+1} F_j e_{i_1} \wedge \cdots \wedge e_{i_{j-1}} \wedge e_{i_{j+1}} \wedge \cdots \wedge e_{i_\alpha}$$

and

$$\delta_1(e_i) = F_i.$$

If we assign the degree $d_{i_1} + \ldots + d_{i_\alpha}$ to $e_{i_1} \wedge \ldots \wedge e_{i_\alpha}$ (d_i is the degree of F_i), the modules Λ_i are graded and the functions δ_i are homogeneous. Hence for every degree d, the restrictions δ_i^d of δ_i to degree d homogeneous parts Λ_i^d of Λ_i form a complex Λ^d of finitely generated K-vector spaces. Let $H_i^d = \ker(\delta_i^d)/\operatorname{im}(\delta_{i+1}^d)$ denote their homology modules. Then $H_0^d = A^d$.

Suppose the equivalent conditions of Theorem 3.2 are satisfied and choose y in A^1 such that multiplication by y from A^d into A^{d+1} is bijective for sufficiently large d (recall that we have assumed K is infinite). Here the element y comes from an element Y in R^1, to which we can associate the exact sequence

$$0 \longrightarrow R^{d-1} \xrightarrow{Y} R^d \longrightarrow (R/YR)^d \longrightarrow 0.$$

By tensoring the complex Λ^d by this exact sequence, we obtain an exact sequence of complexes, and hence an exact homology sequence

$$\dots \longrightarrow \overline{H}_{i+1}^d \longrightarrow H_i^{d-1} \xrightarrow{Y} H_i^d \longrightarrow \overline{H}_i^d \longrightarrow \dots$$

where $\overline{H}_i^d$ denotes the homology of the complex $\overline{\Lambda} = \Lambda \otimes (R/YR)$. But $\overline{\Lambda}$ is the Koszul complex of the ideal $\overline{I}$ generated by the images of $F_0, \dots, F_{k-1}$ in the ring R/YR which is isomorphic to $K[X_0, \ldots, X_{n-1}]$. Therefore, Theorem 3.3 is deduced from Theorem A.5 below, which will itself be proved by induction on n.

Theorem A.5. *Suppose the F_i are sorted by decreasing degree and are not all zero. If $A^d = 0$ for all d sufficiently large, then $H_i^d = 0$,*

(a) *for all d if $i \geq k - n$,*

(b) *for all $d \geq d_0 + \ldots + d_{i+n} - n$ if $i < k - n$.*

Proof. We assume the result is proved for $n = 0$ and show it by induction for all n. The induction hypothesis implies that $\overline{H}_{i+1}^d = 0$ for all d if $i + 1 \geq k - n + 1$ and for $d \geq d_0 + \ldots + d_{i+n} - n + 1$ if $i < k - n$. Suppose $H_i^d = 0$ for large values of d. The exact sequence

$$\overline{H}_{i+1}^d \longrightarrow H_i^{d-1} \longrightarrow H_i^d$$

shows by induction on d that $H_i^{d-1} = 0$ for all d if $i \geq k - n$ and for $d \geq d_0 + \ldots + d_{i+n} - n + 1$ if $i < k - n$. Then we recall the following lemmas.

Lemma A.6. *If $A^d = 0$ for d sufficiently large, the same is true for H_i^d for all i.*

Proof. The hypothesis implies that the ideal $\mathfrak{m} = (X_0, \ldots, X_n)$ is the only prime ideal containing I. If $\mathfrak{p}$ is another prime ideal of R, the complex $\Lambda \otimes R_\mathfrak{p}$ is the Koszul complex of the ideal $I \otimes R_\mathfrak{p}$ of the localization $R_\mathfrak{p}$, but $I \otimes R_\mathfrak{p} = R_\mathfrak{p}$ and $\delta_1 \otimes R_\mathfrak{p}$ is surjective. Let ε be in $\Lambda^1 \otimes R_\mathfrak{p}$ such that $(\delta_1 \otimes R_\mathfrak{p})(\varepsilon) = 1$. Setting $\varepsilon_\alpha(e_{i_1} \wedge \cdots \wedge e_{i_\alpha}) = \varepsilon \wedge e_{i_1} \wedge \cdots \wedge e_{i_\alpha}$, we verify easily that $\varepsilon_{\alpha-1} \circ (\delta_\alpha \otimes R_\mathfrak{p}) + (\delta_{\alpha+1} \otimes R_\mathfrak{p}) \circ \varepsilon_\alpha = \mathrm{id}(\Lambda_\alpha \otimes R_\mathfrak{p})$, which shows that the homology $H_i \otimes R_\mathfrak{p}$ of $\Lambda \otimes R_\mathfrak{p}$ is zero for all $\mathfrak{p} \neq \mathfrak{m}$. Thus H_i is annihilated by a power of $\mathfrak{m}$, which implies the result. $\square$

Lemma A.7. *Theorem A.5 is true for $n = 0$.*

Proof. Assertion (a) is immediate. If $i < k$ the above lemma shows that if $x \in \ker \delta_i$, there exists an integer h such that $X_0^h x = \delta_{i+1}(z)$. Since the degree of the basis elements of Λ_{i+1} is at most $d_0 + \ldots + d_i$, their coefficients are divisible by X_0^h if $\deg(x) \geq d_0 + \ldots + d_i$. We can then divide by X_0^h, which shows that $x \in \mathrm{im}(\delta_{i+1})$. $\square$

$\square$

Theorem 3.3 is now deduced very easily from Theorem A.5 and the exact sequence

$$\overline{H}_1^d \longrightarrow H_0^{d-1} \xrightarrow{Y} H_0^d \longrightarrow \overline{H}_0^d.$$

In fact, Theorem 3.2 implies that $(R/YR)^d$ is zero for d sufficiently large. It follows by Theorem A.5 that $\overline{H}_i^d = 0$ for all d if $i \geq k - n + 1$ and for $d \geq d_0 + \ldots + d_{i+n-1} - n + 1$ otherwise. This implies that multiplication by Y is surjective if $d \geq d_0 + \ldots + d_n + 1 - n$ and injective if $k = n$ or $d - 1 \geq d_0 + \ldots + d_{n-1} - n$. Therefore, we may take (notations from Theorem 3.2)

$$D' = d_0 + \ldots + d_{n-1} + 1 - n$$

and

$$D = d_0 + \ldots + d_{n-1} - n \qquad \text{if } k = n$$

and

$$D = d_0 + \ldots + d_n - n \qquad \text{if } k > n$$

(the hypotheses imply that $k \geq n$).

Then with the convention that $d_n = 1$ if $k = n$, we have $\max(D, D') = d_0 + \ldots + d_n - n$, which proves Theorem 3.3.

Appendix A.4

We give here an explicit form of the algorithm from Sections 4 and 5. The comments between brackets let us make the connection between the two presentations of the algorithm and use the notation from Sections 4 and 5.

1. Input Data

These are:

- the number K of polynomials
- the number N of variables
- the list $D(0), \ldots, D(K-1)$ of degrees of the polynomials

- the polynomials themselves, encoded for example by the list of their coefficients in lexicographic order. Let $F(I,1),\ldots,F(I,M)$ be these coefficients for the I-th polynomial with
$$M = \binom{D(I)+N}{N}.$$

2. Initialization

(I1) IF $K < N$ THEN DO
WRITE 'there are too many variables'
END program
END

(I2) $D(K) \leftarrow 1$

(I3) Let $E(0),\ldots,E(K)$ be the $D(I)$ sorted in decreasing order

(I4) $DD \leftarrow E(0)+\ldots+E(N)-N$

(I5) $NL \leftarrow (DD+1)*\cdots*(DD+N)/(1*\cdots*N)$
[NL is the number of rows of the matrix Φ = dimension R^D]

(I6) Form an $NL \times NL$ matrix C initialized to the identity matrix

(I7) Form a vector PHI of dimension NL
[this will be the current column of Φ]

3. Reduction of the matrix Φ

(R1) FOR $I = 1$ to NL, $B(I) \leftarrow 0$
[the I for which $B(I) \neq 0$ will be the indices of the rows of Φ where we have already found a pivot]

(R2) FOR $I = 0$ to $K-1$ DO

(R3) $DI \leftarrow DD - D(I)$
$NC \leftarrow (DI+1)*\cdots*(DI+N)/(1*\cdots*N)$
[NC is the number of columns of Φ corresponding to the I-th polynomial]

(R4) FOR $J = 1$ to NC DO

(R5) FOR $J1 = 1$ to NL, $PHI(J1) \leftarrow 0$
[re-initialization of PHI]

(R6) Let $E(1),\ldots,E(N)$ be the exponents of the J-th monomial of degree DI (in the lexicographic order)

(R7) FOR $J1 = 1$ to NL DO

(R8) Let $EE(1),\ldots,E(N)$ be the exponents of the $J1$-th monomial of degree DD

(R9) IF $EE(J2) \geq E(J2)$ FOR $J2 = 1$ to N
THEN $PHI(J1) \leftarrow F(I,J3)$
where $F(I,J3)$ is the coefficient of the monomial with exponents $EE(1)-E(1),\ldots,EE(N)-E(N)$ in the I-th polynomial

(R10) END DO (R7)
[PHI is now the J-th column of the part of Φ corresponding to the I-th polynomial]

(R11) $PHI \leftarrow C * PHI$
[The operation $*$ denotes a matrix product]

(R12) IF possible, choose $J1$ such that $1 \leq J1 \leq NL$ and $PHI(J1) \neq 0$ and $B(J1) = 0$
[$PHI(J1)$ is the pivot; choose the maximal absolute value of it if the known quantities are real or complex; choose the numerator and small denominator of it if the known quantities are rational]
ELSE GO TO (R15)

(R13) $B(J1) \leftarrow 1$

(R14) FOR $J2 = 1$ to NL
IF $B(J2) = 0$ and $PHI(J2) \neq 0$ THEN
FOR $J3 = 1$ to NL
$C(J2,J3) \leftarrow C(J2,J3)$
$-(PHI(J2)/PHI(J1)) * PHI(J1,J3)$
[Gauss reduction; if the known quantities are integers or polynomials, there is reason to multiply the two elements by $PHI(J1)$ and divide the result by the previous pivot, in order to limit the growth of the size of the known quantities]

(R15) END DO (R4)

(R16) END DO (R2)
[the actual value of C is the one considered in Section 4]

(R17) $SR \leftarrow \#\{I : B(I) = 0\}$
[SR is the integer $s-r$ of Section 4]

(R18) IF $SR = 0$ THEN DO
WRITE 'no solution'
END program
END

(R19) Let CC be the $R \times NL$ matrix obtained by omitting from C the rows for which $B(J1) = 1$
[this is the matrix of 'the last $s - r$ rows of C']

4. Reduction of Λ

[It is first necessary to write the matrix Λ]

(C1) $NC \leftarrow NL * DD/(DD + N)$
[number of columns of Λ]

(C2) Form an $SR \times NC \times (N + 1)$ matrix LAM
[LAM will be viewed as an $SR \times NC$ matrix of vectors of dimension $N + 1$, the components of such vectors being the coefficients of polynomials of degree 1]

(C3) FOR $I = 1$ to NC DO

(C4) FOR $J = 0$ to N DO

(C5) FOR $I1 = 1$ to SR
$LAM(I1, I, J) = CC(I1, I2)$
where $I2$ is the index among the monomials of degree at most DD of the product of the J-th variable and the monomial of degree at most $DD - 1$ of index I (by convention, the 0-th variable is the element 1)

(C6) END DO (C4)

(C7) END DO (C3)

(C8) Call the procedure REDUC
[The reduction being done recursively, the procedure REDUC calls itself]

The procedure REDUC

(F1) $LPIV \leftarrow 0$ [row of the last pivot]

(F2) FOR $J = 0$ to N DO
[J is the index of the current variable]

(F3) $CPIV \leftarrow 0$ [column of the last pivot]

(F4) IF $LAM(I1, I, J) = 0$
FOR $I1 = LPIV + 1$ to SR
FOR $I = CPIV + 1$ to NC
THEN GO TO (F11)
ELSE choose $I1, I$ such that $LAM(I1, I, J) \neq 0$
[pivot]

(F5) $CPIV \leftarrow CPIV + 1$
$LPIV \leftarrow LPIV + 1$

(F6) Interchange rows $I1$ and $LPIV$ and columns I and $CPIV$ of LAM.

(F7) FOR $I2 = 1$ to SR
FOR $I3 = CPIV + 1$ to NC
FOR $I4 = 0$ to N
$LAM(I2, I3, I4) \leftarrow LAM(I2, I3, I4)$
$-(LAM(LPIV, I3, J)/LAM(LPIV, CPIV, J))$
$*LAM(I2, CPIV, I4)$
[Gauss reduction on the columns]

(F8) IF $LPIV = SR$ THEN GO TO (F13)

(F9) FOR $I2 = LPIV + 1$ to SR
FOR $I3 = 1$ to NC
FOR $I4 = 0$ to N
$LAM(I2, I3, I4) \leftarrow LAM(I2, I3, I4)$
$-(LAM(I2, CPIV, J)/LAM(LPIV, CPIV, J))$
$*LAM(LPIV, I3, I4)$
[Gauss reduction on the rows]

(F10) GO TO (F4)

(F11) END DO (F2)

(F12) IF $LPIV < SR$ THEN DO
WRITE 'there are infinitely many solutions'
END program
END
[indeed, the rank of LAM is less than its number of rows (see Section 6.1(a))]

(F13) Verify that $LAM(I1, I2, I3) = 0$
FOR $I1 = SR - CPIV + 1$ to SR
FOR $I2 = CPIV + 1$ to NC
FOR $I3 = 0$ to N
ELSE DO
WRITE 'there are infinitely many solutions'
END program
END
[it is a matter of verifying (b) of Section 6.1]

(F14) Store the submatrix of LAM comprising the coefficients $LAM(I1, I2, I3)$ for which $I1 > SR - CPIV$ and $I2 \leq CPIV$

(F15) Omit the coefficients $LAM(I1, I2, I3)$ of LAM for which $I1 > SR - CPIV$ or $I2 \leq CPIV$.

(F16) $SR \leftarrow SR - CPIV$
$NC \leftarrow NC - CPIV$

(F17) IF $SR \neq 0$ THEN call REDUC

(F18) END REDUC
[The polynomial G of Theorem 4.1 is the product of the determinants of the matrices stored in (F14)]

Bibliography

[1] N. Bourbaki. *Algèbre*, new edition. Hermann, 1970. Chapter II, §10, Definition 1.

[2] N. Bourbaki. *Algèbre Commutative.* Hermann, 1962. Chapters III–IV.

[3] G.E. Collins. Factoring Univariate Polynomials in Polynomial Average Time. In *Symbolic and Algebraic Computation.* Springer LNCS **72**, 1979, 317–329.

[4] R. Hartshorne. *Algebraic Geometry.* Springer GTM 52, 1977.

[5] I. Kaplansky. *Commutative Rings.* Allyn & Bacon, 1970.

[6] D. Lazard. Algèbre Linéaire sur $K[X_1, \ldots, X_n]$ et Élimination. *Bull. Soc. Math. France* **105** (1977), 165–190.

[7] D. Lazard. Systems of Algebraic Equations. In *Symbolic and Algebraic Computation.* Springer LNCS **72**, 1979, 88–94.

[8] F.S. Macaulay. *The Algebraic Theory of Modular Systems.* Cambridge Tracts in Math. **19**, 1916.

[9] D.G. Northcott. *An Introduction to Homological Algebra.* Cambridge, 1960.

[10] J.P. Serre. *Algèbre Locale, Multiplicités.* Springer LNM **11**, 1965.

[11] B.L. van der Waerden. *Moderne Algebra II*, Third Edition. Springer, 1955. English translation, Ungar 1950.

[12] D.Y.Y. Yun. On Algorithms for Solving Systems of Polynomial Equations. *ACM SIGSAM Bull.* **27** (1973), 19–25.

[13] D. Lazard. Problem #7 and Systems of Algebraic Equations. *ACM SIGSAM Bull.* **54** (1980), 26–29.