J. Symbolic Computation (1999) **28**, 3–44
Article No. jsco.1998.0266
Available online at http://www.idealibrary.com on IDEAL®

Matrices in Elimination Theory

IOANNIS Z. EMIRIS† AND BERNARD MOURRAIN‡

INRIA, SAGA, B.P. 93, Sophia-Antipolis, 06902 France

The last decade has witnessed the rebirth of resultant methods as a powerful computational tool for variable elimination and polynomial system solving. In particular, the advent of sparse elimination theory and toric varieties has provided ways to exploit the structure of polynomials encountered in a number of scientific and engineering applications. On the other hand, the Bezoutian reveals itself as an important tool in many areas connected to elimination theory and has its own merits, leading to new developments in effective algebraic geometry. This survey unifies the existing work on resultants, with emphasis on constructing matrices that generalize the classic matrices named after Sylvester, Bézout and Macaulay. The properties of the different matrix formulations are presented, including some complexity issues, with emphasis on variable elimination theory. We compare toric resultant matrices to Macaulay's matrix and further conjecture the generalization of Macaulay's exact rational expression for the resultant polynomial to the toric case. A new theorem proves that the maximal minor of a Bézout matrix is a non-trivial multiple of the resultant. We discuss applications to constructing monomial bases of quotient rings and multiplication maps, as well as to system solving by linear algebra operations. Lastly, degeneracy issues, a major preoccupation in practice, are examined. Throughout the presentation, examples are used for illustration and open questions are stated in order to point the way to further research.

1. Introduction

Resultants provide an essential tool in constructive algebra and in equation solving. The *resultant* of an overconstrained polynomial system characterizes the existence of common roots as a condition on the input coefficients. If we consider the input coefficients as independent indeterminates, then the solutions lie in a space of dimension $n+m$, where n is the number of variables and m is the number of coefficients. The resultant projects the solutions to an m-dimensional space and is, therefore, also known as a *projection operator*. Since it eliminates the input variables, the resultant is also known as the *eliminant* of the given system.

A number of methods exist for constructing *resultant matrices*, which are matrices whose determinant is the resultant or, more generally, a non-trivial multiple of it. These matrices represent the most efficient way for computing the resultant polynomial and for solving systems of polynomial equations by means of the resultant method. An example of a matrix that gives precisely the resultant is the determinant of the coefficient matrix of $n+1$ linear polynomials, or the Sylvester matrix of a polynomial pair. Resultant matrices have been extensively studied around the turn of the century. Their determinants

†E-mail: First.Last@sophia.inria.fr
‡http://www.inria.fr/saga/First.Last/

0747–7171/99/070003 + 42 $30.00/0

have been known as inertia forms. We give a short historical overview of the impact of resultants in effective algebraic geometry and we refer the reader to Muir (1960) for a more complete historical description.

The first major contribution to resultant theory was probably the work of Bézout (1779) (and Euler). By combining two univariate polynomials $P(x)$ of degree m and $Q(x)$ of degree $n \leq m$, Bézout observed that he could obtain m linearly independent polynomials of degree $\leq m$. This yields an m by m matrix, called Bezoutian matrix by Sylvester, whose determinant is the resultant of P and Q (originally called "la résultante de P et Q" in French, see Lascoux, 1986). It was worked out later by Cayley (1848), who connected it with the polynomial

$$\frac{Q(x)P(y) - Q(y)P(x)}{x - y}.$$

Sylvester (1853) preferred to eliminate directly the monomials $1, x, \ldots, x^{m+n-1}$ in the multiples

$$(x^i P(x))_{0 \leq i \leq n-1}, \qquad (x^j Q(x))_{0 \leq j \leq m-1},$$

of the initial polynomials and took the determinant of the corresponding $(m+n) \times (m+n)$ matrix. Although contemporary to related works (Jacobi, 1835; Richelot, 1840; Cauchy, 1840; ...), this method remains well known as Sylvester's resultant.

A generalization of the Bezoutian in several variables was used in the work of Dixon (1908). For three polynomials P_0, P_1, P_2 in two variables, of the same degree, he took some coefficients of their multivariate Bezoutian and added some multiples of the initial polynomials in order to obtain a square matrix. In Section 3.4, we shall return to this construction.

The work of Macaulay (see Macaulay, 1902, 1916; van der Waerden, 1950) generalizes the Sylvester construction to the multivariate case, which in turn, was extended recently to the case of toric varieties. Indeed, the last two decades have witnessed the flourishing of the theory of sparse elimination (Bernstein, 1975; Gelfand *et al.*, 1994); a more complete account is given below. This theory generalizes several results of classical elimination theory on multivariate polynomial systems by considering the structure of the given polynomials, namely the coefficients which are *a priori* zero and the support and Newton polytope defined by the non-zero coefficients. This leads to stronger algebraic and combinatorial results, in general, whose complexity depends on effective rather than total degree. The toric, or sparse, resultant generalizes the classical resultant of $n + 1$ homogeneous polynomials in $n + 1$ variables in the sense that they coincide when all polynomial coefficients are non-zero. The toric resultant coincides with the Sylvester resultant if the system is comprised of two univariate polynomials. Unlike its classical counterpart, however, the toric resultant depends on the non-zero monomials only and therefore it has lower degree for sparse inputs. The generalization of Bézout matrices has also been considered recently in resultant theory (Jouanolou, 1991, 1993a,b; Elkadi and Mourrain, 1999; Busé *et al.*, 1999), yielding new interesting algorithmic developments.

The renewed interest in elimination theory and the associated matrix methods for system solving is manifold. For small- and medium-sized polynomial systems corresponding to zero-dimensional varieties, the resultant matrix provides one of the most efficient solution methods today. This has been established through a number of concrete applications in the forward and inverse kinematics of robots and mechanisms as well as the computation of their motion plans (Canny, 1988; Raghavan and Roth, 1995), the geometric

structure of molecules (Balbes and Mascarella, 1994; Emiris and Mourrain, 1996), geometric and solid modeling, graphics, and computer-aided design (Bajaj *et al.*, 1988; Hoffmann, 1989; Manocha and Demmel, 1995), as well as quantifier elimination (Renegar, 1992; Canny, 1993; Basu *et al.*, 1997), and the solution of systems of inequalities (Grigoryev and Vorobjov, 1988). Bezoutian matrices are used in complexity theory in effective algebraic geometry (Fitchas *et al.*, 1993; Lickteig and Meer, 1995; see also Sabia and Solerno, 1995). They appear to be a fundamental tool in several other domains such as residue theory (Scheja and Storch, 1975; Kunz, 1986) and complex analysis (Aizenberg and Kytmanov, 1981; Berenstein and Yger, 1991; Berenstein *et al.*, 1993). Its use in important applications from an algorithmic point of view is illustrated in Cardinal (1993); Becker *et al.* (1996); Cardinal and Mourrain (1996); Elkadi and Mourrain (1998); and Elkadi and Mourrain (1999).

This survey is organized as follows. The next section sketches the basic notions of elimination theory, starting with notations, then the classical theory and, finally, discusses sparse elimination theory. Section 3 is the main part, where the different matrix formulations are detailed and analyzed: Sylvester and Macaulay matrices in Section 3.1, two algorithms for the toric resultant (or Newton) matrix in Sections 3.2 and 3.3, Bézout and Dixon matrices in the following two sections and, finally, a comparison between different matrix formulations. Then we consider how these matrices can be used for constructing monomial bases, multiplication maps and, ultimately, for solving systems of polynomial equations by different methods in Sections 4.1 to 4.3. Section 4.4 presents methods for handling input degeneracy. We conclude with a summary.

2. Elimination Theory

2.1. NOTATIONS

- $\mathbb{K}$ is the coefficient field; unless otherwise stated, it is arbitrary. $\mathbb{P}^n$ denotes the projective space of dimension n, obtained as a quotient of $\mathbb{K}^{n+1}$ (non-zero multiple vectors are identified). $\overline{\mathbb{K}}$ is the algebraic closure of $\mathbb{K}$. $\mathbb{K}^* = \mathbb{K} - \{0\}$ is the field $\mathbb{K}$ except 0.
- n is the number of variables of the polynomial rings of interest. $x_1, \ldots, x_n$ denote the n independent variables and $\mathbf{x}$ is a shorthand for all of them. The notation $\mathbf{x}^{-1}$ represents $x_1^{-1}, \ldots, x_n^{-1}$. More sets of variables are sometimes needed and denoted by $\mathbf{y}$ and $\mathbf{u}$.
- The set of polynomials in the variables $x_1, \ldots, x_n$, with coefficients in $\mathbb{K}$, will be denoted by $R = \mathbb{K}[x_1, \ldots, x_n]$. $f_1, \ldots, f_{n+1}$ are the input polynomials lying in $R = \mathbb{K}[\mathbf{x}]$. These can be Laurent polynomials lying in $R = \mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]$. Their coefficients are denoted by $\mathbf{c} = [c_{ij}]$.
- I, J are ideals of R and $\mathcal{A} = R/I$ denotes the quotient algebra of polynomials modulo the ideal I. $\mathcal{Z}_{\mathbb{K}}(f_1, \ldots, f_s)$ stands for the zero-set of $f_1 = 0, \ldots, f_s = 0$; that is, the algebraic variety of points $\zeta \in \mathbb{K}^n$ such that $f_1(\zeta) = \cdots = f_s(\zeta) = 0$. If not specified, $\mathcal{Z}(f_1, \ldots, f_s)$ means the zero-set over the algebraic closure $\overline{\mathbb{K}}$ of $\mathbb{K}$. The set of roots ζ of $f_1, \ldots, f_s$ which are in an algebraic variety X is denoted by $\mathcal{Z}_X(f_1, \ldots, f_s)$.
- $\mathbb{I}$ is the identity matrix whose dimension is clear from context or specified by $\mathbb{I}_k$, if it equals k.

- A_i, for $i = 1, \ldots, n+1$, denote the Newton polytopes of the given polynomials, defined in Section 2.4. $Q = \sum_{i=1}^{n+1} A_i$ is the Minkowski sum of the given $n+1$ Newton polytopes and $Q_{-i} = \sum_{j \neq i} A_j$ is the Minkowski sum of n of them, introduced in Section 3.3.
- $V(\cdot)$ is the standard Euclidean volume function. $\mathrm{MV}(\cdot)$ denotes the mixed volume operation on polytopes. Given a set of $n+1$ polytopes, MV_{-i}, for $i = 1, \ldots, n+1$, denotes the mixed volume of n Newton polytopes excluding the ith one. When these operators are applied to polynomials or point sets, we understand the mixed volume of the corresponding Newton polytopes or, respectively, the convex hulls of the points. These operators are defined in Section 2.4.

2.2. ELIMINATION THEORY

Elimination theory deals with the problem of finding conditions on parameters of a polynomial system, so that these equations have a common solution in an algebraic set, that we denote hereafter by X. A typical situation is the case of $n+1$ polynomials

$$\begin{cases} f_1(\mathbf{x}) &= \sum_{j=0}^{k_1} c_{1,j} \psi_{1,j}(\mathbf{x}) \\ & \vdots \\ f_{n+1}(\mathbf{x}) &= \sum_{j=0}^{k_{n+1}} c_{n+1,j} \psi_{n+1,j}(\mathbf{x}) \end{cases}$$

where

- $\mathbf{c} = (c_{i,j})$ are parameters,
- $\mathbf{x}$ is a point of the projective variety $X \subset \mathbb{P}^N$, of dimension n,
- the functions $\psi_{i,j}(\mathbf{x})$ $(j = 0, \ldots, k_i)$ are homogeneous polynomials, independent of the parameters $\mathbf{c}$, and of the same degree in the coordinates of $\mathbf{x} \in \mathbb{P}^N$.

Let us denote by $\mathcal{L}_i(\mathbf{x})$ the vector of polynomial functions $\mathcal{L}_i(\mathbf{x}) = (\psi_{i,j}(\mathbf{x}))_{j=0,\ldots,k_i}$ [†] and by $\mathbf{f_c}(\mathbf{x}) = \mathbf{0}$ the global system of equations.

The elimination problem consists, in this case, in finding necessary (and sufficient) conditions on the parameters $\mathbf{c} = (c_{i,j})_{i,j}$ such that the equations $f_1 = 0, \ldots, f_{n+1} = 0$ have a common root in X. Note that if the number of equations is not greater than the dimension of X, then there is no condition on the parameters. This is the reason why we choose X of dimension n.

The classical situation is the case where $\mathcal{L}_i(\mathbf{x}) = (\psi_{i,j}(\mathbf{x}))_{j=0,\ldots,k_i}$ is the vector of all monomials of degree d_i and where $X = \mathbb{P}^n$ is the projective space of dimension n. The functions $f_i(\mathbf{x})$ are *generic* homogeneous polynomials of degree d_i and the necessary and sufficient condition on the parameters $\mathbf{c} = (c_{i,j})_{i,j}$ such that the homogeneous polynomials $f_1, \ldots, f_{n+1}$ have a common root in $X = \mathbb{P}^n$ is $\mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_{n+1}) = 0$, where $\mathrm{Res}_{\mathbb{P}^n}$ is the *classical projective resultant*.

Considering a geometric point of view, we are looking for the set of parameters $\mathbf{c} = (c_{i,j})$ such that there exists $\mathbf{x} \in X$ with $\sum_j c_{i,j} \psi_{i,j}(\mathbf{x}) = 0$ for $i = 1, \ldots, n+1$. In other words, the parameter vector $\mathbf{c}$ is the projection of the point $(\mathbf{c}, \mathbf{x})$ of the variety

$$W_X = \{(\mathbf{c}, \mathbf{x}) \in \mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_{n+1}} \times X \text{ s.t. } \sum_{j=0}^{k_i} c_{i,j} \psi_{i,j}(\mathbf{x}) = 0 \ ; \ i = 1, \ldots, n+1\}.$$

[†]This notation refers to line bundles, as we will see in the following.

This variety W_X is called the *incidence variety* and we have two projections:

$$\pi_1 : W_X \to \mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_{n+1}},$$
$$\pi_2 : W_X \to X.$$

The image of W_X by π_1 is precisely the set of parameters $\mathbf{c}$ for which the system has a root. The image by π_2 of a point of W_X is a solution in X of the associated system. Any polynomial in $\mathbf{c} = (c_{i,j})_{i,j}$ which vanishes on the projection $\pi_2(W_X)$ is called an *inertia form* (see van der Waerden, 1950). The inertia forms are homogeneous polynomials in each subset $(c_{i,j})_{j=0,\ldots,k_i}$ of parameters.

Affine algebraic sets, defined by polynomial equations, may project onto sets defined by equalities and inequalities (e.g. the hyperbola defined by $xy - 1 = 0$ projects onto the x-axis except 0). But projective varieties project onto projective varieties, defined by (homogeneous) polynomial equations (see Shafarevitch, 1974; Harris, 1992). Their projection is closed for the Zarisky topology. That is the reason why we assume here that *X is a subvariety of a projective space.* Moreover, we will assume that *X is irreducible*, for we can reduce our study to this case via $W_{X_1 \cup X_2} = W_{X_1} \cup W_{X_2}$. Note that the variety W_X, defined by multihomogeneous equations, is also a projective variety. Therefore, its projection by π_1 is a closed subvariety of $\mathbb{P}^{k_0} \times \cdots \times \mathbb{P}^{k_n}$.

Definition 2.1. Let $Z = \pi_1(W_X)$. If Z is an hypersurface, then its equation (unique up to a scalar) will be called the *resultant* of $f_1, \ldots, f_{n+1}$ on X. It will be denoted by $\mathrm{Res}_X(f_1, \ldots, f_{n+1})$.

In other words, when $\mathrm{codim}(Z) = 1$, Res_X is defined up to scaling, and $f_1(\mathbf{x}) = \cdots = f_{n+1}(\mathbf{x}) = 0$ have a solution in X iff $\mathrm{Res}_X(f_1, \ldots, f_{n+1}) = 0$. This generalizes the definition of the classical resultant (over $\mathbb{P}^n$) to any irreducible projective variety.

In order to arrive at the case $\mathrm{codim}(Z) = 1$, we impose the following conditions:

Condition 2.2.

(1) For any point $\mathbf{x} \in X$, and any $i = 1, \ldots, n+1$, the vector $\mathcal{L}_i(\mathbf{x})$ is not zero.
(2) For a generic value of the parameters $\mathbf{c}$, the system $\mathbf{f_c}(\mathbf{x}) = \mathbf{0}$ has no solution in X.

The first condition is needed to derive easily the properties on W_X. The second condition is there to avoid degenerate cases, such as $\mathcal{L}_i = \mathcal{L}_1$ for $i \neq 1$.

Proposition 2.3. *Under conditions 2.2, the projection Z is of codimension* 1 *and the* resultant Res_X *of $f_1, \ldots, f_{n+1}$ on X is uniquely defined, up to a scalar. It is an irreducible polynomial of $\mathbb{Z}[c_{i,j}]$.*

Proof. As for any $\mathbf{x} \in X$ and any $i = 1, \ldots, n+1$, the vector $\mathcal{L}_i(\mathbf{x})$ is non-zero, the set $\pi_2^{-1}(\mathbf{x})$ is a linear space of $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_{n+1}} \times \{\mathbf{x}\}$ of dimension $\sum_{i=1}^{n+1} k_i - n - 1$. Consequently, by the fiber theorem (see Shafarevitch, 1974, pp. 60, 61; Harris, 1992, p. 139), we deduce that W_X is irreducible and of dimension $\sum_{i=0}^{n} k_i - 1$.

Thus, its projection $Z = \pi_1(W_X)$ is an irreducible variety of dimension $\leq \sum_{i=0}^{n} k_i - 1$, or equivalently of codimension ≥ 1.

Let U be the set of parameters $\mathbf{c} \in \mathbb{P}^{k_0} \times \cdots \times \mathbb{P}^{k_n}$ such that $\mathbf{f_c}(\mathbf{x}) = 0$ has no solution on X. Let U' be the set of parameters $\mathbf{c}$ such that the system $f_2 = \cdots = f_{n+1} = 0$

has a finite number of solutions (in X). Note that $U \subset U'$, for if the solution set of $f_2 = \cdots = f_{n+1} = 0$ is of dimension ≥ 1, then the solution set of $\mathbf{f}(\mathbf{x}) = 0$ is of dimension ≥ 0. By Condition 2.2 (2), U and therefore U' are dense subsets of $\mathbb{P}^{k_0} \times \cdots \times \mathbb{P}^{k_n}$. Then $W_X \cap (U' \times X)$ is a dense subset of W_X and projects by π_1 onto $Z \cap U'$. As $\mathcal{Z}(f_2 = \cdots = f_{n+1} = 0)$ is finite, for any $\mathbf{c} \in Z \cap U'$, $\pi_1^{-1}(\mathbf{c}) = \{(\mathbf{c}, \zeta)\ ; \zeta \in \mathcal{Z}_X(f_1 = \cdots = f_n = 0) \cap \mathcal{Z}_X(f_0 = 0)\}$ is finite. Therefore, W_X and Z are of the same dimension and Z is a hypersurface of $\mathbb{P}^{k_0} \times \cdots \times \mathbb{P}^{k_n}$, defined by a unique equation $\mathrm{Res}_X(f_0, \ldots, f_n)$ (up to a scalar). Since Z is irreducible, this polynomial is irreducible, and since the equations defining W_X are in $\mathbb{Z}[c_{i,j}, \mathbf{x}]$, $\mathrm{Res}_X \in \mathbb{Z}[c_{i,j}]$. □

In the language of modern algebraic geometry, the vectors $\mathcal{L}_i(\mathbf{x})$ define line bundles on X. If the line bundles are ample, then the conditions 2.2 are satisfied (see Gelfand *et al.*, 1994).

2.3. RESULTANT OVER $\mathbb{P}^n$

To illustrate the previous developments, we consider here the classical case $X = \mathbb{P}^n$. The polynomials f_i are generic homogeneous polynomials of degree d_i:

$$f_i = \sum_{a_0 + \cdots + a_n = d_i} c_{i,\mathbf{a}} x_0^{a_0} \cdots x_n^{a_n}, \qquad i = 1, \ldots, n+1.$$

Thus, the vector $\mathcal{L}_i$ is up to a permutation $\mathcal{L}_i = (x_0^{d_i}, x_0^{d_i - 1} x_1, \ldots, x_n^{d_i})$; that is, the vector of all monomials of degree d_i in the variables $x_0, \ldots, x_n$. We easily check that for generic values of $\mathbf{c} = (c_{i,\mathbf{a}})$ the system has no solution in $\mathbb{P}^n$ and that $\mathcal{L}_i(\mathbf{x}) = 0$ iff $\mathbf{x} \equiv 0$. Therefore, according to Proposition 2.3, the resultant $\mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_{n+1})$ is well defined (up to a scalar) and vanishes iff the polynomials $f_1, \ldots, f_{n+1}$ have a common root in $\mathbb{P}^n$.

We list some fundamental properties of the classical resultant, discussed in the aforementioned references.

- The resultant is an irreducible polynomial in the variables $\mathbf{c} = (c_{i,\mathbf{a}})$, with integer coefficients (Proposition 2.3)..
- The classical resultant is invariant under linear transformations of the variables (invariance of the elimination problem by a change of coordinates).
- If f_{n+1} is written as a polynomial product $f'_{n+1} f''_{n+1}$, then the resultant is also factored into the corresponding product $\mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_n, f_{n+1}) = \mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_n, f'_{n+1}) \mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_n, f''_{n+1})$. Note that this does not contradict irreducibility, since the coefficients of f_{n+1} are no longer free parameters. They are sums of products of the coefficients of f'_{n+1} and f''_{n+1}.
- The latter property generalizes to the case of a system comprised of polynomials in an ideal, yielding a divisibility property:

  $$f_1, \ldots, f_{n+1} \in (b_1, \ldots, b_{n+1}) \Rightarrow \mathrm{Res}_{\mathbb{P}^n}(b_1, \ldots, b_{n+1}) | \mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_{n+1}).$$

  The seeming paradox is explained again by the non-genericity of the coefficients.

2.4. RESULTANT OVER A TORIC VARIETY

We move now to the case of a toric variety X. As we will see, this variety X is not given explicitly as in the previous section, but defined implicitly by the monomials that

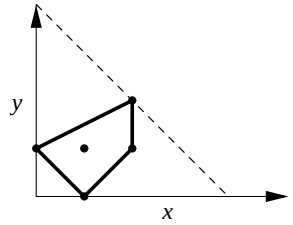

Figure 1. The support and Newton polytope of polynomial $c_1y + c_2x^2y^2 + c_3x^2y + c_4x + c_5xy$. The dotted triangle is the Newton polytope of the completely dense polynomial of the same total degree.

actually appear in the polynomials f_i. The possibility to tune the resultant construction to the actual set of monomials in the equations is one of the strengths of the theory, but it also makes it more difficult to understand. We will sketch here the main results and refer to the work of Gelfand *et al.* (1994) for a complete treatment.

TORIC VARIETY

A toric variety can be defined as the closure of the image of the torus $(\mathbb{K}^*)^n$ by a *monomial* parameterization, in a projective space $\mathbb{P}^N$:

$$\sigma : \mathbf{t} = (t_1, \ldots, t_n) \in (\overline{\mathbb{K}}^*)^n \mapsto (\mathbf{t}^{\mathbf{a}_0} : \cdots : \mathbf{t}^{\mathbf{a}_N}) \in \mathbb{P}^N.$$

See Fulton (1993) or Cox (1995) for a more intrinsic construction. By definition, the image of the torus $(\mathbb{K}^*)^n$, denoted by X^0, is dense in this variety.

POLYTOPES

Elimination theory on a toric variety (also called sparse elimination theory[†]) considers *Laurent* polynomials in n variables, where the exponents are allowed to be arbitrary integers. The polynomial ring is $\mathbb{K}[x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}] = \mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]$, for some base field $\mathbb{K}$.

DEFINITION 2.4. Let f be a polynomial in $K[\mathbf{x}, \mathbf{x}^{-1}]$. The finite set $A \subset \mathbb{Z}^n$ of all monomial exponents corresponding to non-zero coefficients is the *support* of f. The *Newton polytope* of f is the convex hull of A, denoted $Q = \mathrm{conv}(A) \subset \mathbb{R}^n$.

If we use $\mathbf{x}^e$ to denote the monomial $x_1^{e_1} \ldots x_n^{e_n}$, where $e = (e_1, \ldots, e_n) \in \mathbb{Z}^n$ is an exponent vector, then

$$f = \sum_{a_j \in A} c_j \mathbf{x}^{a_j}, \qquad \forall c_j \neq 0.$$

Figure 1 depicts the support and Newton polytope for a bivariate polynomial and compares it with the Newton polytope of the completely *dense* polynomial with the same total degree, i.e. a polynomial in which every coefficient is non-zero. Newton polytopes provide a bridge from algebra to geometry since they permit certain algebraic problems to be cast in geometric terms. For background information on polytope theory and any

[†]The word "sparse" is a misnomer in this context, for it refers to polynomials which could be dense with respect to their polytopes.

unproved properties of mixed volumes the reader may refer to Grünbaum (1967) and Schneider (1993).

MIXED VOLUME

The Minkowski sum $A + B$ of convex polytopes A and B in $\mathbb{R}^n$ is the set $A + B = \{a + b | a \in A, b \in B\} \subset \mathbb{R}^n$ (also denoted by $A \oplus B$). $A + B$ is a convex polytope.

DEFINITION 2.5. Given convex polytopes $Q_1, \ldots, Q_n \subset \mathbb{R}^n$, there is a unique, up to multiplication by a scalar, real-valued function $\mathrm{MV}(Q_1, \ldots, Q_n)$, called the *mixed volume* of the given polytopes, which is multilinear with respect to Minkowski addition and scalar multiplication, i.e. for $\mu, \rho \in \mathbb{R}_{\geq 0}$ and convex polytope $Q'_k \subset \mathbb{R}^n$

$$\begin{aligned}\mathrm{MV}(Q_1, \ldots, \mu Q_k + \rho Q'_k, \ldots, Q_n) = \mu \mathrm{MV}(Q_1, \ldots, Q_k, \ldots, Q_n) \\ + \rho \mathrm{MV}(Q_1, \ldots, Q'_k, \ldots, Q_n).\end{aligned}$$

To define mixed volume completely we require that

$$\mathrm{MV}(Q_1, \ldots, Q_1) = n! V(Q_1),$$

where $V(\cdot)$ is the Euclidean n-dimensional volume function that assigns the unit volume to the hypercube of unit edge length.

Mixed volume generalizes the standard volume function on a single polytope. Indeed, it is the multilinear function associated with the volume function. It has been extensively studied by combinatorial geometers, though its definition sometimes differs by a factor of $n!$.

Let us first give the intuition which is behind the computation of mixed volume, before going into details. The mixed volume of polytopes $Q_1, \ldots, Q_n$ is, in some sense, the multilinear part in the volume of the sum $Q = Q_1 + Q_2 + \cdots + Q_n$. Thus, the idea is to subdivide the polytope Q into a union of polytopes which are sums of faces of the polytopes Q_i. We will keep those polytopes which contribute "multilinearly to the volume"; that is, those which are sums of edges of the polytopes Q_i, and compute the sum of their volumes.

Note that the operation of Minkowski addition on n polytopes is a many-to-one function from $(\mathbb{R}^n)^n$ onto $\mathbb{R}^n$, mapping an n-tuple of polytopes Q_i into their Minkowski sum by sending an n-tuple of points $p_i \in Q_i$ into their vector sum:

$$\begin{aligned}Q_1 \times \cdots \times Q_n \to Q = \sum_{i=1}^{n} Q_i \\ (a_1, \ldots, a_n) \mapsto \sum_{i=1}^{n} a_i.\end{aligned}$$

To subdivide the polytope Q, we define a *section* of this map, i.e. a unique tuple for every point in Q, using a standard *lifting* method. Below we describe an efficient version of this technique, introduced in Billera and Sturmfels (1992), and which is used in several algorithms in toric elimination: select n generic linear lifting forms $l_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, n$. Then define the lifted polytopes

$$\widehat{Q}_i = \{(p_i, l_i(p_i)) : p_i \in Q_i\} \subset \mathbb{R}^{n+1}, \qquad i = 1, \ldots, n.$$

Their Minkowski sum is an $(n+1)$-dimensional polyhedral complex whereas its lower envelope is an n-dimensional polyhedral complex defined as the union of all n-dimensional faces, or facets (whose inner normal vector has positive last component). The genericity of the l_i ensures that the lower envelope projects bijectively onto the Minkowski sum $\sum_{i=1}^n Q_i$ of the original polytopes. Moreover, it ensures that every lower envelope facet is a unique sum of faces $\widehat{F}_i$ from the $\widehat{Q}_i$ for $i = 1, \ldots, n$ such that $\sum_{i=1}^n \dim \widehat{F}_i = n$.

The subdivision of the lower envelope into facets induces a subdivision of $\sum_{i=1}^n Q_i$ into cells by projecting each facet onto an n-dimensional cell. These are maximal cells, whereas cells of dimension $k < n$ are defined as the projection of lower envelope faces of dimension k. This cell complex is a *mixed subdivision*. Every maximal cell is a unique Minkowski sum of faces $F_i \subset Q_i$, where each F_i corresponds to $\widehat{F}_i$ that appears in the unique sum defining the corresponding lower envelope facet. This Minkowski sum is said to be optimal since it minimizes the value of the aggregate lifting function over all possible n-tuples of faces whose Minkowski sum equals the given cell. Due to the linearity of the lifting functions l_i, $\dim F_i = \dim \widehat{F}_i$ therefore $\sum_{i=1}^n \dim F_i = n$. We define the *mixed cells* to be precisely those where all summand faces are one-dimensional, i.e. all F_i are edges. We are now ready to state a property (or equivalent definition) of mixed volume (Billera and Sturmfels, 1992).

PROPOSITION 2.6. *The sum of the n-dimensional Euclidean volumes of all mixed cells in a mixed subdivision of $\sum_{i=1}^n Q_i$ is the* mixed volume $\mathrm{MV}(Q_1, \ldots, Q_n)$ *of the given polytopes $Q_1, \ldots, Q_n \subset \mathbb{R}^n$.*

The shorthands $\mathrm{MV}(f_1, \ldots, f_n)$ and $\mathrm{MV}(A_1, \ldots, A_n)$ are occasionally used for the mixed volume $\mathrm{MV}(Q_1, \ldots, Q_n)$.

BKK BOUND

The Newton polytopes offer a convenient model for the "sparseness" of a polynomial system in light of Bernshtein's upper bound on the number of common roots (Bernstein, 1975). This bound is also known as the BKK bound to underline the contributions of Kushnirenko (1976) and Khovanskii (1978) in its development and proof.

THEOREM 2.7. *Let $f_1, \ldots, f_n \in \mathbb{K}[x_1, x^{-1}, \ldots, x_n, x_n^{-1}]$ with Newton polytopes $A_1, \ldots, A_n$. The number of isolated common zeros in $(\overline{\mathbb{K}}^*)^n$, multiplicities counted, is either infinite, or does not exceed* $\mathrm{MV}(A_1, \ldots, A_n)$*, where $\overline{\mathbb{K}}$ is the algebraic closure of $\mathbb{K}$. For almost all specializations of the coefficients the number of common zeros is exactly* $\mathrm{MV}(A_1, \ldots, A_n)$.

Indeed this result is not so surprising, for it is easy to see that mixed volume behaves like the generic number of roots, if we replace sum of polytopes by product of polynomials and integer multiplication of polytopes by exponentiation of polynomials.

Interesting extensions to this theorem concern the weakening of the genericity condition (Canny and Rojas, 1991; Rojas, 1994), arbitrary affine roots, and the case of arbitrary fields, including fields of positive characteristic (Danilov, 1978; Huber and Sturmfels, 1997; Rojas, 1999).

The mixed volume is typically significantly lower than Bézout's bound, which bounds the number of projective solutions by $\prod_i \deg f_i$, where $\deg f_i$ is the total degree of f_i.

One example is the simple and generalized eigenproblems on $n \times n$ matrices. Both can be expressed by systems of $n+1$ polynomials of total degree two. Hence, the Bézout bound in both cases is 2^{n+1}, while the number of solutions is $2n$ because to each eigenvalue correspond two right eigenvectors of opposite sign. A mixed volume computation yields precisely $2n$ (see, e.g. Li *et al.*, 1996).

The two bounds coincide for completely dense polynomials, because each Newton polytope is an n-dimensional unit simplex scaled by $\deg f_i$. By definition, the mixed volume of the dense system is

$$\mathrm{MV}(\deg f_1 S, \ldots, \deg f_n S) = \prod_{i=1}^{n} \deg f_i \mathrm{MV}(S, \ldots, S) = \prod_{i=1}^{n} \deg f_i,$$

where S is the unit simplex in $\mathbb{R}^n$ with vertex set $\{(0,\ldots,0),(1,0,\ldots,0),\ldots,(0,\ldots,0,1)\}$.

Several efficient algorithms exist for computing mixed volume (Verschelde *et al.*, 1994; Emiris and Canny, 1995; Verschelde *et al.*, 1996; Li *et al.*, 1996; Emiris and Verschelde, 1999). The main idea of all algorithms is to use a lifting in order to apply Definition 2.6. In terms of complexity classes, mixed volume is #P-complete (Pedersen, 1994).

A mixed subdivision provides not only the mixed volume, but also a monomial basis for the coordinate ring associated to the ideal of the given polynomials. This is explained in Section 4.1. The same computation also specifies the start system of a homotopy continuation for numerically approximating all common roots (Huber and Sturmfels, 1995; Verschelde *et al.*, 1996). These homotopies are called sparse because the number of paths followed depends on the monomial structure of the system, in particular, on its mixed volume.

Clearly, mixed volume captures the inherent complexity of algebraic problems in the context of toric elimination and thus provides lower bounds on the complexity of algorithms. In dealing with mixed volumes, some fundamental results can be found in Burago and Zalgaller (1988) and Schneider (1993). In particular, the Aleksandrov–Fenchel inequality leads to the following bound (Emiris, 1996): $\mathrm{MV}(A_1,\ldots,A_n) \geq (n!)(V(A_1)\cdots V(A_n))^{1/n}$. On the other hand, it shall become clear that several toric elimination algorithms rely on some Minkowski sum of Newton polytopes. It turns out that a crucial question in deriving output-sensitive upper bounds is the relation between mixed volume and the volume of these Minkowski sums.

For a set of n or $n+1$ Newton polytopes A_i, define its *scaling factor* s to be the minimum real value so that $A_i + t_i \subset sA_\mu$ for all A_i, where A_μ is the polytope of minimum Euclidean volume and the $t_i \in \mathbb{R}^n$ are arbitrary translation vectors. Clearly, $s \geq 1$ and s is finite iff all polytopes have an affine span of the same dimension. Let e denote the basis of natural logarithms, and suppose that $V(A_i) > 0$ for all i. Then Emiris (1996) proved

$$V\left(\sum_{i=1}^{n} A_i\right) = O(e^n s^n)\mathrm{MV}(A_1,\ldots,A_n), \qquad V\left(\sum_{i=1}^{n+1} A_i\right) = O\left(\frac{e^n s^n}{n}\right)\sum_{i=1}^{n+1} \mathrm{MV}_{-i},$$

where MV_{-i} stands for the mixed volume $\mathrm{MV}(A_1,\ldots,A_{i-1},A_{i+1},\ldots,A_{n+1})$, $i = 1,\ldots,n+1$.

THE TORIC RESULTANT

For a system of $n+1$ (Laurent) polynomials in n variables, the toric resultant characterizes the existence of non-trivial common zeros in a toric variety X.

We associate to a polytope Q, the toric variety parameterized by all the monomials of this polytope. We denote this variety by $\mathcal{T}_Q$. In the context of toric elimination theory, we will consider the toric variety $X = \mathcal{T}_Q$, where $Q = A_1 + \cdots + A_{n+1}$ and A_i is the Newton polytope of f_i, $i = 1, \ldots, n+1$. If the supports A_i are equal, then we can instead consider $\mathcal{T}_{A_1}$.

Let $\mathbf{c} = (c_{i,j})$ be the vector of all polynomial coefficients, regarded as indeterminates, and let Z^0 be the set of coefficients $\mathbf{c}$ such that the system $\mathbf{f_c}(\mathbf{x}) = \mathbf{0}$ has a root in $X^0 = \mathrm{im}(\sigma)$, or equivalently such that there exists $\mathbf{t} \in (\overline{\mathbb{K}}^*)^n$ with $\mathbf{f_c}(\sigma(\mathbf{t})) = \mathbf{0}$. Let Z be the Zariski closure of Z^0, which can also be defined as the projection $\pi_2(W_X)$ of the incidence variety W_X (see Section 2.2) over the toric variety $X = \mathcal{T}_Q$ (Kapranov *et al.*, 1992; Gelfand *et al.*, 1994; Rojas, 1999).

A technical assumption is that, without loss of generality, the affine lattice generated by $\sum_{i=1}^{n+1} A_i$ is n-dimensional. This lattice is identified with $\mathbb{Z}^n$ possibly after a change of variables, which can be implemented by computing the appropriate Smith's Normal form (Sturmfels, 1994). Then we have the following theorem (see also Gelfand *et al.*, 1994 and Theorem 2.3):

PROPOSITION 2.8. (PEDERSEN AND STURMFELS, 1993) *Assume that the affine lattice generated by* $\sum_{i=1}^{n+1} A_i$ *is* n*-dimensional. Then the* toric resultant $\mathrm{Res}_X(f_1, \ldots, f_{n+1})$ *of polynomials* $f_i \in \mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]$ *with supports* A_i *for* $i = 1, \ldots, n+1$*, is well defined (up to a scalar). It is an irreducible polynomial over their coefficients* $\mathbf{c}$*, itself with integer coefficients. Furthermore, the degree of* $\mathrm{Res}_X(f_1, \ldots, f_{n+1})$ *in the coefficients of polynomial* f_i *equals* MV_{-i}*, for* $i = 1, \ldots, n+1$.

Generic polynomials are specified with respect to their support instead of the total degree of the classical theory. Hence, the resultant will also be defined for a set of supports, assuming all non-zero coefficients are generic. The resultant $\mathrm{Res}_X(f_1, \ldots, f_{n+1})$ will be denoted by $\mathrm{Res}(A_1, \ldots, A_{n+1})$, where A_i is the Newton polytope of f_i. The vanishing of $\mathrm{Res}(A_1, \ldots, A_{n+1})$ is a necessary and sufficient condition for the existence of roots in the toric variety $X = \mathcal{T}_Q$ (Kapranov *et al.*, 1992). See also Cox (1995) and Rojas (1999, Section 4.2).

Some fundamental properties of the toric resultant (as a polynomial in the coefficients $\mathbf{c}$) are as follows.

- The toric resultant subsumes the classical resultant in the sense that they coincide if the (generic) polynomials are dense (van der Waerden, 1950). Section 3.2 expands on this topic.
- Just as in the classical case, when all coefficients are generic, the resultant is irreducible (for X is necessarily irreducible). This is essentially stated in Proposition 2.8.
- While the classical resultant is invariant under linear transformations of the variables, the toric resultant is invariant under invertible transformations of the variables that preserve the polynomial support (Sturmfels, 1994; Gelfand *et al.*, 1994).

- In the case of non-generic coefficients, analogous divisibility properties hold as in the case of the classic resultant. In particular, when a system of polynomials lies in the ideal generated by another system, then the latter resultant is divisible by the former resultant.

3. Matrix Formulations

The computation of resultants typically relies on obtaining matrices whose determinant is either the exact resultant polynomial or, more generally, a non-trivial multiple of it. In addition, for solving polynomial systems these matrices are sufficient, since they reduce the given non-linear problem to a question in linear algebra. The focus of this survey are methods for constructing such matrices and their properties.

Resultant matrices can be classified into two large families, though the distinction is not always completely clear. The matrices that generalize Sylvester's and Macaulay's formulations shall be the topic of the next section. Algorithms for constructing matrices for the toric (or sparse) resultant, also known as Newton matrices, are discussed in Sections 3.2 and 3.3. The second family of matrices, which generalize Bézout constructions, will be discussed in the following section, whereas a method combining the two approaches will be discussed in Section 3.5. Finally, Section 3.6 compares some of the different formulations.

There is a command for computing the Sylvester matrix on most general computer algebra systems, including AXIOM, MAPLE, MATHEMATICA, and REDUCE. There is also a command for computing the Bezoutian in most of these systems (e.g. MAPLE), but it seems to be less used, probably because of the complex definition of the tool. Their generalization to the multivariate case is implemented in the MAPLE package `multires` and available at `http://www.inria.fr/saga/logiciels/multires.html`.

In the first of the following sections, we will consider what we call *Sylvester-type* matrices. For two univariate polynomials, their resultant equals the determinant of the well-known Sylvester matrix, a very widespread tool for variable elimination; refer to Sylvester (1853) or Knuth (1981); Kapur and Lakshman (1992); and Zippel (1993). The *Sylvester-type* matrices, generalize this construction to multivariate polynomials $f_1, \ldots, f_{n+1}$. As in the univariate case, the matrices that we construct represent monomial multiples of the polynomials f_i.

Let $\langle \mathbf{x}^A \rangle \subset \mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]$ be the set of all Laurent polynomials in n variables with support $A \subset \mathbb{Z}^n$; that is, the vector space generated by the monomials $\mathbf{x}^A = \{\mathbf{x}^a : a \in A\}$. Now fix supports $B_1, \ldots, B_{n+1} \subset \mathbb{Z}^n$ and consider the following linear transformation:

$$\begin{aligned} S : \langle \mathbf{x}^{B_1} \rangle \times \cdots \times \langle \mathbf{x}^{B_{n+1}} \rangle &\to \langle \mathbf{x}^B \rangle \\ (g_1, \ldots, g_{n+1}) &\mapsto g = \sum_{i=1}^{n+1} g_i f_i, \end{aligned}$$

where B is a subset of $\mathbb{Z}^n$ containing the support of all $\mathbf{x}^b f_i$, for $b \in B_i, i = 1, \ldots, n+1$. The resultant matrices that we consider in this first part, are precisely the matrices of such transformations and to define them fully we have to specify supports B_i. In the next sections, we will describe several formulations, specifying these supports.

We fill in the matrix entries as follows. Every column of S is indexed by an element of some B_i, $i = 1, \ldots, n+1$ and every row by an element of B; equivalently, the columns and rows are indexed respectively by the monomials of g_i and the monomials of g. The

coefficient in the row corresponding to $b \in B$ and in the columns corresponding to $b' \in B_j$ is the coefficient of x^b in the polynomial $x^{b'} f_j$. The coefficients of monomials which do not explicitly appear in $x^{b'} f_j$ have a zero entry. In other words, the column corresponding to $b' \in B_j$ is the coefficient vector of $x^{b'} f_j$ in the basis $\mathbf{x}^B$.

Thus, the matrix S can be divided in blocks $S = [S_1, \ldots, S_{n+1}]$, each S_i depending only on the coefficients of polynomial f_i. In the case $n = 1$, we recover the usual Sylvester matrix, whose block S_1 represents the multiples $f_1(x_1), x_1 f_1(x_1), \ldots, x_1^{d_2-1} f_1(x_1)$ and block S_2 the multiples $f_2(x_1), x_1 f_2(x_1), \ldots, x_1^{d_1-1} f_2(x_1)$.

The number of columns equals the sum of the cardinalities of supports B_i while the number of rows equals the cardinality of B. In the sequel we restrict ourselves to matrices S with *at least as many columns as rows*.

Let us describe now an important property of these matrices, for the construction of the resultant on an irreducible variety X:

THEOREM 3.1. *Assume that:*

(1) there exists a dense open subset X^0 of the variety X such that the zero-set in X^0 of the elements of $\mathbf{x}^B$ is empty: $X^0 \cap \mathcal{Z}(\mathbf{x}^B = 0) = \emptyset$,
(2) Conditions 2.2 are satisfied.

Then every minor D of size $|B|$ of matrix S is a multiple of the resultant $\mathrm{Res}_X(f_1, \ldots, f_{n+1})$.

PROOF. We distinguish two cases. Either S is always of rank $< |B|$ (for any value of the parameters $\mathbf{c}$) and any minor of size $|B|$ is zero. Then, the theorem is obviously true.

Or, we may assume that S is generically of rank $|B|$. Let Z^0 be the set of coefficient specializations such that $f_1, \ldots, f_{n+1}$ have a common solution in $X^0 \subset X$. Assume that for a specialization of the system $\mathbf{f_c}$ with coefficients in Z^0, there exists a non-vanishing maximal minor of S. Then S is surjective and any element of $\mathbf{x}^B$ is a polynomial combination of the polynomials $f_1, \ldots, f_{n+1}$. Since these polynomials have a root $\zeta \in X^0$, the elements $\mathbf{x}^B$ vanish at ζ, which contradicts assumption (1).

Therefore, any maximal minor D of S is zero on Z^0, thus it is zero on its (Zariski) closure, which is $Z = \pi_1(W_X)$ for X^0 is a dense subset of X. The set of coefficients $\mathbf{c} \in Z$ is, by definition, the zero set of $\mathrm{Res}_X(f_i, \ldots, f_{n+1})$. Since the latter is irreducible, it divides D in $\mathbb{Z}[c_{i,j}]$ where $(c_{i,j})_{j=0,\ldots,k_i}$ are the coefficients of f_i. □

In the examples that we will consider, Condition (1) will obviously be true. In the projective case ($X = \mathbb{P}^n$), X^0 will be the affine space $\mathbb{A}^n$ and the set of monomial $\mathbf{x}^B$ will contain 1. In the toric case, X^0 will be the "monomial" image of $(\overline{\mathbb{K}}^*)^n$ in X. As the monomials $\mathbf{x}^B$ do not vanish on $(\overline{\mathbb{K}}^*)^n$, Condition (1) is satisfied.

REMARK 3.2. This property is interesting only in the case where the minors of size $|B|$ are not all identically zero. Assume this, for a moment, and let $\deg_{f_i}(D)$ denote the degree in the coefficients of polynomial f_i of a non-zero maximal minor D of S. Then, $\deg_{f_i}(D) > \deg_{f_i}(\mathrm{Res}_X(\mathbf{f_c}))$, which is also the number of generic roots of $\mathcal{Z}(f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_n)$. If, moreover, $|B_1| = \deg(\mathrm{Res}_X(\mathbf{f_c}))$, we have $\deg_{f_1}(D) = \deg(\mathrm{Res}_X(\mathbf{f_c}))$ and by cyclic permutation of the polynomials and a gcd computation, we can recover the resultant (see Macaulay, 1902; van der Waerden, 1950; Canny and Emiris, 1993).

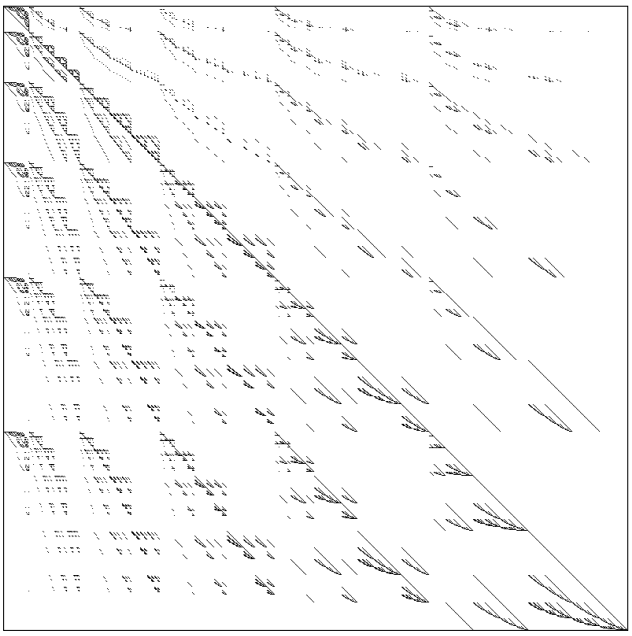

Figure 2. Structure of the Macaulay matrix of six quadrics of $\mathbb{P}^5$.

An important property of this type of matrix from a computational point of view is its structure in the sense of Toeplitz and Hankel matrices (see Bini and Pan, 1994). For general resultant matrices, this kind of structure was established by defining quasi-Toeplitz and quasi-Hankel matrices in Mourrain and Pan (1997a,b, 1999); see also Canny *et al.* (1989); Emiris and Pan (1999). In particular, Macaulay and toric resultant matrices exhibit quasi-linear complexity for vector multiplication. It has been exploited in order to reduce by an order of magnitude the complexity of solving some polynomial systems (Mourrain and Pan, 1998). These matrices are also usually very sparse (and quasi-Toeplitz as illustrated in Figure 2) and this feature has been exploited in the algorithm proposed in Bondyfalat *et al.* (1998), for selecting the root(s) which maximize or minimize a given criterion.

Many questions in this direction need further investigation in order to understand deeply the structure of these matrices. For instance, solving a Toeplitz system can be done in the univariate case in almost linear time, whereas for their generalization to the multivariate case, only quasi-quadratic algorithms are known. Moreover, can we exploit both quasi-Toeplitz structure and sparsity?

3.1. Macaulay matrices

Macaulay's construction (1902) of the resultant of $n+1$ polynomials $f_1, \ldots, f_{n+1}$ of degree $d_1, \ldots, d_{n+1}$ in the variables $\mathbf{x} = (x_1, \ldots, x_n)$ proceeds as follows (we give the non-homogeneous version). Let $\nu = \sum_{i=1}^{n+1} d_i - n$ and let $\mathbf{x}^B$ be the set of all monomials in $\mathbf{x}$ of degree $\leq \nu$. Let $x_n^{d_{n+1}}\mathbf{x}^{B_{n+1}}$ be the set of all monomials of $\mathbf{x}^B$ which are divisible by $x_n^{d_{n+1}}$. Among the remaining monomials in $\mathbf{x}^B - x_n^{d_{n+1}}\mathbf{x}^{B_{n+1}}$, let us denote by $x_{n-1}^{d_n}\mathbf{x}^{B_n}$ those which are divisible by $x_{n-1}^{d_n}$. Similarly, for $i = n+1, \ldots, 2$, we define by induction

$x_{i-1}^{d_i}\mathbf{x}^{B_i}$ to be the set of monomials of $\mathbf{x}^B - x_n^{d_{n+1}}\mathbf{x}^{B_{n+1}} - \cdots - x_i^{d_{i+1}}\mathbf{x}^{B_{i+1}}$ which are divisible by $x_{i-1}^{d_i}$. The set $\mathbf{x}^B - x_n^{d_n}\mathbf{x}^{B_{n+1}} - \cdots x_1^{d_2}\mathbf{x}^{B_2}$ is denoted by $\mathbf{x}^{B_1}$ and is equal to

$$\mathbf{x}^{B_1} = \{x_1^{b_1}\cdots x_n^{b_n} \ ; \ 0 \le b_i \le d_{i+1} - 1\}.$$

It has $d_2 \cdots d_{n+1}$ monomials. Consider now the matrix S associated to these subsets B_i. It is a square matrix (for $\mathbf{x}^B = \mathbf{x}^{B_1} \cup x_1^{d_2}\mathbf{x}^{B_2} \cdots x_n^{d_{n+1}}\mathbf{x}^{B_{n+1}}$) of size equal to the number of monomials of degree $\le \nu$; that is, $\begin{pmatrix} \nu + n \\ n \end{pmatrix}$.

In order to prove that its determinant is a non-trivial multiple of the resultant $\mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_{n+1})$, we first prove that the conditions of Theorem 3.1 are true, and then that the determinant of S is not zero. Condition (1) is obviously true for $1 \in \mathbf{x}^B$. Condition (2) is true according to Section 2.3. Now, the determinant of S as a polynomial in $\mathbf{c} = (c_{i,j})$ is not 0, for when f_i is specialized to $x_{i-1}^{d_i}$ for $i = 2, \ldots, n+1$, and f_1 to 1, we obtain the identity matrix. Thus, according to Theorem 3.1, the determinant of this matrix is a non-trivial multiple of the resultant $\mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_{n+1})$.

This determinant is homogeneous of degree $d_2 \cdots d_{n+1}$ in the coefficients of f_1; that is, exactly the degree of the resultant in f_1. Therefore, the resultant of these polynomials can be computed as the gcd of determinants of such matrices, obtained by a cyclic permutation of the polynomials $f_1, \ldots, f_{n+1}$ (see Remark 3.2).

In fact, Macaulay proved (1902) that the resultant is the ratio of $\det(S)$ by a determinant of a submatrix of S.

Example 3.3. For the system

$$\begin{aligned} f_1 &= c_{1,1} + c_{1,2}x_1 + c_{1,3}x_2 + c_{1,4}x_1x_2 \\ f_2 &= c_{2,1} + c_{2,2}x_1 + c_{2,3}x_2 + c_{2,4}{x_1}^2 \\ f_3 &= c_{3,1} + c_{3,2}x_1 + c_{3,3}x_2 + c_{3,4}{x_2}^2, \end{aligned}$$

the *transposed* 15×15 Macaulay matrix has the form:

$$S^t = \left[\begin{array}{ccccccccccccccc}
c_{1,1} & c_{1,3} & c_{1,2} & c_{1,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & c_{1,1} & 0 & c_{1,2} & 0 & c_{1,3} & 0 & c_{1,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & c_{1,1} & c_{1,3} & c_{1,2} & 0 & c_{1,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & c_{1,1} & 0 & 0 & c_{1,2} & c_{1,3} & 0 & c_{1,4} & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & c_{2,1} & 0 & c_{2,3} & 0 & 0 & 0 & 0 & c_{2,2} & 0 & c_{2,4} & 0 \\
0 & 0 & c_{2,1} & c_{2,3} & c_{2,2} & 0 & 0 & 0 & 0 & 0 & 0 & c_{2,4} & 0 & 0 & 0 \\
0 & 0 & 0 & c_{2,1} & 0 & 0 & c_{2,2} & c_{2,3} & c_{2,4} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & c_{2,1} & 0 & c_{2,2} & 0 & c_{2,3} & c_{2,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
c_{2,1} & c_{2,3} & c_{2,2} & 0 & c_{2,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & c_{3,1} & 0 & c_{3,2} & 0 & 0 & 0 & 0 & c_{3,3} & 0 & c_{3,4} \\
0 & c_{3,1} & 0 & c_{3,2} & 0 & c_{3,3} & 0 & 0 & 0 & 0 & 0 & 0 & c_{3,4} & 0 & 0 \\
0 & 0 & 0 & c_{3,1} & 0 & 0 & c_{3,2} & c_{3,3} & 0 & 0 & c_{3,4} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & c_{3,1} & 0 & c_{3,3} & 0 & 0 & c_{3,4} & 0 & c_{3,2} & 0 & 0 & 0 \\
0 & 0 & c_{3,1} & c_{3,3} & c_{3,2} & 0 & 0 & c_{3,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
c_{3,1} & c_{3,3} & c_{3,2} & 0 & 0 & c_{3,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right].$$

The ith block in the matrix corresponds to polynomials f_i. The size of these blocks are four for f_1, five for f_2 and six for f_3. We have $\nu = 4$, $B_1 = \{1, x_1, x_2, x_1x_2\}$, $B_2 = \{1, x_1, x_2, x_1x_2, x_1^2\}$, $B_3 = \{1, x_1, x_2, x_1x_2, x_1^2, x_2^2\}$.

Here is a table giving the Bézout number d^n and the size of the Macaulay matrices for a linear form and n polynomials of degree d, in n variables.

$n\backslash d$	2	3	4	5	6	7	8
2	4 10	9 21	16 36	25 55	36 78	49 105	64 136
3	8 35	27 120	64 286	125 560	216 969	343 1540	512 2300
4	16 126	81 715	256 2380	625 5985	1296 12650	2401 23751	4096 40920
5	32 462	243 4368	1024 20349	3125 65780	7776 169911	16807 376992	32768 749398
6	64 1716	729 27132	4096 177100	15625 736281	46656 2324784	117649 6096454	262144 13983816
7	128 6435	2187 170544	16384 1560780	78125 8347680	279936 32224114	823543 99884400	2097152 264385836
8	256 24310	6561 1081575	65536 13884156	390625 95548245	1679616 450978066	5764801 1652411475	16777216 5047381560
9	512 92378	19683 6906900	262144 124403620	1953125 1101716330	10077696 6358402050	40353607 27540584512	134217728 97082021465
10	1024 352716	59049 44352165	1048576 1121099408	9765625 12777711870	60466176 90177170226	282475249 461738052776	1073741824 1878392407320

These figures show the limit of the size of problems that can effectively be treated by such methods. The next sections are devoted to methods which usually lead to smaller matrices.

3.2. Newton matrices

The construction of Newton matrices is similar to the construction of Macaulay matrices, except that it uses more intricate geometry on the monomials in the f_i. We briefly sketch it, before going into details.

Let us fix $n+1$ Laurent polynomials $f_1, \ldots, f_{n+1} \in \mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]$, with support respectively in the polytopes $A_1, \ldots, A_{n+1}$. Here is a short description of this algorithm:

construction of the Newton matrices

(1) Subdivide the monomials of the Minkowski sum $Q = A_1 + \cdots A_{n+1}$ into cells (mixed subdivision).
(2) Decompose each of these cells as a sum $\mathbf{a}_{i_0} + B_{i_0}$ of a vertex of some A_{i_0} and faces of the other polytopes A_i, $i \neq i_0$. This gives a partition of $\mathbf{x}^B$ as an union of sets $\mathbf{x}^{\mathbf{a}_{i_0}}\mathbf{x}^B$ (to be compared with the partition of $\mathbf{x}^B$ into the union of the sets $x_i^{d_i}\mathbf{x}^{B_i}$ in the previous section).
(3) For all these cells, replace the monomial $\mathbf{x}^{\mathbf{a}_{i_0}}$ by the polynomial f_{i_0} and construct the corresponding coefficient matrix of all these polynomials.

Let us describe, now more precisely, the algorithm that guarantees most properties for the resultant matrix. The original version of Canny and Emiris (1993) was subsequently

improved and generalized in Canny and Pedersen (1993) as explained at the end of the section. A further generalization can be found in Sturmfels (1994).

MIXED SUBDIVISION

First, we need to describe how to subdivide the Minkowski sum of all input Newton polytopes $Q = A_1 + \cdots + A_{n+1} \subset \mathbb{R}^n$. The basic construction extends that of mixed subdivision, described in Section 2.4, to an overconstrained system. To apply the lifting technique, select $n+1$ linear lifting forms $l_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, n+1$. Then define the lifted Newton polytopes

$$\widehat{A}_i = \{(p_i, l_i(p_i)) : p_i \in A_i\} \subset \mathbb{R}^{n+1}, \qquad i = 1, \ldots, n+1,$$

and their Minkowski sum

$$\widehat{Q} = \widehat{A}_1 + \cdots + \widehat{A}_{n+1} \subset \mathbb{R}^{n+1}.$$

The lower envelope of $\widehat{Q}$ projects bijectively onto Q by analogy to the case of n polytopes discussed in Section 2.4. The subdivision of the lower envelope into facets induces a subdivision of Q into cells. This is known as a *mixed subdivision* and shall be assumed fixed in the course of the algorithm.

The lifting functions l_i are chosen to be sufficiently generic so that every point $p \in Q$ can be uniquely written as a sum of Newton polytope points

$$p = p_1 + \cdots + p_{n+1} : \qquad p_i \in A_i, i = 1, \ldots, n+1,$$

where the lifted images of the p_i add up to the unique point on the lower envelope of $\widehat{Q}$ which projects to p. The above sum is called optimal because the p_i minimize the aggregate lifting function $\sum_i l_i(p_i)$ over all $(n+1)$-tuples of points in $\mathbb{R}^n$ whose sum equals p. Analogously, each cell σ of the mixed subdivision is uniquely expressed as an optimal sum

$$\sigma = F_1 + \cdots + F_{n+1} \subset \mathbb{R}^n : \qquad F_i \text{ is a face of } A_i, i = 1, \ldots, n+1.$$

The lower envelope facet projecting onto σ is the Minkowski sum of those faces in the $\widehat{A}_i$ corresponding to F_i. Genericity implies that $\dim \sigma = \sum_i \dim F_i$ and, since $\dim \sigma = n$, we deduce that at least one F_i is a vertex. Cells where exactly one summand face is a vertex are called *mixed* and, in particular, i-mixed iff F_i is a vertex. All other summand faces in a mixed cell must be edges. Recall that by Definition 2.6,

$$\mathrm{MV}_{-i} = \mathrm{MV}(A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_{n+1}) = \sum_{i\text{-mixed } \sigma} V(\sigma), \qquad i = 1, \ldots, n+1.$$

The Minkowski sum contains all the required information for the system, as will be made clear below.

PARTITION OF THE MONOMIALS

In order to remove the ambiguity for monomials in the border of two cells (step 2 of the algorithm), we consider the set

$$B = (Q + \delta) \cap \mathbb{Z}^n,$$

where $\delta \in \mathbb{Q}^n$ is a sufficiently small and generic vector. The partition of B is the one induced by the mixed subdivision.

MATRIX CONSTRUCTION

The rows and columns of the toric resultant matrix S will be indexed by B. Every point $\mathbf{b} \in B$ is in an i-mixed cell. Let us choose i maximal. Then $\mathbf{x}^{\mathbf{b}} = \mathbf{x}^{\mathbf{a}_{i_0}} \mathbf{x}^{\mathbf{b}-\mathbf{a}_{i_0}}$, where $\mathbf{a}_{i_0}$ is a vertex of A_{i_0} and $\mathbf{b} - \mathbf{a}_{i_0} \in \sum_{i \neq i_0} A_i$. Note that $f_{i_0} \mathbf{x}^{\mathbf{b}-\mathbf{a}_{i_0}} \in \langle \mathbf{x}^B \rangle$, for $\mathbf{b} - \mathbf{a}_{i_0} + A_i \subset Q + \delta$. By definition, the column of S indexed by $p \in B$ is the coefficient vector of $f_{i_0} \mathbf{x}^{\mathbf{b}-\mathbf{a}_{i_0}}$ in the basis $\mathbf{x}^B$. The exponent $\mathbf{a}_{i_0}$ will be called the column content of p in S. In summary, we have associated to every point in B the product of some polynomial with a monomial, thus defining a matrix column. The above arguments establish the fact that the supports of all these products lie in B, thus yielding the required *closure property* to guarantee that the matrix is well-defined and square. Point set B and the corresponding Minkowski sum Q are the smallest point set and polytope, respectively, where this closure property holds. The role of the mixed subdivision also becomes obvious, namely in order to specify a unique sum of points defining each point in B.

NON-DEGENERACY

Every principal minor of matrix S, including its determinant, is non-zero when the polynomials have generic coefficients (Canny and Emiris, 1993). The proof is based on the following technical lemma, which captures the geometric properties of the construction. Specialize S to a matrix $S(t)$, where t is a new variable, by specializing every coefficient $c_{i,j}$ in S to a power $t^{l_i(\mathbf{a}_{ij})}$, where $\mathbf{a}_{i,j} \in \mathrm{supp}(f_i)$ is the exponent vector of the corresponding monomial in f_i. For every $p \in E$, let $\widehat{p}$ be the point on the lower envelope of $\widehat{Q}$ lying directly above p and let $h(\widehat{p})$ express its $(n+1)$st coordinate. Consider any column of $S(t)$ indexed by $p \in B$, corresponding to the polynomial $\mathbf{x}^{\mathbf{b}-\mathbf{a}_{i_0}} f_{i_0}$. Then define matrix S' by multiplying the column of $S(t)$ indexed by p by $t^{h(p)-l_i(\mathbf{a}_{i,j})}$.

LEMMA 3.4. *Assume the above notation and denote by S'_{pq} the entry of S' with row index p and column index q, for $p, q \in E$. Then, for all non-zero elements S'_{pq} with $p \neq q$, $\deg_t(S'_{pq}) > \deg_t(S'_{qq})$.*

According to Theorem 3.1, the determinant of S is divisible by the toric resultant $\mathrm{Res}(f_1, \ldots, f_{n+1})$. By construction, the degree of $\det(S)$ in f_1 is the number of points $p \in B$ of the 1-mixed cells $\sigma + \delta$, where $\sigma = F_1 + F_2 + \cdots + F_{n+1}$, $\dim(F_1) = 0$ and $\dim(F_2) = \cdots = \dim(F_{n+1}) = 1$. This number is precisely the mixed volume of $A_2, \ldots, A_n$ (see Section 2.4). Thus, the degree of $\det S$ in the coefficients of f_1 equals the degree of the toric resultant in the same coefficients. The determinant degree in the coefficients of f_i for $i = 2, \ldots, n+1$, is greater or equal to the respective degree of R. According to Remark 3.2, we obtain the resultant $\mathrm{Res}_X(\mathbf{f_c})$ by cyclic permutation of the f_i and gcd computation.

EXAMPLE 3.5. The Newton matrix construction is illustrated for a system of three polynomials in two unknowns:

$$
\begin{aligned}
f_1 &= c_{11} + c_{12}xy + c_{13}x^2y + c_{14}x, \\
f_2 &= c_{21}y + c_{22}x^2y^2 + c_{23}x^2y + c_{24}x, \\
f_3 &= c_{31} + c_{32}y + c_{33}xy + c_{34}x.
\end{aligned}
$$

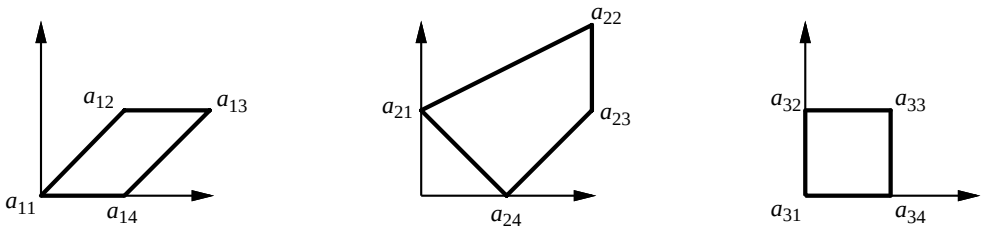

Figure 3. The Newton polytopes and the exponent vectors $\mathbf{a}_{ij}$.

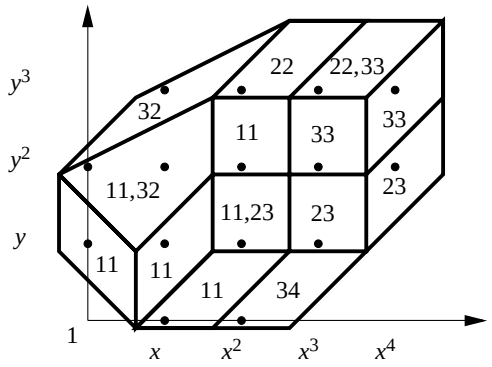

Figure 4. The mixed subdivision Δ_δ of $Q+\delta$, where $\delta = (-3/8, -1/8)$. Each cell is labeled with the indices of the Newton polytope vertices that appear in its optimal sum: ij denotes vertex a_{ij}.

The Newton polytopes are shown in Figure 3. The mixed volumes are $\mathrm{MV}(A_1, A_2) = 4$, $\mathrm{MV}(A_2, A_3) = 4$, $\mathrm{MV}(A_3, A_1) = 3$, so the toric resultant's total degree is 11. Compare this with the Bézout numbers of the 2×2 subsystems: $8, 6, 12$; hence, the classical resultant degree is 26. We pick generic functions $l_1(x, y) = Lx + L^2 y, l_2(x, y) = -L^2 x - y, l_3(x, y) = x - Ly$, where L is a sufficiently large positive integer. We construct the Minkowski sum $\sum_{i=1}^{3} \widehat{A}_i$ in three dimensions; its lower envelope is two-dimensional and projects bijectively onto Minkowski sum $Q = \sum_{i=1}^{3} A_i$. Then we apply a perturbation by vector $\delta = (-3/8, -1/8)$.

The mixed subdivision of $Q+\delta$ into two-dimensional cells and the indices of the Newton polytope vertices in the optimal sums for each cell are shown in Figure 4.

Matrix S, the Newton matrix associated to f_1, appears below with rows and columns indexed by the integer points in B, and has dimension 15. S contains, by construction, the minimum number of f_1 rows, namely four. The total number of rows is $4+4+7 = 15$.

Here is the *transpose* of S:

$$S^t = \begin{array}{c} \\ 1,0 \\ 2,0 \\ 0,1 \\ 1,1 \\ 2,1 \\ 3,1 \\ 0,2 \\ 1,2 \\ 2,2 \\ 3,2 \\ 4,2 \\ 1,3 \\ 2,3 \\ 3,3 \\ 4,3 \end{array} \begin{array}{c} \begin{array}{ccccccccccccccc} 1,0 & 2,0 & 0,1 & 1,1 & 2,1 & 3,1 & 0,2 & 1,2 & 2,2 & 3,2 & 4,2 & 1,3 & 2,3 & 3,3 & 4,3 \end{array} \\ \left[\begin{array}{ccccccccccccccc} c_{11} & c_{14} & 0 & 0 & c_{12} & c_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ c_{31} & c_{34} & 0 & c_{32} & c_{33} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_{11} & c_{14} & 0 & 0 & 0 & c_{12} & c_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{11} & c_{14} & 0 & 0 & 0 & c_{12} & c_{13} & 0 & 0 & 0 & 0 & 0 \\ c_{24} & 0 & c_{21} & 0 & c_{23} & 0 & 0 & 0 & c_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_{24} & 0 & c_{21} & 0 & c_{23} & 0 & 0 & 0 & c_{22} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_{31} & c_{34} & 0 & 0 & c_{32} & c_{33} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{31} & c_{34} & 0 & 0 & c_{32} & c_{33} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_{11} & c_{14} & 0 & 0 & 0 & c_{12} & c_{13} \\ 0 & 0 & 0 & 0 & c_{31} & c_{34} & 0 & 0 & c_{32} & c_{33} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{24} & 0 & 0 & c_{21} & 0 & c_{23} & 0 & 0 & 0 & c_{22} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_{31} & c_{34} & 0 & 0 & c_{32} & c_{33} & 0 & 0 \\ 0 & 0 & 0 & c_{24} & 0 & 0 & c_{21} & 0 & c_{23} & 0 & 0 & 0 & c_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_{31} & c_{34} & 0 & 0 & c_{32} & c_{33} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_{31} & c_{34} & 0 & 0 & c_{32} & c_{33} \end{array}\right] \end{array}$$

Both row and column indexed by $(1,3)$ can be removed, thus factoring c_{32} out of the determinant. This is, in fact, the matrix obtained by a greedy variant of the algorithm proposed by Canny and Pedersen (1993). It constructs matrices whose size is typically smaller than, and can never exceed, that of the original algorithm. Moreover, the greedy algorithm works on arbitrary supports, thus removing the technical requirement set before Definition 2.8 that the integer lattice they generate should be full-dimensional. $\deg_{f_i} \det(S) \geq \deg_{f_i} \mathrm{Res}$, for $i = 2, \ldots, n+1$.

3.3. INCREMENTAL NEWTON MATRICES

In the previous section, we describe sets B_i and B which lead to a square non-degenerate resultant matrix. However, we have no guarantee that its size is minimal. A method proposed in Emiris and Canny (1995) and implemented in Emiris (1997) consists of searching incremental subsets of the B_i and B, which also yield a square non-degenerate matrix. This approach produces matrices whose dimension never exceeds that of the subdivision-based algorithms, and which are typically significantly smaller. The flexibility of the construction makes it suitable for overconstrained systems. On the other hand, this is the reason that certain *a priori* properties of the subdivision-based construction are not guaranteed.

We define

$$Q_{-i} = \sum_{j=1, j\neq i}^{n+1} A_j \subset \mathbb{R}^n \qquad \text{and} \qquad E_i = Q_{-i} \cap \mathbb{Z}^n, \qquad \text{for } i = 1, \ldots, n+1.$$

The algorithm shall restrict B_i to be a subset of E_i, just as in the case of the subdivision-based algorithms. This ensures that $A_i + B_i \subset Q$, for all i. The second concept used is the v-distance of points in $P \cap \mathbb{Z}^n$, where $P \in \mathbb{R}^n$ is any convex polytope and $v \in \mathbb{Q}^n$ is a given vector:

$$v\text{-distance}(p) = \max\{s \in \mathbb{R}_{\geq 0} : p + sv \in P\}.$$

This is the distance of point p from the polytope boundary along direction v.

The main issue is to choose the points of E_i that make up B_i. The construction is incremental, in the sense that successively larger candidate matrices are defined and

tested for validity on whether they express a non-trivial multiple of the toric resultant. Supposing that we are given a direction vector v, we can partition the points in every E_i with respect to their v-distance. At every step, the algorithm adds to B_i all points in E_i whose v-distance exceeds some bound $\beta \in \mathbb{R}_{\geq 0}$, for $i = 1, \ldots, n+1$. Incrementing the sets B_i is equivalent to decreasing β, until a valid matrix is found.

For given sets B_i a rectangular matrix S is well defined, and constitutes a useful candidate only if the number of columns is at least as large as the number of rows. If so, the algorithm tests whether S has full rank for generic coefficients, in other words, whether there exists a maximal minor D which is generically non-zero. The algorithm terminates if S has generically full rank and returns a non-singular maximal square submatrix. This submatrix is a toric resultant matrix, since its determinant D is a non-trivial multiple of R. Observe that the process of deleting the extra columns does not affect the validity of Theorem 3.1.

For D to be a multiple of Res, its degree must be at least MV_{-i} in the coefficients of f_i, for $i = 1, \ldots, n+1$. Hence, the initial sets B_i contain the MV_{-i} points of largest v-distance in E_i, for $i = 1, \ldots, n+1$. The number of points comprising each increment to the B_i values has been studied in Emiris and Pan (1999). Large increments speed up the construction but may miss the smallest possible matrix, so some further tests may be needed once a valid matrix has been found in order to decrease the matrix dimension. It can be shown that for $v = -\delta$, where δ is the perturbation vector in the subdivision-based algorithm, the incremental construction yields a matrix at most as large. Therefore, if at some stage, $B_i = E_i$ for $i = 1, \ldots, n+1$, then this v is rejected. For arbitrary systems, a random vector usually produces a smaller matrix. But there is a class of systems for which a deterministic vector guarantees the construction of an optimal matrix, i.e. a matrix whose determinant equals the resultant or, equivalently, a matrix whose dimension is minimum. This class includes all systems for which an optimal matrix of Sylvester type does exist, as discussed hereafter and in Emiris and Canny (1995).

There are two potential bottlenecks in this construction. First, enumerating all integer lattice points in E_i, for $i = 1, \ldots, n+1$ or, rather, an appropriate and sufficiently large subset of E_i. Certain heuristics are proposed and implemented in Emiris and Canny (1995) based on linear programming. A more complete treatment of the problem may be based on enumerative geometric techniques (Gruber and Wills, 1993; Barvinok, 1993). The second bottleneck is the full-rank test. It can be implemented in an incremental fashion based on an LU decomposition of rectangular matrices (Emiris and Canny, 1995). Extending Canny *et al.* (1989), the results of Emiris and Pan (1999) establish a quasi-linear complexity for vector premultiplication, which reduces to computing the sum of polynomial products, see also, Mourrain and Pan, 1999. Applying standard results this yields quasi-quadratic complexity for calculating the rank and the determinant of a quasi-Toeplitz matrix (Wiedemann, 1986; Bini and Pan, 1994; Golub and Van Loan, 1996). Therefore, the time complexity for matrix construction is $O^*(2^{O(n)} \deg Rt)$, where t is the total number of rank tests during construction and vector v is assumed given. This bound becomes $O^*(2^{O(n)} \deg R)$ when the number of rows in the final matrix is bounded by a constant multiple of $\deg R$, as is often the case in practice.

Compared to the matrix produced by a mixed subdivision, the incremental matrix has the following features.

- There exist deterministic choices for vector v that yield optimal matrices for a

subclass of multihomogeneous systems, as illustrated in Example 3.7 and, in full detail, in Emiris and Canny (1995). This subclass includes linear systems and pairs of arbitrary polynomials, thus, the algorithm of this section generalizes Sylvester's construction.

- For a sufficiently generic vector v, this algorithm subsumes both the original subdivision-based algorithm and its greedy variant, so it produces a matrix at most as large as those algorithms. A brief explanation is provided above and a proof in Emiris and Canny (1995).

- Unlike the previous constructions, where we could establish a closure property like the one just in Section 3.2, the heuristic nature of the incremental algorithm cannot provide the guarantee that every principal submatrix is generically non-singular (Emiris and Canny, 1995). To show an analogous closure property here would constitute a major enhancement to this algorithm.

- If some set B_i is fixed to its optimal size, then we can apply the two alternatives in Canny and Emiris (1993) to recover the actual resultant polynomial.

EXAMPLE 3.6. (CONTINUED FROM SECTION 3.5) Figure 5 shows Q_{-1} in a bold and randomly chosen vector $v = (20, 11)$. The different subsets of E_1 with respect to v-distance are shown by the thin-line polygons inside Q_{-1}. In fact, the thin lines represent contours of fixed v-distance.

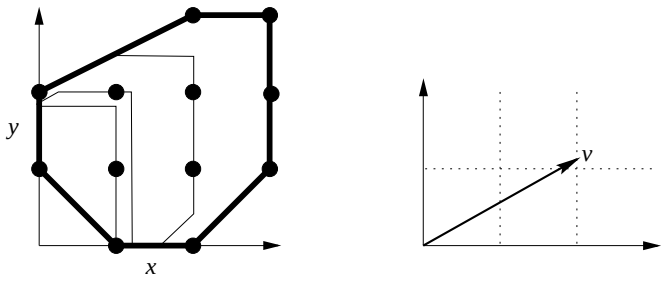

Figure 5. E_1 subsets with different v-distance bounds and vector v.

The final set B_1 includes all integer points in Q_{-1} whose v-distance is larger than or equal to $1/11$; here is B_1 with the v-distances: $\{(0, 1; 3/20), (1, 0; 1/10), (1, 1; 1/10), (1, 2; 1/11)\}$. This v leads to a 13×12 non-singular matrix S shown below with B_i cardinalities $4, 4, 5$. Recall that, in general, the algorithm constructs a rectangular matrix from which it extracts a generically non-singular maximal submatrix. Here, deleting the last row defines the 12×12 resultant submatrix. The first line below displays the integer points indexing the columns. Recall that the subdivision and greedy algorithms give matrices of dimension 15 and 14 respectively, whereas the degrees of the toric and the classic resultant are 11 and 26, respectively. Here is the *transpose* of the constructed

matrix:

$$
\begin{array}{c|cccccccccccc}
 & 1,2 & 2,2 & 0,1 & 1,1 & 2,1 & 3,1 & 1,0 & 2,0 & 3,2 & 2,3 & 3,3 & 0,2 \\
0,1 & c_{12} & c_{13} & c_{11} & c_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1,0 & 0 & 0 & 0 & 0 & c_{12} & c_{13} & c_{11} & c_{14} & 0 & 0 & 0 & 0 \\
1,1 & 0 & c_{12} & 0 & c_{11} & c_{14} & 0 & 0 & 0 & c_{13} & 0 & 0 & 0 \\
1,2 & c_{11} & c_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_{12} & c_{13} & 0 \\
\hline
0,0 & 0 & c_{22} & c_{21} & 0 & c_{23} & 0 & c_{24} & 0 & 0 & 0 & 0 & 0 \\
1,0 & 0 & 0 & 0 & c_{21} & 0 & c_{23} & 0 & c_{24} & c_{22} & 0 & 0 & 0 \\
1,1 & c_{21} & 0 & 0 & 0 & c_{24} & 0 & 0 & 0 & c_{23} & 0 & c_{22} & 0 \\
0,1 & 0 & c_{23} & 0 & c_{24} & 0 & 0 & 0 & 0 & 0 & c_{22} & 0 & c_{21} \\
\hline
0,1 & c_{33} & 0 & c_{31} & c_{34} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_{32} \\
1,1 & c_{32} & c_{33} & 0 & c_{31} & c_{34} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1,0 & 0 & 0 & 0 & c_{32} & c_{33} & 0 & c_{31} & c_{34} & 0 & 0 & 0 & 0 \\
2,1 & 0 & c_{32} & 0 & 0 & c_{31} & c_{34} & 0 & 0 & c_{33} & 0 & 0 & 0 \\
2,2 & 0 & c_{31} & 0 & 0 & 0 & 0 & 0 & 0 & c_{34} & c_{32} & c_{33} & 0
\end{array}
$$

EXAMPLE 3.7. (MULTIHOMOGENEOUS SYSTEMS) We now focus on a special class of multihomogeneous polynomial systems for which Sylvester-type matrices provably exist for the toric resultant. The incremental algorithm produces these matrices and, additionally, finds rather compact matrices for arbitrary multihomogeneous systems.

A homogeneous polynomial is said to be multihomogeneous if the set of variables can be partitioned into r subsets $X_1, \ldots, X_r$, so that the polynomial is homogeneous when considered as a polynomial in each subset. This is sometimes called an r-homogeneous polynomial. Suppose that the number of individual variables in X_k is $l_k + 1$, where one of them is the homogenizing variable, then, in our notation, $n = l_1 + \cdots + l_r$. If the total degree in X_k is d_k, then the polynomial is said to be of type $(l_1, \ldots, l_r; d_1, \ldots, d_r)$. There is a rich theory of multihomogeneous systems (Morgan, 1987; Morgan *et al.*, 1994), namely systems where each polynomial is multihomogeneous and has the same support. If the degree of the ith polynomial in X_j is d_{ij}, then the number of common isolated solutions for the n polynomials is bounded by

$$\text{the coefficient of } \prod_{j=1}^{r} x_j^{l_j} \text{ in polynomial } \prod_{i=1}^{n} \left(\sum_{j=1}^{r} d_{ij} x_j \right).$$

Sturmfels and Zelevinsky (1994) studied, in particular, the subclass of systems for which

$$l_k = 1 \text{ or } d_k = 1, \qquad \text{for } k = 1, \ldots, r.$$

They showed that every such system has a number of Sylvester type matrices for its toric resultant, i.e. matrices whose determinant is precisely the resultant. Furthermore, they conjectured that no other class of systems has optimal matrices of Sylvester type, i.e. where the matrix dimension is minimum and every entry is either zero or an input coefficient. It can be proven that the optimal matrices, whenever they exist, can be constructed by the incremental algorithm for a deterministic choice of v.

THEOREM 3.8. (EMIRIS AND CANNY, 1995) *Consider a multihomogeneous system in the subclass specified above, comprised of $n + 1$ polynomials in n variables. The variable*

Table 1. Incremental matrix construction on multihomogeneous systems.

type	vector $v \in \mathbb{Q}^n$	$\deg R$	$\dim S$
$(2,1;2,1)$	$\sim (1,1,3)$	48	52
$(2,1;2,2)$	$\sim (1,1,5)$	96	104
$(2,1,1;2,1,1)$	$\sim (3,3,2,1)$	240	295
$(2,1,1;2,2,1)$	$\sim (3,3,3,1)$	480	592

set is partitioned in r subsets of l_k variables each, for $k = 1, \ldots, r$, so that all polynomials have total degree d_k in the kth variable subset. Using the above notation, define vector $v \in \mathbb{Q}^n$ as a concatenation of r subvectors of cardinality l_k, $k = 1, \ldots, r$, where all coordinates in the kth subvector are set to $1 - d_k + (d_k/l_k) \sum_{j=k}^{r} l_k$. Then, the incremental algorithm constructs an optimal matrix for the input system.

Experimental results show that for arbitrary multihomogeneous systems, the same procedure for determining v can be applied to yield matrices of satisfactory size (Emiris and Canny, 1995). Table 1 displays some such systems, where the vector is calculated by the procedure above, after a random perturbation. $\deg R$ indicates the total degree of the toric resultant and $\dim S$ is the dimension of the constructed matrix.

3.4. Bézout matrices

In this section, we recall some basic definitions from the theory of Bezoutians referring the reader to Cardinal and Mourrain (1996) and Elkadi and Mourrain (1996) for further details. As mentioned at the beginning, this tool is fundamental in many recent algorithmic developments (Aizenberg and Kytmanov, 1981; Berenstein and Yger, 1991; Berenstein *et al.*, 1993; Cardinal, 1993; Fitchas *et al.*, 1993; Kapur *et al.*, 1994, 1996; Becker *et al.*, 1996; Cardinal and Mourrain, 1996; Kapur and Saxena, 1997; Elkadi and Mourrain, 1998, 1999).

It generalizes the construction of E. Bézout to the multivariate case. In addition to the vector of variables $\mathbf{x}$, consider the vector $\mathbf{y} = (y_1, \ldots, y_n)$ and write $\mathbf{x}^{(0)} = \mathbf{x}$, $\mathbf{x}^{(1)} = (y_1, x_2, \ldots, x_n), \ldots, \mathbf{x}^{(n)} = \mathbf{y}$. For a polynomial $p \in R$, define $\theta_i(p) = \frac{p(\mathbf{x}^{(i)}) - p(\mathbf{x}^{(i-1)})}{y_i - x_i}$, the *discrete differentiation* of p. For a sequence of $n+1$ polynomials $\mathbf{f} = (f_1, f_2, \ldots, f_{n+1}) \in R$, construct the following polynomial in $\mathbf{x}$ and $\mathbf{y}$:

$$\Theta_{\mathbf{f}} = \det \begin{pmatrix} f_1(\mathbf{x}) & \theta_1(f_1) & \cdots & \theta_n(f_1) \\ \vdots & \vdots & & \vdots \\ f_{n+1}(\mathbf{x}) & \theta_1(f_{n+1}) & \cdots & \theta_n(f_{n+1}) \end{pmatrix} = \sum_{\alpha,\beta} \theta^{\mathbf{f}}_{\alpha,\beta} \mathbf{x}^{\alpha} \mathbf{y}^{\beta}, \qquad (3.1)$$

where $\det(\cdot)$ denotes the determinant of the corresponding matrix, $\mathbf{f} = (f_1, f_2, \ldots, f_{n+1})$, and α and β vary in fixed ranges. This polynomial of $\mathbb{K}[\mathbf{x}, \mathbf{y}]$ is called the *Bezoutian* of $f_1, f_2, \ldots, f_{n+1}$. Note that it depends on the order that we have imposed on the variables. Its expansion in the monomial basis of the form

$$\Theta_{\mathbf{f}}(\mathbf{x}, \mathbf{y}) = \sum_{\alpha \in E} \mathbf{y}^{\alpha} \mathbf{w}_{\alpha}(\mathbf{x})$$

defines a map

$$\begin{aligned}\sigma : \mathbb{K}^E &\to R\\ (\lambda_\alpha)_{\alpha\in E} &\mapsto \sum_{\alpha\in E} \lambda_\alpha \mathbf{w}_\alpha.\end{aligned}$$

The matrix of this map in the monomial basis is precisely the matrix of the coefficients $[\theta^{\mathbf{f}}_{\alpha,\beta}]_{\alpha,\beta}$.

Definition 3.9. For any sequence of $n+1$ polynomials $\mathbf{f} = (f_1, f_2, \ldots, f_{n+1})$ in n variables, we denote by $B_{\mathbf{f}}$ the matrix $[\theta^{\mathbf{f}}_{\alpha,\beta}]_{\alpha,\beta}$ of the Bezoutian in the monomial basis. We call it the Bezoutian matrix of $\mathbf{f}$.

Example 3.10. (Continued from Section 3.5) The Bezoutian of

$$\begin{aligned}f_1 &= c_{10} + c_{11}x_1x_2 + c_{12}x_1^2x_2 + c_{13}x_1,\\ f_2 &= c_{20}x_2 + c_{21}x_1^2x_2^2 + c_{22}x_1^2x_2 + c_{23}x_1,\\ f_3 &= c_{30} + c_{31}x_2 + c_{32}x_1x_2 + c_{33}x_1.\end{aligned}$$

is of the form

$$\begin{aligned}&\theta_1(\mathbf{c})x_1{x_2}^2{y_1}^3y_2 + \theta_2(\mathbf{c})x_1{x_2}^2{y_1}^2y_2 + \theta_3(\mathbf{c})x_1x_2{y_1}^3y_2\\ &+\theta_4(\mathbf{c}){x_2}^2{y_1}^3y_2 + \theta_5(\mathbf{c})x_1{x_2}^2{y_1}^2 + \theta_6(\mathbf{c})x_1x_2{y_1}^2y_2\\ &+\theta_7(\mathbf{c}){x_2}^2{y_1}^2y_2 + \theta_8(\mathbf{c})x_2{y_1}^3y_2 + \theta_9(\mathbf{c})x_1{x_2}^2y_1 - \theta_{10}(\mathbf{c})x_1x_2{y_1}^2 + \theta_{11}(\mathbf{c}){x_2}^2{y_1}^2 + \theta_{12}(\mathbf{c})x_2{y_1}^2y_2\\ &+\theta_{13}(\mathbf{c})x_1{x_2}^2 + \theta_{14}(\mathbf{c})x_1x_2y_1 + \theta_{15}(\mathbf{c}){x_2}^2y_1 + \theta_{16}(\mathbf{c})x_2{y_1}^2\\ &+\theta_{17}(\mathbf{c}){y_1}^2y_2 + \theta_{18}(\mathbf{c})x_1x_2 + \theta_{19}(\mathbf{c})x_2y_1\\ &+\theta_{20}(\mathbf{c}){y_1}^2 + \theta_{21}(\mathbf{c})x_2 + \theta_{22}(\mathbf{c})y_1 + \theta_{23}(\mathbf{c})\end{aligned}$$

where $\theta_i(\mathbf{c})$ is a polynomial in the parameters $\mathbf{c} = (c_{i,j})$. The Bezoutian matrix associated to this polynomial is the following 5×5 matrix

$$\begin{bmatrix}
[142]-[114]+[411] & [431]-[134]-[341] & [143]-[241] & [421]-[124] & 0\\
[122] & -[224]+[423] & -[221]+[123]+[422] & [322] & [323]\\
[211]-[113] & [432]+[421]-[342]-[124]+[314] & [311]-[242]-[413]+[214]+[132] & [422]+[221]-[123] & [321]\\
[132]+[311] & -[343]+[433]-[234] & [314]-[231]+[432]+[133]-[342] & [321] & [324]\\
0 & [422] & [122] & [222] & [322]
\end{bmatrix}$$

where the symbol $[ijk]$ stands for the product $c_{1(i-1)}c_{2(j-1)}c_{3(k-1)}$. Its determinant equals $c_{12} * c_{21}^2 * c_{31} * R$.

This matrix is usually of much smaller size than the resultant matrix, as illustrated in this example.

Let us denote by I the ideal generated by $(f_2, \ldots, f_{n+1})$ and by $\mathcal{A}$ the quotient algebra $\mathcal{A} = R/I$[†]. An important property of these Bezoutians is given in the following proposition (see Cardinal and Mourrain, 1996), which is used in the next theorem.

Proposition 3.11. *Assume that the quotient $\mathcal{A} = R/I$ is a finite vector space of dimension d. Then there exist bases (a_i), (b_i) of R such that $(a_1, \ldots, a_d)$ (resp. $(b_1, \ldots, b_d)$)*

[†]Set of equivalence classes, modulo I.

is a basis of $\mathcal{A}$, $a_{d+1}, \ldots \in I$ *(resp.* $b_{d+1}, \ldots \in I$*), and for any polynomial* $f_1 \in R$*, the matrix of* $B_{f_1,f_2,\ldots,f_{n+1}}$ *in these bases is of the form*

$$\begin{pmatrix} M_{f_1} & 0 \\ 0 & L_{f_1} \end{pmatrix}, \tag{3.2}$$

where M_{f_1} *is the matrix of multiplication by* f_1 *in the basis* $(a_1, \ldots, a_d)$ *of* $\mathcal{A}$*.*

Let us first describe a direct application of this proposition, related to the *Chow Form* and based on the fact that the eigenvalues of M_{f_1} are $f_1(\zeta)$, for $\zeta \in \mathcal{Z}(f_2 = 0, \ldots, f_{n+1} = 0)$ (see Stetter, 1996; Auzinger and Stetter, 1988; Mourrain, 1998): let us recall that the Chow form of $(f_2, \ldots, f_{n+1})$ is

$$\prod_{\zeta \in \mathcal{Z}(f_2=0,\ldots,f_n=0)} (u_0 + u_1\zeta_1 + \cdots + u_n\zeta_n)^{\mu_\zeta}$$

where μ_ζ is the multiplicity of $\zeta \in \mathcal{Z}(f_2 = 0, \ldots, f_{n+1} = 0)$ (see Hodge and Pedoe, 1952). It can be shown that it is also the determinant $\det(u_0\mathbb{I}_d + u_1 M_{x_1} + \cdots + u_n M_{x_n})$, where M_{x_l} is the matrix of multiplication by x_l in the quotient $\mathcal{A} = R/(f_2, \ldots, f_{n+1})$. Therefore, it is a homogeneous polynomial in $\mathbf{u} = (u_0, \ldots, u_n)$ with coefficients in $\mathbb{K}$, of total degree the dimension $D = \sum_{\zeta \in \mathcal{Z}(f_2=0,\ldots,f_{n+1}=0)} \mu_\zeta$ of $\mathcal{A}$.

If we are able to compute this Chow form (or a multiple of it), then by factorization of this polynomial in $\mathbf{u}$, over $\overline{\mathbb{K}}$, we can recover the linear factors

$$u_0 + u_1\zeta_1 + \cdots + u_n\zeta_n$$

for $\zeta = (\zeta_1, \ldots, \zeta_n) \in \mathcal{Z}(I)$ and thus the coordinates $(\zeta_1, \ldots, \zeta_n)$ of the root ζ. Therefore, computing this Chow form (or a multiple of it) can lead to an efficient way to find the roots of the polynomial equations. An algorithm for factoring such polynomials has been proposed, for instance, in Carstensen (1992). Given the Bezoutian, computing a multiple of this Chow form is a direct consequence of Proposition 3.11, as explained below.

PROPOSITION 3.12. *Any non-zero minor of maximal size of the Bezoutian matrix* $B_{u_0+u_1x_1+\cdots+u_nx_n,f_2,\ldots,f_{n+1}}$ *is divisible by the Chow form* $C_I(\mathbf{u}) = \prod_{\zeta \in \mathcal{Z}(f_2=0,\ldots,f_{n+1}=0)} (u_0 + u_1\zeta_1 + \cdots + u_n\zeta_n)$*.*

The vanishing of the Chow form is a condition on the coordinates $\mathbf{u} = (u_0, \ldots, u_n)$ of a hyperplane to contain a root of the system of equations $f_2 = 0, \ldots, f_{n+1} = 0$. It can be seen as a special case of an eliminant condition.

We are going to show that the Bezoutian can be used to obtain a non-trivial multiple of the general resultant over a variety X, when this is meaningful (see Section 2.2). We have defined Bezoutians for *affine* polynomials but resultants are defined over projective varieties. To work on an affine space, we will consider a polynomial map $\sigma : \mathbb{A}^n \to X$, such that $\sigma(\mathbb{A}^n) = X^0$ is dense in X. Then $\tilde{f}_i = f_i \circ \sigma$ is a polynomial in the variables $x = (x_1, \ldots, x_n)$ and the Bezoutian $\Theta_{\tilde{f}_1,\ldots,\tilde{f}_{n+1}}$ is well defined. The next theorem shows that the resultant $\mathrm{Res}_X(f_1, \ldots, f_{n+1})$ can be recovered from the Bezoutian matrix $B_{\tilde{f}_1,\ldots,\tilde{f}_{n+1}}$. This result, although a direct consequence of Proposition 3.11, is new, to the best of our knowledge. It generalizes the result of Kapur *et al.* (1994) where, under a technical hypothesis, a multiple of the toric resultant is computed.

THEOREM 3.13. *Assume that Conditions 2.2 are satisfied and that* $\sigma : \mathbb{A}^n \to X$ *is a polynomial map such that its image is dense in* X. *Then any maximal minor of the Bezoutian matrix* $B_{\tilde{f}_1,\ldots,\tilde{f}_{n+1}}$ *is divisible by the resultant* $\mathrm{Res}_X(f_1,\ldots,f_{n+1})$.

PROOF. According to Conditions 2.2, the set of coefficients $\mathbf{c} = (c_{i,j})$ of $f_1,\ldots,f_{n+1}$ such that $\mathcal{Z}(f_2 = \cdots = f_{n+1} = 0)$ is finite is a dense subset of $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_n}$. As $X^0 = \sigma(\mathbb{A}^n)$ is a dense subset of X, the set of coefficients $c_{i,j}$ such that $\mathcal{Z}(f_2 = \cdots = f_{n+1} = 0)$ is finite and in X^0 is also a dense subset. Let us choose "generic" coefficients in this dense subset, for $f_2,\ldots,f_{n+1}$.

Then, the $\mathbb{K}$-vector space $\mathbb{K}[x_1,\ldots,x_n]/(\tilde{f}_2,\ldots,\tilde{f}_{n+1})$ is of finite dimension. Let us denote by D_g the generic dimension of this quotient. For any $\tilde{f}_1 \in R$, we denote by r_g the generic rank of the Bezoutian matrix $B_{\tilde{f}_1,\ldots,\tilde{f}_{n+1}}$. The minors of size r_g of $B_{\tilde{f}_1}$ are polynomials in $\mathbf{c}$, which are not all identically zero and any minor of size $r_g + 1$ is identically zero.

According to Proposition 3.11, for generic values of $\mathbf{c}$, the matrix $B_{\tilde{f}_1}$ can be decomposed as in (3.2), so that the rank of this matrix is

$$\mathrm{rank}(M_{\tilde{f}_1}) + \mathrm{rank}(L_{\tilde{f}_1}),$$

where $M_{\tilde{f}_1}$ is the matrix of multiplication modulo $f_2,\ldots,f_{n+1}$ and $L_{\tilde{f}_1}$ is the lower diagonal block in the decomposition (3.2). Since for generic values of $\mathbf{c}$, the variety $\mathcal{Z}(\tilde{f}_1 = \cdots = \tilde{f}_{n+1} = 0)$ is empty (Condition 2.2(2)), the multiplication matrix $M_{\tilde{f}_1}$ is generically invertible (the eigenvalues of $M_{\tilde{f}_1}$ are the values of f_1 at the roots of $\tilde{f}_2,\ldots,\tilde{f}_{n+1}$), that is of rank $D_g = \dim_{\mathbb{K}}\left(R/(\tilde{f}_2,\ldots,\tilde{f}_{n+1})\right)$.

Let us choose now $f_2,\ldots,f_{n+1}$ such that their roots are in X^0 and f_1 has a common root with $f_2,\ldots,f_{n+1}$. In this case, $\mathrm{Res}_X(f_1,\ldots,f_{n+1}) = 0$. Moreover, we have $\mathrm{rank}(M_{\tilde{f}_1}) < D_g$ (for $\tilde{f}_1$ vanishes at one of the roots of $\tilde{f}_2,\ldots,\tilde{f}_{n+1}$), and by specialization the rank of $L_{\tilde{f}_1}$ cannot exceed the generic rank. Thus, the matrix $B_{\tilde{f}_1}$ is of rank $< r_g$ and all the $r_g \times r_g$ minors vanish.

The set of systems $(f_1,\ldots,f_{n+1})$, such that $\mathcal{Z}(f_2 = \cdots = f_{n+1} = 0) \subset X^0$ and f_1 vanishes at one of these points, is a dense subset of the resultant variety $\mathcal{Z}(\mathrm{Res}_X(f_1,\ldots,f_{n+1}) = 0)$. Therefore any maximal minor of the Bezoutian matrix vanishes on this resultant variety. Consequently, any maximal minor (of size r_g) is divisible by the resultant, which proves the theorem. □

An important aspect of this construction is the size and the rank of the resulting Bezoutian matrix. We give here a bound on the number of monomials when we fix the degree d_i of the polynomials f_i. See Cardinal and Mourrain (1996) for more details. In this case, we consider generic polynomials of degree d_i and the bounds given below have to be compared with those of Macaulay matrices (or with the matrices associated with simplices in the toric case). If $d = \max_i(d_i)$ and n is the number of variables, then the size of the Bezoutian matrix is bounded by

$$e^n d^n$$

which is an order of magnitude less than the Macaulay matrix. Here is a table of the size of the Bezoutian matrix of $1, f_1,\ldots,f_n$ for small values of the degree $d_1,\ldots,d_n$. We give the size of the matrix, its rank and we compare it with the Bézout bound N, which is a lower bound on the generic rank of these matrices:

(d_i)	size	rank	N
2×2	5	4	4
3×2	9	6	6
4×2	14	9	8
4×3	18	13	12
$2 \times 2 \times 1$	9	5	4
$3 \times 2 \times 1$	19	8	6
$2 \times 2 \times 2$	14	8	8
$3 \times 2 \times 2$	28	14	12
$3 \times 3 \times 2$	43	20	18
$4 \times 3 \times 2$	70	30	24
$3 \times 3 \times 3$	55	29	27
$2 \times 2 \times 2 \times 2$	42	19	16
$3 \times 3 \times 3 \times 3$	273	99	81

A combinatorial result (Habsieger, 1998) shows that indeed the rank of $B_{1,f_2,\ldots,f_{n+1}}$ can not exceed $n^{-\frac{1}{2}}\left(\frac{e}{2}\right)^n d^n$ where $d = max_{i=2,\ldots,n+1}(d_i)$. For small values of n, this is not too far from d^n.

We end with some open questions. The maximal non-zero minors of the Bezoutian matrix are non-trivial multiples of the resultant. Can we obtain exactly the resultant by computing the gcd of these maximal minors, like in the Macaulay, or Newton formulation? Changing the variable order, replacing the variables x_i by powers $x_i^{e_i}$ (see Kapur and Saxena, 1997), can be helpful to remove some of the extraneous factors. But even then, the gcd of all possible maximal minors may remain reducible, as illustrated in the following example:

$$\begin{cases} f_1 = c_{1,1} + c_{1,2}x + c_{1,3}y \\ f_2 = c_{2,1} + c_{2,2}x + c_{2,3}y + c_{2,4}\left(x^2 + y^2\right) \\ f_3 = c_{3,1} + c_{3,2}x + c_{3,3}y + c_{3,4}\left(x^2 + y^2\right) \end{cases}$$

The Bezoutian matrix is a 3×3 matrix whose determinant factors as:

$$\begin{aligned} &(c_{1,3}c_{3,2}c_{2,4} - c_{1,3}c_{2,2}c_{3,4} - c_{1,2}c_{2,4}c_{3,3} + \cdots) \\ &\times \left(-2{c_{1,3}}^2 c_{2,4}c_{2,1}c_{3,1}c_{3,4} + 2c_{1,1}c_{1,3}c_{2,1}c_{2,4}c_{3,3}c_{3,4} + \cdots\right). \end{aligned}$$

Therefore, an important problem is to describe explicitly the factors that appear in such a maximal minor.

3.5. Dixon matrices

The two kinds of matrices that we have seen (ie. Bezoutian and Sylvester type matrices) can be mixed together, by choosing some of the coefficients $\mathbf{w}_\alpha(\mathbf{x})$ of the monomials $\mathbf{y}^\alpha$ in the Bezoutian and some multiples of the initial polynomials f_i in order to build a square matrix.

We call such matrices, which combine blocks of Sylvester and Bezout matrices, *Dixon matrices*, after the work of Dixon (1908), who proposed such matrix formulations for computing the resultant of three polynomials over $\mathbb{P}^2$. In Kapur *et al.* (1994), Dixon matrix refers to our Bezoutian matrix. In Jouanolou (1993a), they are also called Morley Matrices. Our terminology of Dixon matrix applies for all these kinds of matrices. It generalizes both Bézout and Sylvester construction of the resultant. We consider here two formulations of this type.

The first construction, which yields a smaller matrix than the Macaulay matrix, from which the resultant over $X = \mathbb{P}^n$ can also be constructed, is as follows. We consider the map

$$\begin{aligned}\Phi : \langle \mathbf{x}^{B_1} \rangle \times \cdots \times \langle \mathbf{x}^{B_{n+1}} \rangle \times \mathbb{K} &\to \langle \mathbf{x}^B \rangle \\ (q_1, \ldots, q_{n+1}, \lambda) &\mapsto \sum_{i=1}^{n+1} q_i f_i + \lambda \mathbf{w}_0,\end{aligned}$$

where $\mathbf{w}_0$ is the constant coefficient in $\mathbf{y}$ of the Bezoutian of the homogeneous polynomials $f_1, \ldots, f_{n+1}$, where $\mathbf{x}^{B_i}$ is the set of all monomials of degree $\sum_{j \neq i} d_j - n - 1$ and V is the set of all monomials of degree $\nu - 1$ where $\nu = \sum_{i=0} d_i - n$. Compared with Macaulay's formulation, the matrix of Φ is of size $\begin{pmatrix} \nu + n - 1 \\ n \end{pmatrix}$, which is less than the size of Macaulay's matrix.

The following theorem is essentially due to Macaulay:

PROPOSITION 3.14. (MACAULAY, 1902) *If Φ is surjective, then* $\mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_{n+1}) \neq 0$.

The converse is also true. See also Jouanolou (1991, 1993a,b) for more details on different formulations of this type and on projective resultants in several variables.

It is used in our case in the following way: $\mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_{n+1}) = 0$ implies that Φ is not surjective or equivalently that all the maximal minors of the matrix of Φ are divisible by the resultant.

PROPOSITION 3.15. *The maximal minors of the matrix of Φ are divisible by the resultant* $\mathrm{Res}_{\mathbb{P}^n}(f_1, \ldots, f_{n+1})$ *of $f_1, \ldots, f_{n+1}$ over $\mathbb{P}^n$.*

This construction has also been generalized to the case of toric varieties (with some restriction on the support of the polynomials) in Cattani and Dickenstein (1996).

In the second construction, the number of columns from the Bezoutian matrix is greater and the number of monomial multiples of the f_i (from the Sylvester matrix) is smaller. This is precisely the construction proposed by Dixon (1908) (for three polynomials in two variables), that we generalize slightly. The matrix that we consider is the matrix of a map mixing the Bezoutian and the Sylvester approach, of the form

$$\begin{aligned}\Phi : \langle \mathbf{x}^{B_1} \rangle \times \cdots \times \langle \mathbf{x}^{B_{n+1}} \rangle \times \mathbb{K} \times \cdots \times \mathbb{K} &\to \langle \mathbf{x}^B \rangle \\ (q_1, \ldots, q_{n+1}, \lambda_1, \ldots, \lambda_k) &\mapsto \sum_{i=1}^{n+1} q_i f_i + \sum_{i=1}^{k} \lambda_i \mathbf{w}_i,\end{aligned}$$

where $\mathbf{w}_i$ are polynomials which vanish when the polynomials $f_1, \ldots, f_{n+1}$ have a common root, and k is the number of polynomials $\mathbf{w}_i$. Here again to define this map, we have to specify the set of monomials $B_1, \ldots, B_{n+1}, B$ and the polynomials $\mathbf{w}_1, \ldots, \mathbf{w}_k$.

The polynomials $\mathbf{w}_i$ which were used in Dixon (1908) are precisely the coefficients of the first k monomials $\mathbf{y}^\alpha$ of smallest degree in the Bezoutian $\Theta(f_1, \ldots, f_{n+1}) = \sum_\alpha \mathbf{y}^\alpha \mathbf{w}_\alpha(\mathbf{x})$. We assume for a moment that the polynomials are of the same degree d. We will take for $\mathbf{x}^B$ the set of monomials of degree $\leq nd - n - u$ (where $u \in \mathbb{N}$ will be defined later). In order to obtain a construction which is symmetric in $f_1, \ldots, f_{n+1}$ we will assume that

the sets B_i are equal: $B_1 = \cdots = B_{n+1}$. Let $l = |B_1| = \cdots = |B_{n+1}|$. Now, we adjust the parameters k and l in such a way that:

- the number $k + (n+1) \times l$ of columns of the matrix is the number $|B|$ of rows,
- and the degree of the determinant of this matrix with respect to the coefficients of f_i is exactly d^n; that is, the degree of the resultant in the coefficients of f_i.

Thus, we obtain the following constrains :

$$k + (n+1)l = \tbinom{nd-u}{n}, \qquad k + l = d^n, \tag{3.3}$$

for some $u \in \mathbb{N}$, or

$$l = \frac{\binom{nd-u}{n} - d^n}{n} \in \mathbb{N}, \qquad k = \frac{(n+1)d^n - \binom{nd-u}{n}}{n} \in \mathbb{N},$$

for some value of $u \in \mathbb{N}$.

Example 3.16. For $n = 3$, $d = 2$, and for the system of the form

$$f_j(x_1, x_2, x_3) = \sum_{i_1,i_2,i_3 \in \mathbb{N}, 0 \le i_1+i_2+i_3 \le 2} c^j_{i_1,i_2,i_3,i_4} x_1^{i_1} x_2^{i_2} x_3^{i_3}, \qquad \text{for } 1 \le j \le 4,$$

we obtain $k = l = 4$ and the following 20×20 matrix:

	x_1^3	$x_1^2x_2$	$x_1^2x_3$	.....	x_1	x_2	x_1	1
f_1	0	0	0	.....	$c^1_{1,0,0}$	$c^1_{0,1,0}$	$c^1_{0,0,1}$	$c^1_{0,0,0}$
⋮	⋮				⋮	⋮	⋮	⋮
f_4	0	0	0	.....	$c^4_{1,0,0}$	$c^4_{0,1,0}$	$c^4_{0,0,1}$	$c^4_{0,0,0}$
x_1f_1	$c^1_{2,0,0}$	$c^1_{1,1,0}$	$c^1_{1,0,1}$	.....	$c^1_{0,0,0}$	0	0	0
⋮	⋮							⋮
x_1f_4	$c^4_{2,0,0}$	$c^4_{1,1,0}$	$c^4_{1,0,1}$	.....	$c^4_{1,1,0}$	0	0	0
x_2f_1	0	$c^1_{2,0,0}$	0	.....	0	$c^1_{0,0,0}$	0	0
⋮	⋮							⋮
x_2f_4	0	$c^4_{2,0,0}$	0	.....	0	$c^1_{0,0,0}$	0	0
x_3f_1	0	0	$c^1_{2,0,0}$	.....	0	0	$c^1_{0,0,0}$	0
⋮	⋮							⋮
x_3f_4	0	0	$c^4_{2,0,0}$	.....	0	0	$c^4_{0,0,0}$	0
$\mathbf{w}_{0,0,0}$	·	·	·	.....	·	·	·	·
$\mathbf{w}_{1,0,0}$	·	·	·	.....	·	·	·	·
$\mathbf{w}_{0,1,0}$	·	·	·	.....	·	·	·	·
$\mathbf{w}_{0,0,1}$	·	·	·	.....	·	·	·	·

where $\mathbf{w}_{j_1,j_2,j_3}$ is the coefficient of $y_1^{j_1} y_2^{j_2} y_3^{j_3}$ in the Bezoutian $\Theta(f_1, f_2, f_3, f_4)$. Thus each coefficient of $\mathbf{w}_{j_1,j_2,j_3}$ is a 4×4 determinant of the matrix of coefficients of the quadratic forms $f_1, \ldots, f_4$. The determinant of the matrix is of degree 8 with respect to each polynomial. This matrix has been used to deduce a polynomial of degree 40 in the direct kinematic problem of a parallel robot (see Mourrain, 1993, 1996a). Note that the corresponding Macaulay matrix is of size 56.

We have the property:

PROPOSITION 3.17. *It is not zero, the determinant of such a matrix (when it exists) is the resultant of the polynomials $f_1, \ldots, f_{n+1}$ over the projective space $X = \mathbb{P}^n$.*

An interesting challenge is to extend this construction to systems of equations f_i of degrees d_i not necessarily equal, as well as for the computation of resultants over $X = \mathbb{P}^{l_1} \times \cdots \times \mathbb{P}^{l_s}$ (of multihomogeneous equations). Dixon treated the case $X = \mathbb{P}^1 \times \mathbb{P}^1$.

3.6. COMPARISON BETWEEN DIFFERENT MATRICES

This section focuses on some properties of the toric resultant matrices and compares them to the matrix formulations of the classic resultant and to the Bézout type matrices. We also conjecture the extension of Macaulay's exact rational expression to the context of toric resultants.

The main characteristic of the Sylvester-type matrices is that the coefficients are either 0 or the coefficients of the input polynomials. This type includes the Macaulay and toric resultant (or Newton) matrices. The columns of these matrices correspond to multiples of the polynomial f_i and the difficulty for constructing these matrices relies on the choice of these multiples. In the case of toric resultants, this is performed by geometric considerations on the support of the polynomials.

The subdivision algorithm generalizes the classical Macaulay construction in the sense that it produces the same matrix on completely dense systems. For instance, if we take the following lifting and perturbation:

$$\begin{aligned} \delta &= (\epsilon, \ldots, \epsilon), \\ l_i &= L_i x_1 + \cdots + L_i x_{i-1} + x_i + L_i x_{i+1} + \cdots + L_i x_n, i = 1, \ldots, n, \\ l_{n+1} &= L_{n+1} x_1 + \cdots + L_{n+1} x_n, \\ &\quad \text{where } L_1 \gg L_2 \gg \cdots L_n \gg L_{n+1} \gg 1 \gg \epsilon > 0, \end{aligned}$$

we obtain the Macaulay matrix of $f_{n+1}, \ldots, f_1$. Thus, Macaulay matrices are a special case of Newton matrices. The latter matrices are usually smaller but require more computation on the polytopes of the f_i.

Macaulay's impressive result is the derivation of the extraneous factor in the matrix determinant as a minor of the matrix. The extension of this formula to the case of the toric resultant matrix is a major open question. The natural way to generalize Macaulay's result is by defining $E^{nm} \subset B$ to be the subset of these points that do not lie in any i-mixed cell, for any $i \in \{1, \ldots, n+1\}$. Let S^{nm} denote the square submatrix of S that includes all entries whose row and column indices lie in E^{nm}. As noted above, S^{nm} is generically non-singular. Based on empirical results, Canny and Emiris (19996) have stated the following conjecture.

CONJECTURE 3.18. *There exists a perturbation vector δ and lifting functions $l_1, \ldots, l_{n+1}$ for which the determinant of matrix S^{nm} divides exactly the determinant of a Newton matrix S and, hence, the toric resultant of the given polynomial system is $R = \det S / \det S^{nm}$.*

The proposed rational expression has the same degree as R in every polynomial and, furthermore, equals the resultant in the completely dense case by Macaulay's result.

Example 3.19. (Continued from Section 3.5) Referring to the example of Section 3.5, we have

$$S^{nm} = \begin{bmatrix} c_{23} & 0 & 0 & 0 \\ 0 & c_{32} & c_{33} & 0 \\ c_{34} & 0 & c_{32} & 0 \\ 0 & 0 & 0 & c_{33} \end{bmatrix},$$

and $\det S / \det S^{nm}$ produces the correct toric resultant. Hence the conjecture is verified here.

The Sylvester type matrices admit a natural generalization to overconstrained systems $f_1, \ldots, f_m$ $(m \geq n+1)$, just by adding new blocks S_i depending on the coefficients $f_{n+2}, \ldots, f_m$. See, for instance, Lazard (1981), where following Macaulay (1902), a rectangular matrix with $|B|$ rows and more than $|B|$ columns has been used. See also Grigoryev and Vorobjov (1988); Grigoryev (1988).

Let us compare Sylvester-type matrices with Bezoutian-type matrices. In practice, the size of the former is usually larger than the size of the Bézout-type matrices. The monomials in $\mathbf{x}$ (resp. $\mathbf{y}$) of $\Theta_f(\mathbf{x}, \mathbf{y})$ also depend on the polytope of the f_i. In Kapur *et al.* (1996) for instance, it is shown that the monomials in $\mathbf{x}$ (resp. $\mathbf{y}$) of the Bezoutian are in the Minkowski sum of projections on some coordinate hyperplanes of the polytopes generated by the polynomials f_i.

The entries of the Bezoutian are sums of determinants of the coefficients of the input polynomials f_i, for the Bezoutian is obtained by expansion of a determinant of a matrix in polynomials. Thus this object is more difficult to compute.

A numerical comparison in solving methods based on these matrices has been initiated in Emiris and Mourrain (1996), and seems to give an advantage to the Bezoutian approach in terms of accuracy. The structure of the Bezoutian is more difficult to analyze, but just as in the univariate case, the reduction of $\Theta(f_1, \ldots, f_{n+1})$ modulo the ideal $(f_2, \ldots, f_{n+1})$ leads to a matrix whose inverse is of Hankel type.

Besides all these properties, one major advantage of Bezoutians is that they directly give informations on the isolated points of the variety (see Section 4.4 and Elkadi and Mourrain, 1998). Moreover, many interesting and powerful properties of the quotient algebra $\mathcal{A} = R/(f_2, \ldots, f_{n+1})$ are connected to these matrices, including duality, algebraic residues (Scheja and Storch, 1975; Kunz, 1986; Elkadi and Mourrain, 1996), real root counting (Becker *et al.*, 1996), multiplicities (Mourrain, 1996b), ...

4. Applications of Resultant Matrices

4.1. Monomial bases and multiplication maps

This section establishes certain facts about the coordinate ring of the variety associated to a well-constrained system. Particularly useful are monomial bases, since they index matrices that define multiplication maps in the corresponding coordinate ring. Multiplication maps are directly obtained from resultant matrices, given by any of the formulations seen so far.

The basic property of all resultant matrices is that premultiplication with certain row vectors expresses evaluation of the polynomials whose coefficients have filled in the matrix columns. For an arbitrary $\zeta \in \overline{\mathbb{K}}^n$, and S of Sylvester type:

$$[\ \cdots\ \zeta^p\ \cdots\ ] \begin{bmatrix} & \cdots q \cdots & \\ & S & \end{bmatrix} \begin{matrix} \vdots \\ p \\ \vdots \end{matrix} = [\ \cdots\ \zeta^q f_i(\zeta)\ \cdots\ ], \tag{4.4}$$

where p and q range over the points indexing rows and columns in S, respectively, and $\mathbf{x}^q f_i(\mathbf{x})$ defines the contents of the column indexed by q. The vector $[\cdots, \zeta^q, \cdots]$ is indexed by the points q and contains the values of column monomials $\mathbf{x}^q$ at ζ. The vector product is indexed by points p and contains, at the entry indexed by p, the value of $\mathbf{x}^p f_i(\mathbf{x})$ at ζ. For Macaulay's matrix, points p lie in the disjoint union of all sets B_i and q ranges over B, following the notation of Section 3.1. In the toric context, points p and q range over B, in the notation of Section 3.2. For the Bézout-type matrices, we have to replace the multiples $\mathbf{x}^p f_i(\mathbf{x})$ by some polynomials $\mathbf{w}_\alpha(\mathbf{x})$ which appear in the expansion of $\Theta(f_1, \ldots, f_{n+1})$.

The premultiplication property has almost reduced the calculation of all common roots to computing the kernel vectors of S. In Section 4.3 we see that the problem may be reduced to computing the eigenvalues and eigenvectors of a square matrix. For now, the premultiplication property is used in connection to monomial bases and multiplication maps.

Consider a well-constrained system $f_2, \ldots, f_{n+1} \in \mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]$. If the ideal I generated by these polynomials corresponds to a variety of zero dimension in $\overline{\mathbb{K}}^n$, then its coordinate ring $\mathcal{A} = \mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]/I$ is a vector space over $\mathbb{K}$, of finite dimension (Cox *et al.*, 1992). Its dimension is the number of roots counted with multiplicity. In the toric case, generically, the vector space's dimension is $\mathrm{MV}(f_2, \ldots, f_{n+1})$, whereas in the projective case, it is $\prod_{i=2}^{n+1} \deg f_i$. In the toric elimination context there is a stronger property demonstrated in Pedersen and Sturmfels (1996) and also proven directly from the subdivision-based construction in Emiris and Rege (1994).

PROPOSITION 4.1. *Assume that* $\mathcal{Z}_{(\overline{\mathbb{K}}^*)^n}(f_2, \ldots, f_{n+1})$ *is a finite set of* simple *roots. Consider the integer lattice points lying in the mixed cells of the mixed subdivision of* $\sum_{i=1}^n A_i$ *after applying an infinitesimal perturbation as in Section 3.2. These points are in bijective correspondence with monomials in the* n *variables* $\mathbf{x}$ *and this monomial set forms a vector-space basis for* $\mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]/I$.

In establishing this proposition, the proof of Emiris and Rege (1994) defines an overconstrained system by adding a generic polynomial f_1. Then, in appropriate mixed subdivisions of $\sum_{i=1}^{n+1} A_i$ such as those of Section 3.2, the integer points defining the monomial basis of $\mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]/I$ are precisely those lying in the 1-mixed cells.

Let this point set be denoted $B_1 \subset B$, with cardinality $\mathrm{MV}(f_2, \ldots, f_{n+1})$. The sets B_1 and $B \setminus B_1$ partition S into four blocks $S_{i,j}$, for $i = 1, 2$, $j = 1, 2$, where $S_{1,1}$ and $S_{2,2}$ are square of dimension $|B_1|$ and $|B \setminus B_1|$, respectively. As discussed in Section 3.2, the subdivision-based algorithm guarantees that submatrix $S_{2,2}$ is generically non-singular, so we can define $M = S_{1,1} - S_{1,2} S_{2,2}^{-1} S_{2,1}$, of dimension $|B_1|$. Then, it is easy to show that M defines a multiplication map in coordinate ring $\mathcal{A}$ for the polynomial f_1. This is

an endomorphism in $\mathcal{A} = \mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]/I$, such that

$$g \mapsto gf_1 \bmod I.$$

In other words, if the polynomial $g \in \mathcal{A}$ is represented by a column vector with respect to monomial basis B_1, then premultiplication of M by this vector yields another column vector expressing $gf_1 \bmod I$ in the same basis.

Information on a basis of the coordinate ring can also be recovered from the Bezoutian.

PROPOSITION 4.2. (ELKADI AND MOURRAIN, 1996) *Assume that $\mathcal{Z}_{\overline{\mathbb{K}}^n}(f_2, \ldots, f_{n+1})$ is a finite set of points in $\mathbb{A}^n$. Then the set of monomials in $\mathbf{x}$, respectively in $\mathbf{y}$, which appear in the Bezoutian of $1, f_2, \ldots, f_{n+1}$ contains a basis of the coordinate ring $\mathbb{K}[\mathbf{x}, \mathbf{x}^{-1}]/I$, or $\mathbb{K}[\mathbf{x}]/I$.*

This also holds in the non-generic case, for any complete intersection. The advantage of obtaining a monomial basis in the non-generic case, provided the n polynomials express a complete intersection, is important in practical applications of resultants. We return to genericity issues and their significance in Section 4.4.

4.2. SYSTEM SOLVING BY THE u-RESULTANT

The main goal in system solving is to find all common isolated roots of a well-constrained system. The computation of the u-resultant is a standard tool for finding all isolated roots. A related approach reduces the problem to solving a single equation in one variable, then "lifting" these solutions to the common roots of the original system. We restrict our attention to zero-dimensional systems defining a complete intersection. Yet, the approach can be extended to arbitrary systems, including overconstrained ones, through the techniques of Section 4.4. This section continues the discussion above and uses the same notation.

Let f_1 be linear with all coefficients u_i being symbolic,

$$f_1 = u_0 + u_1 x_1 + \cdots + u_n x_n,$$

so that $f_1 \notin I = (f_2, \ldots, f_{n+1})$ generically. We consider a specialization of the coefficients of $f_2, \ldots, f_{n+1}$. Whenever f_1 vanishes at a common root $\zeta = (\zeta_1, \ldots, \zeta_n) \in \overline{\mathbb{K}}^n$ of the original system, then the resultant must vanish. Therefore, $u_0 + u_1\zeta_1 + \cdots + u_n\zeta_n$ divides $\mathrm{Res}(u_0, \ldots, u_n)$. This is the classic approach based on the u-resultant for system solving (Macaulay, 1916; van der Waerden, 1950; Lazard, 1981; Canny, 1988; Renegar, 1992). The last three methods yield single-exponential algorithms (in the problem dimension) for finding all isolated roots in the generic cases. The same is true with the toric resultant construction. It is also possible to use the non-zero maximal minors of the Bezoutian of these polynomials, as explained in Proposition 3.12. In both cases, we obtain a polynomial in $\mathbf{u}$, which is divisible by the linear factors $u_0 + u_1\zeta_1 + \cdots + u_n\zeta_n$, for $\zeta \in \mathcal{Z}(f_2, \ldots, f_{n+1})$.

Let us describe now two approaches, which can be used at this point to recover the roots.

First, we compute explicitly the minor and factor it over $\overline{\mathbb{K}}$. From the linear factors, we deduce all the coordinates of the roots. Of course, factoring this polynomial may give more linear factors than those corresponding to roots, but this is not a severe limitation because we still obtain a superset of all solutions. Algorithms for computing a

numerical approximation of such factorization can be found in Carstensen (1992), for instance.

Another approach consists in computing the coefficient $d(u_0)$ of 1 and the coefficients $d_i(u_0)$ of x_i $i = 1, \ldots, n$ of this minor, when we replace u_i by $u_i + v_i$ for some $v_i \in \mathbb{K}$. For this, we do not necessarily need to compute the whole polynomial, which may be huge. Then we deduce a rational representation of the roots $d_0(u_0) = 0, x_i = f_i(u_0)$ $i = 1, \ldots, n$, where f_i is a rational fraction in u_0. See Macaulay (1912, p. 88); Rouillier (1996); Elkadi and Mourrain (1998).

The u-resultant, however, is identically zero when the variety has positive dimension, even if the positive-dimensional component lies at infinity. This is projective infinity for the classical context and toric infinity for the toric resultant, as mentioned in Section 2.4. Then the techniques of Section 4.4 have to be applied. Note, however, that the Bezoutian construction is still valid in the case where a positive-dimensional component lies at infinity or even in the affine part. What we obtain here are the isolated roots of the variety (see Elkadi and Mourrain, 1998, 1999).

4.3. system solving by eigenvectors

This section details the reduction of solving the initial non-linear problem to an eigenproblem. This is more efficient than the above method because it does not require factorization of large polynomials.

For this, observe that the constant polynomial $f_1 = u_0$ has multiplication map $u_0\mathbb{I}$ where $\mathbb{I}$ is the identity matrix of dimension $\dim_{\mathbb{K}}(\mathcal{A})$. If now $f_1 = u_0 + v_1x_1 + \cdots + v_nx_n$ (where v_i are random constants) by the linearity property, the multiplication map of f_1, is of the form $M_{f_1+u_0} = M_l + u_0\mathbb{I}$ with $l = v_1x_1 + \cdots + v_nx_n$. Therefore, when $f_1(\zeta) = 0$ then $-u_0$, is an eigenvalue of the matrix M_l.

The obvious question then is, what are the eigenvectors expressing? Recall that premultiplication by a vector representing the evaluation of the row monomials at some $\zeta \in \overline{\mathbb{K}}^n$ yields the values of the column polynomials at ζ. In practice, only the constant coefficient of f_1 is an indeterminate, whereas all other coefficients are specialized to random values. Then for any root ζ of $f_1 = \cdots = f_n = 0$ all vectors of the form $[\ldots, \zeta^b, \ldots]$, where b ranges over B, are eigenvectors of M_l, based on expression (4.4). The latter matrix has numeric entries, so the problem can be solved by linear algebra methods. Hence the eigenproblem offers a necessary condition for the roots ζ. By exploiting the fact that the eigenvectors of interest exhibit a special structure, it is possible to recover all root coordinates (Auzinger and Stetter, 1988; Manocha and Canny, 1992; Emiris, 1996; Mourrain, 1998).

If we use Bezoutian-type matrices instead of Sylvester-type matrices, we have to replace the eigenproblem $(M_l + u_0\mathbb{I}_d)v = 0$ by the generalized eigenproblem $(B_l + u_0B_1)w = 0$. When B_1 is invertible, these two problems are equivalent, and when B_0 is not invertible, compression techniques can be used in order to reduce the generalized eigenproblem to a non-singular one (see Mourrain, 1998).

We have reduced root finding to a problem in linear algebra by adding the u-form to the given well-constrained system. An alternative is to "hide" one of the n variables in the coefficient field. This produces an overconstrained system without increasing the problem dimension. Our experience with systems in robotics and vision suggests that this is preferable in many practical situations (Emiris, 1997). Formally, we consider the

given polynomials as

$$f_1, \ldots, f_n \in (\mathbb{K}(x_n))\,[x_1, x_1^{-1}, \ldots, x_{n-1}, x_{n-1}^{-1}].$$

The variable x_n is chosen so that all roots are separated by projection on x_n, if possible. Otherwise, we have to deal with the case of multiple roots. The construction of S is as before, and any algorithm can be used. There is a generating set of the coordinate ring defined and made up of monomials. Moreover, a multiplication map of larger-than-optimal dimension is in general obtained, and a non-trivial multiple of the u-resultant can be computed. S' is defined by eliminating the largest possible number of rows in S, which are constant with respect to x_n. Then the problem is reduced to an eigendecomposition of S'.

Row and column permutations do not affect the matrix properties so we apply them to obtain a maximal S_{11}. Gaussian elimination gives $S'(x_n) = S_{22}(x_n) - S_{21}(x_n) S_{11}^{-1} S_{12}(x_n)$. In contrast to the previous approach, the matrix may be non-linear in the hidden variable, so $S'(x_n)$ is a matrix polynomial $A_d x_n^d + \cdots + A_1 x_n + A_0$, for some $d \geq 1$. If A_d is numerically non-singular, we reduce the equation $w(\mathbb{I} x_n^d + A_d^{-1} A_{d-1} x_n^{d-1} + \cdots + A_d^{-1} A_0) = 0$ to the following eigenproblem:

$$w \begin{bmatrix} 0 & \cdots & 0 & -A_0 A_d^{-1} \\ \mathbb{I} & \ddots & \vdots & \vdots \\ \vdots & \ddots & 0 & -A_{d-2} A_d^{-1} \\ 0 & \cdots & \mathbb{I} & -A_{d-1} A_d^{-1} \end{bmatrix} = x_n w,$$

where $w = [w, x_n w, \ldots, x_n^{d-1} w]$. If A_d is numerically singular, we may change the variable x_n to $(t_1 y + t_2)/(t_3 y + t_4)$ for random $t_1, \ldots, t_4$. This *rank balancing* improves the numeric rank of A_d. If the latter is still non-singular, we have to consider a generalized eigenproblem on the following matrix pencil:

$$C(y) = \begin{bmatrix} \mathbb{I} & & & \\ & \ddots & & \\ & & \mathbb{I} & \\ & & & A_d \end{bmatrix} y + \begin{bmatrix} 0 & \cdots & 0 & A_0 \\ -\mathbb{I} & \ddots & \vdots & A_1 \\ \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & -\mathbb{I} & A_{d-1} \end{bmatrix}.$$

For every eigenvalue λ with associated left eigenvector $w = [w_1, \ldots, w_d]$ of $C(y)$, we have $w_i = \lambda^{i-1} w_1$ for $i = 2, \ldots, d$. Moreover, $s'(y)$ has the same eigenvalue λ and has left eigenvector w_1. These are all standard operations in numerical linear algebra (Golub and Van Loan, 1996). This is also valid for Bezoutian-type matrices if we replace eigenvector problems by generalized eigenvector problems.

Indeed, one can apply quite general methods to a number of different resultant matrices for solving polynomial systems. These numerical matrix manipulations manage to obtain smaller matrices that contain all relevant information through a Schur factorization, Jordan decomposition or singular value decomposition, respectively (Manocha and Demmel, 1995; Corless *et al.*, 1997; Mourrain, 1998), or computing maximal non-singular minors (Kapur *et al.*, 1994; Manocha, 1995; Cardinal and Mourrain, 1996).

Several implementations of resultant matrices use these matrix manipulations in order to solve systems of arbitrary polynomial equations. We mention those by Krishnan and Manocha (1995), and by Emiris (1997). A similar approach, generalizing Lazard's (1981) method and implemented in MATLAB, is proposed by Corless *et al.* (1997). The

second author is developing a program for system solving based on the Bézout matrix (Mourrain, 1998), and so are Kapur *et al.* (1996). A C++ library linking some linear algebra packages such as LAPACK or UMFPACK, called ALP, is currently under development, in order to gather all these different techniques (see `http://www.inria.fr/saga/logiciels/ALP/`).

4.4. GENERICITY ISSUES

We have seen that resultants provide exact conditions for the existence of roots only in projective space or in some toric variety. Otherwise, they provide only necessary conditions. The computation of resultant matrices may be fruitless in the case of specific coefficient specializations, for instance when the underlying system has positive-dimensional components in the projective variety X. For instance, a positive-dimensional component at infinity causes the resultant polynomial to vanish identically. In these degenerate cases, the constructed matrices may be identically singular, thus offering no indication of the vanishing of the resultant.

Recent work focuses on the degenerate cases and adapts resultant theory so that it applies to arbitrary inputs. One of the goals is to extract useful information on the isolated roots, even in the presence of positive-dimensional components, by generalizing the classic resultant and the toric resultant.

In the context of the classical theory, Canny analyzed the generalized characteristic polynomial for computing the resultant over general algebraically closed fields (Renegar, 1989; Canny, 1990). Its name is due to the fact that it generalizes the characteristic polynomial of linear systems. See also Chistov (1986), Grigoryev (1986). It provides a projection operator that is not identically zero in the presence of positive-dimensional components at infinity or elsewhere. The input polynomial system is perturbed by polynomials defined on the new supports and multiplied by ϵ. The resultant of the perturbed system is an ϵ-polynomial, and its constant term is identically zero iff the system's variety has positive dimension in an appropriate variety. Its non-zero coefficient of lowest degree generalizes the resultant and, in particular, the u-resultant, in the case that a u-form is the first polynomial. This operator generalizes the u-resultant and offers a necessary condition for the existence of isolated roots. Although the construction was proposed for Macaulay's matrix, it can also be coupled with Lazard's (1981) matrix. The generalized characteristic polynomial has been recently adapted to the toric context by Rojas (1999).

Let us mention here that these perturbation techniques are not required for Bezoutian matrices. Indeed, in the case of any affine complete intersection (whatever the situation at infinity is), the maximal minors of these matrices yield a multiple of the u-resultant or Chow form (see Proposition 3.12). But even when the variety has a positive-dimensional component in the affine part, these minors give a non-trivial multiple of the Chow form of the *isolated roots* (see Elkadi and Mourrain, 1998).

These methods define a perturbed determinant in terms of some parameter. Recovering the trailing term of this determinant typically requires some exact computation that increases the practical complexity of the problem. On the contrary, the approach (Mourrain, 1998) avoids degeneracies purely by matrix operations and applies for different kind of formulations. It works directly with the (degenerate) resultant matrices and more precisely with pencils of matrices of the form $M_i - z_i M_0$ associated to these resultant matrices. The so-called Kronecker decomposition of such pencils yields a left and a right singular part and a regular part, which can be used to solve the system. This

method has the advantage of being practical even with approximate coefficients, stable algorithms being available for computing this regular part (e.g. Demmel and Kågström, 1993).

5. Conclusion

Sparse elimination theory is a comparatively recent algebraic approach, studied in the past two decades and dealing with polynomials described by their monomial supports. This leads to more efficient algorithms in practice, and calls for several combinatorial and geometric techniques. Bézout and Dixon matrices provide more compact conditions and may prove to be numerically more stable. In addition to the general overview here, we have established a new result concerning the relevance of minors in the Bézout matrix. This survey has reviewed the state of the art in constructing resultant matrices, the major step in reducing system solving to a problem in linear algebra as well as for computing the resultant polynomial. We have also described several methods for solving arbitrary systems of polynomial equations, including the degenerate cases. We have also pointed out main open issues in this domain, which is expected to be equally active in the years to come.

Acknowledgements

The first author thanks Maurice Rojas for eagerly providing preprints and explanations of his work. Both authors are partially supported by European ESPRIT project FRISCO (LTR 21.024).

References

Aizenberg, L. A., Kytmanov, A. M. (1981). Multidimensional analogues of Newton's formulas for systems of non-linear algebraic equations and some of their applications. *Trans. Sib. Mat. Zhurnal*, **22**, 19–39.

Auzinger, W., Stetter, H. J. (1988). An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations. In *Proc. Int. Conf. on Numerical Mathematics,*volume 86,*International Series of Numerical Mathematical*, pp. 12–30. Basel, Birkhäuser Verlag.

Bajaj, C., Garrity, T., Warren, J. (1988). On the applications of multi-equational resultants. Technical Report 826, Purdue University.

Balbes, L. M., Mascarella, S. W., Boyd, D. B. (1994). A perspective of modern methods in computer-aided drug design. In Lipkowitz, K. B. and Boyd, D. B., eds, *Reviews in Computational Chemistry*, volume 5, pp. 337–379. New York, VCH Publishers.

Barvinok, A. I. (1993). A polynomial time algorithm for counting integral points in polyhedra when the dimension is fixed. In *Proc. 34th Ann. Symp. on Foundations of Computer Science, Palo Alto, CA*, pp. 566–572. Los Alamitos, CA, IEEE Computer Society Press.

Basu, S., Pollack, R., Roy, M.-F. (1997). Computing roadmaps of semi-algebraic sets on a variety. In Cucker, F. and Shub, M., eds, *Proc. Workshop on Foundations of Computational Mathematics*, pp. 1–15. Berlin, Springer Verlag.

Becker, E., Cardinal, J. P., Roy, M. F., Szafraniec, Z. (1996). Multivariate Bezoutians, Kronecker symbol and Eisenbud-Levin formula. In González-Vega, L. and Recio, T., eds, *Algorithms in Algebraic Geometry and Applications,* volume 143, *Progress in Mathematics*. pp. 79–104. Basel, Birkhäuser.

Berenstein, C. A., Gay, R., Vidras, A., Yger, A. (1993). *Residue Currents and Bezout Identities*. volume 114, *Progress in Mathematics*, Basel, Birkhäuser.

Berenstein, C. A., Yger, A. (1991). Effective Bezout identities in $\mathbb{Q}[z_1, \ldots, z_n]$. *Acta. Math.*, **166**, 69–120.

Bernstein, D. N. (1975). The number of roots of a system of equations. *Funct. Anal. Appl.*, **9**, 183–185.

Bézout, E. (1779). *Théorie Générale des Équations Algébriques.* Paris.

Billera, L. J., Sturmfels, B. (1992). Fiber polytopes. *Ann. Math.*, **135**, 527–549.

Bini, D., Pan, V. (1994). *Polynomial and Matrix Computations,* volume 1*, Fundamental Algorithms*. Boston, Birkhäuser.

Bondyfalat, D., Mourrain, B., Pan, V. Y. (1998). Controlled iterative methods for solving polynomial systems. In *Proceedings of ISSAC*, pp. 252–259. New York, ACM Press.

Burago, Y. D., Zalgaller, V. A. (1988). *Geometric Inequalities*. Grundlehren der mathematischen Wissenschaften, **285**, Berlin, Springer Verlag.

Busé, L., Elkadi, M., Mourrain, B. Generalized resultant over an algebraic variety.Submitted, 1999.

Canny, J. (1990). Generalised characteristic polynomials. *J. Symb. Comput.*, **9**, 241–250.

Canny, J. (1993). Improved algorithms for sign determination and existential quantifier elimination. *The Comput. J.*, **36**, 409–418.

Canny, J. F. (1988). *The Complexity of Robot Motion Planning*. Cambridge, MA, M.I.T. Press.

Canny, J., Emiris, I. (1993). An efficient algorithm for the sparse mixed resultant. In Cohen, G., Mora, T. and Moreno, O., eds, *Proc. Int. Symp. on Appl. Algebra, Algebraic Algorithms and Error-Corr. Codes, Puerto Rico*, LNCS **263**, pp. 89–104. Berlin, Springer Verlag.

Canny, J. F., Emiris, I. Z. (1999). A subdivision-based algorithm for the sparse resultant. *J. ACM*, to appear.

Canny, J. F., Kaltofen, E., Lakshman, Y. (1989). Solving systems of non-linear polynomial equations faster. In *Proceedings of ISSAC*, pp. 121–128. New York, ACM Press.

Canny, J., Pedersen, P. (1993). An algorithm for the Newton resultant. Technical Report 1394, Computer Science Dept., Cornell University.

Canny, J., Rojas, J. M. (1991). An optimal condition for determining the exact number of roots of a polynomial system. In *Proceedings of ISSAC, Bonn, 1991*, pp. 96–102. New York, ACM Press.

Cardinal, J. P. (1993). Dualité et algorithmes itératifs pour la résolution de systèmes polynomiaux. Ph.D. Thesis, Université de Rennes.

Cardinal, J. P., Mourrain, B. (1996). Algebraic approach of residues and applications. In Renegar, J., Shub, M. and Smale, S., eds, *Proceedings AMS-SIAM Summer Seminar on Mathematics of Numerical Analysis, Park City, UT, 1995*, volume 32, *Lectures in Applied Mathematics*, pp. 189–210. New York, Americal Mathematical Society Press.

Carstensen, C. (1992). On Grau's method for simultaneous factorization of polynomials. *SIAM J. Numer. Anal.*, **29**, 601–613.

Cattani, E., Dickenstein, A. (1996). A global view of residues in the torus. *J. Pure Appl. Algebra*, **117–118**, 119–144.

Cayley, A. (1848). On the theory of elimination. *Dublin Math. J.*, **II**, 116–120.

Chistov, A. L. (1986). Algorithm of polynomial complexity for factoring polynomials and finding the components of varieties in subexponential time. *J. Sov. Math.*, **34**, 1838–1882.

Corless, R. M., Gianni, P. M., Trager, B. M. (1997). A reordered Schur factorization method for zero-dimensional polynomial systems with multiple roots. In Küchlin, W. W., ed., *Proceedings of ISSAC*, pp. 133–140. New York, ACM Press.

Cox, D., Little, J., O'Shea, D. (1992). *Ideals, Varieties, and Algorithms,* Undergraduate Texts in Mathematics. New York, Springer Verlag.

Cox, D. (1995). The homogeneous coordinate ring of a toric variety. *J. Algebr. Geom.*, **4**, 17–50.

Danilov, V. I. (1978). The geometry of toric varieties. *Russian Math. Surveys*, **33**, 97–154.

Demmel, J., Kågström, B. (1993). The generalized Schur decomposition of an arbitrary pencil $A - \Lambda B$: robust software with error bounds and applications I,II. *ACM Trans. Math. Software*, **19**, 160–201.

Dixon, A. L. (1908). The eliminant of three quantics in two independent variables. *Proc. Lond. Math. Soc.*, **6**, 49–69, 473–492.

Elkadi, M., Mourrain, B. (1996). Approche effective des résidus algébriques. Rapport de Recherche 2884, INRIA.

Elkadi, M., Mourrain, B. Some applications of bezoutians in effective algebraic geometry. Rapport de Recherche 3572, INRIA, 1998.

Elkadi, M., Mourrain, B. (1999). Algorithms for residues and Lojasiewicz exponents. *J. Pure Appl. Algebra*, to appear.

Elkadi, M., Mourrain, B. (1999). A new algorithm for the geometric decomposition of a variety,*Proceedings of ISSAC*. New York, ACM Press, to appear.

Emiris, I. Z. (1996). On the complexity of sparse elimination. *J. Complexity*, **12**, 134–166.

Emiris, I. Z. (1997). A general solver based on sparse resultants: Numerical issues and kinematic applications. Technical Report 3110, France, INRIA Sophia-Antipolis.

Emiris, I. Z., Canny, J. F. (1995). Efficient incremental algorithms for the sparse resultant and the mixed volume. *J. Symb. Comput.*, **20**, 117–149.

Emiris, I. Z., Mourrain, B. Polynomial system solving; the case of a 6-atom molecule. Rapport de Recherche 3075, INRIA, 1996.

Emiris, I., Pan, Y. V. (1999). Symbolic and numeric methods for exploiting structure in constructing resultant matrices. *J. Symb. Comput.*, to appear.

Emiris, I. Z., Rege, A. (1994). Monomial bases and polynomial system solving. In *Proc. ACM Int. Symp. on Symbolic and Algebraic Computation, Oxford, July 1994*, pp. 114–122.

Emiris, I., Verschelde, J. (1999). How to count efficiently all affine roots of a polynomial system. *Discrete Appl. Math.*, to appear.
Fitchas, N., Giusti, M., Smietanski, M. (1993). Sur la complexité du théoréme des zéros. In Gudatt, J., ed., *Proc. Second. Int. Conf. on Approximation and Optimization, La Habana, 1993,* pp. 274–329. Peter Lang Verlag.
Fulton, W. (1993). *Introduction to Toric Varieties,* number 131, *Annals of Mathematics*. Princeton, NJ, Princeton University Press.
Gelfand, I. M., Kapranov, M. M., Zelevinsky, A. V. (1994). *Discriminants, Resultants and Multidimensional Determinants*. Basel, Birkhäuser.
Giusti, M., Heintz, J., Morais, J. E., Pardo, L. M. (1995). When polynomial equation systems can be "Solved" Fast ?. In *AAECC'95,* LNCS **948**, pp. 205–231. Springer Verlag.
Grigoryev, D. Y. (1986). Factorization of polynomials over finite field and the solution of systems of algebraic equations. *J. Sov. Math.*, **34**, 1762–1803.
Grigoryev, D. Y. (1988). The complexity of deciding Tarski algebra. *J. Symb. Comput.*, **5**, Special Issue on Decision Algorithms for the Theory of Real Closed Fields, pp. 65–108.
Grigoryev, D. Y., Vorobjov, N. N. (1988). Solving systems of polynomial inequalities in subexponential time. *J. Symb. Comput.*, **5,** Special Issue on Decision Algorithms for the Theory of Real Closed Fields, pp. 37–64.
Golub, G. H., Van Loan, C. F. (1996). *Matrix computations,* 3rd edn. Baltimore, MD, John Hopkins, University Press.
Gruber, P. M., Wills, J. M., eds, (1993). *Handbook for Convex Geometry*. Amsterdam, North Holland.
Grünbaum, B. (1967). *Convex Polytopes*. New York, Wiley-Interscience.
Habsieger, L. (1998). Sur un problème combinatoire. Communication personnelle.
Harris, J. (1992). *Algebraic Geometry, a first course,* volume 133, *Graduate Texts in Mathematics*. Springer Verlag.
Hodge, W., Pedoe, D. (1952). *Methods of Algebraic Geometry*. Cambridge, Cambridge University Press.
Hoffmann, C. M. (1989). *Geometric and Solid Modeling*. San Mateo, Morgan Kaufmann.
Huber, B., Sturmfels, B. (1995). A polyhedral method for solving sparse polynomial systems. *Math. Comput.*, **64**, 1542–1555.
Huber, B., Sturmfels, B. (March 1997). Bernstein's theorem in affine space. *Discr. Comput. Geom.*, **17**, 137–142.
Jouanolou, J. P. (1991). Le formalisme du résultant. *Adv. Math.*, **90**, 117–263.
Jouanolou, J. P. (1993a). Formes d'inertie et Résultants: Un formulaire. Prépublication de l'IRMA (Strasbourg).
Jouanolou, J. P. (1993b). Résultant anisotrope, compléments et appplications. Prépublication de l' IRMA (Strasbourg).
Kapranov, M. M., Sturmfels, B., Zelevinsky, A. V. (1992). Chow polytopes and general resultants. *Duke Math. J.*, **67**, 189–218.
Kapur, D., Lakshman, Y. N. (1992). In Donald, B., Kapur, D. and Mundy, J., eds, *Symbolic and Numerical Computation for Artifitial Intelligence,* pp. 45–89. New York, Academic Press.
Kapur, D., Saxena, T. (1997). Extraneous factors in the Dixon resultant formulation. In Küchlin, W. W., ed., *Proceedings of ISSAC*, pp. 141–148. New York, ACM Press.
Kapur, D., Saxena, T., Yang, L. (1994). Algebraic and geometric reasoning using Dixon resultants. In von zur Gathen, J., ed., *Proceedings of ISSAC, Oxford,* pp. 99–107, New York, ACM Press.
Kapur, D., Saxena, T., Yang, L. (1996). Sparsity Considerations in Dixon Resultants. In *Proceedings of STOC*, pp. 184–191. New York, ACM Press.
Khovanskii, A. G. (1978). Newton polyhedra and the genus of complete intersections. *Funktsional'nyi Analiz i Ego Prilozheniya*, **12**, 51–61.
Knuth, D. E. (1981). *The Art of Computer Programming: Seminumerical Algorithms,* volume 2. Reading, MA, Addison-Wesley.
Krishnan, S., Manocha, D. (1995). In Levelt, A. H. M., ed., *Proceedings of ISSAC, Montreal, Canada*. pp. 59–67. New York, ACM Press.
Krick, T., Pardo, L. M. (1996). A computational method for diophatine approximation. In González-Vega, L. and Recio, T., eds, *Algorithms in Algebraic Geometry and Applications*. volume 143, *Progress in Mathematics*, pp. 193–254. Basel, Birkhäuser.
Kunz, E. (1986). In Kähler differentials. *Advanced Lectures in Mathematics*, Friedr, Vieweg and Sohn.
Kushnirenko, A. G. (1976). Newton polytopes and the Bézout theorem. *Funktsional'nyi Analiz i Ego Prilozheniya*, **10**.
Lascoux, A. (1986). La résultante de deux polynômes. In *Séminaire M. P. Malliavin,* LNM **1202**. pp. 56–72.
Lazard, D. (1981). Résolution des systèmes d'équations algébriques. *Theor. Comput. Sci.*, **15**, 77–110.
Li, T. Y., Wang, T., Wang, X. (1996). Random product homotopy with minimal BKK bound. In Renegar, J., Shub, M. and Smale, S., eds, *The Mathematics of Numerical Analysis,* volume 32, *Lectures in Applied Mathematics*. New York, American Mathematical Society.

Lickteig, T., Meer, K. (1995). A note on testing the resultant. *J. Complexity*, **11**, 344–351.
Macaulay, F. S. (1902). Some formulae in elimination. *Proc. London Math. Soc.*, **1**, 3–27.
Macaulay, F. S. (1912). On the resolution of a given modular system into primary systems including some properties of Hilbert numbers. *Proc. London Math. Soc.*, 66–121.
Macaulay, F. S. (1916). *The Algebraic Theory of Modular Systems*. Cambridge, Cambridge University Press.
Manocha, D., Zhu, Y., Wright, W. (1995). Conformational analysis of molecular chains using nano-kinematics. *Comput. Appl. Biol. Sci.*, **11**, 71–86.
Manocha, D., Canny, J. (1992). Multipolynomial resultants and linear algebra. In *Proceedings of ISSAC*, pp. 96–102. New York, ACM Press.
Manocha, D., Demmel, J. (1995). Algorithms for intersecting parametric and algebraic curves II: multiple intersections. *Graph. Models Image Process*, **57**, 81–100.
Morgan, A. P., Sommese, A. J. (1987). A homotopy for solving general polynomial systems that respects m-homogeneous structures. *Appl. Math. Comput.*, **24**, 101–113.
Morgan, A. P., Sommese, A. J., Wampler, C. W. (1994). A product-decomposition bound for Bézout numbers. *SIAM J. Numer. Anal.*, **32**.
Mourrain, B. (1993). The 40 generic positions of a parallel robot. In Bronstein, M., ed., *Proceedings of ISSAC, Kiev, Ukraine*, pp. 173–182. New York, ACM Press.
Mourrain, B. (1996a). Enumeration problems in Geometry, Robotics and Vision. In González, L. and Recio, T., eds, *Algorithms in Algebraic Geometry andApplications*, volume 143, *Progress in Mathematics*, pp. 285–306. Basel, Birkhäuser.
Mourrain, B. (1996b). Isolated points, duality and residues. *J. Pure Appl. Algebra*, **117–118**, 469–493. Special Issue for the Proc. of the 4th Int. Symp. on Effective Methods in Algebraic Geometry (MEGA).
Mourrain, B., Pan, V. Y. (1997a). Solving special polynomial systems by using structured matrices and algebraic residues. In Cucker, F. and Shub, M., eds, *Foundations of Computational Mathematics, Rio de Janeiro*, pp. 287–304. Springer Verlag.
Mourrain, B., Pan, V. Y. (1997b). Multidimensional structured matrices and polynomial systems. *Calcolo, Special Issue for the workshop: Structure, Algorithms and Applications*, **33**, 389–401.
Mourrain, B. (1998). Computing isolated polynomial roots by matrix methods. *J. Symb. Comput.,* Special Issue on Symbolic-Numeric Algebra for Polynomials, **26**, 715–738.
Mourrain, B., Pan, V. Y. (1998). Asymptotic acceleration of solving multivariate polynomial systems of equations. In *Proceedings of STOC*, pp. 488–496. New York, ACM Press.
Mourrain, B., Pan, V. Y. (1999). Multivariate polynomials, duality and structured matrices. *J. Complexity*, **136**, 67–100.
Mourrain, B., Stolfi, N. (1995). Applications of Clifford algebra in robotics. In Merlet, J. P. and Ravani, B., eds, *Computational Kinematics'95, Solid Mechanics and its applications*, pp. 141–150. Sophia-Antipolis, Kluwer.
Muir, T. (1960). History of determinants. 5 volumes, Dover reprints.
Pedersen, P. *AMS–IMS–SIAM Summer Conference on Continuous Algorithms and Complexity, Mt. Holyoke, MA, July 1994. Lecture notes.*
Pedersen, P. S., Sturmfels, B. (1993). Product formulas for resultants and Chow forms. *Math. Zeitschrift*, **214**, 377–396.
Pedersen, P., Sturmfels, B. (1996). Mixed monomial bases. In González-Vega, L. and Recio, T., eds, *Effective Methods in Algebraic Geometry*, volume 143, *Progress in Mathematics*, pp. 307–316. Boston, Birkhäuser, (Proc. MEGA '94, Santander, Spain).
Raghavan, M., Roth, B. (June 1995). Solving polynomial systems for the kinematics analysis and synthesis of mechanisms and robot manipulators. *Trans. ASME, Special 50th Annivers. Design Issue*, **117**, 71–79.
Renegar, J. (1989). On the worst-case arithmetic complexity of approximating zeros of system of polynomials. *SIAM J. Computing*, **18**, 350–370.
Renegar, J. (1992). On the computational complexity and geometry of the first order theory of reals (I, II, III). *J. Symb. Comput.*, **13**, 255–352.
Rojas, J. M. (1994). A convex geometric approach to counting the roots of a polynomial system. *Theor. Comput. Sci.*, **133**, 105–140.
Rojas, J. M. (1999). Toric laminations, sparse generalized characteristic polynomials, and a refinement of Hilbert's tenth problem. In Cucker, F. and Shub, M., eds, *Proc. Workshop on Foundations ofComputational Mathematics*, pp. 369–381. Berlin, Springer.
Rojas, J. M. (1997). Toric intersection theory for affine root counting. *J. Pure Appl. Algebra*, to appear.
Rouillier, F. (1996). Algorithmes efficaces pour l'étude des zéros réels des systèmes polynomiaux. Ph.D. Thesis, Université de Rennes.
Sabia, J., Solerno, P. (1995). Bounds for traces in complete intersections and degrees in the Nullstellenstaz. *AAECC-6*, **948**, 353–376.

Scheja, G., Storch, U. (1975). Über Spurfunktionen bei vollständigen Durschnitten. *J. Reine Ang. Math.*, **278**, 174–190.

Schneider, R. (1993). *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge, Cambridge University Press.

Shafarevitch, I. R. (1974). *Basic Algebraic Geometry*. Berlin, Springer Verlag.

Stetter, H. J. (1996). Eigenproblems are at the heart of polynomial system solving. *SIGSAM Bull.*, **30**, 22–25.

Sturmfels, B. (1994). On the Newton polytope of the resultant. *J. Algebr. Combin.*, **3**, 207–236.

Sturmfels, B., Zelevinsky, A. (1994). Multigraded resultants of Sylvester type. *J. Algebra*, **163**, 115–127.

Sylvester, J. J. (1853). On a theory of syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraic common measure. *Phil. Trans.*, **143**, 407–548.

van der Waerden, B. L. (1950). *Modern Algebra,* 3rd edn. New York, F. Ungar Publishing Co..

Verschelde, J., Verlinden, P., Cools, R. (1994). Homotopies exploiting Newton polytopes for solving sparse polynomial systems. *SIAM J. Numer. Anal.*, **31**, 915–930.

Verschelde, J., Gatermann, K., Cools, R. (1996). Mixed volume computation by dynamic lifting applied to polynomial system solving. *Discrete and Comput. Geom.*, **16**, 69–112.

Wiedemann, D. H. (1986). Solving sparse linear equations over finite fields. *IEEE Trans. Inf. Theory*, **32**, 54–62.

Zippel, R. (1993). *Effective Polynomial Computation*. Boston, Kluwer.