# INTRODUCTION TO RESULTANTS

Bernd Sturmfels

Department of Mathematics
University of California
Berkeley, CA 94720; U.S.A
bernd@math.berkeley.edu

Abstract. These are notes for my lecture at the AMS short course on *Applications of Computational Algebraic Geometry*, San Diego, January 6–7, 1997. We consider the problem of eliminating $k$ variables from a system of $k$ polynomial equations. This can be done using lexicographic Gröbner bases, but it is more efficient to use classical resultant techniques. We present an introduction to the Sylvester resultant (for $k = 1$) and the multivariate resultant (for $k \geq 2$). In the last section we discuss sparse resultants and we give a guide to the current literature on elimination theory.

## 1. Why classical elimination theory ?

Elimination theory deals with the problem of eliminating one or more variables from a system of polynomial equations, thus reducing the given problem to a smaller problem in fewer variables. For instance, if we are interested in solving

$$a_0 + a_1 x + a_2 x^2 \quad = \quad b_0 + b_1 x + b_2 x^2 \quad = \quad 0, \tag{1.1}$$

where $a_2 \neq 0$ and $b_1 \neq 0$, then we can eliminate the variable $x$ to get the expression

$$a_0^2 b_2^2 - a_0 a_1 b_1 b_2 - 2 a_0 a_2 b_0 b_2 + a_0 a_2 b_1^2 + a_1^2 b_0 b_2 - a_1 a_2 b_0 b_1 + a_2^2 b_0^2. \tag{1.2}$$

This expression is called the *Sylvester resultant* of (1.1). It vanishes if and only if the polynomials in (1.1) have a common complex root $x$. This assertion can be derived from the Elimination Theorem in [Cox, §4] as follows. Consider a polynomial ring in seven variables which are ordered $x > a_0 > a_1 > a_2 > b_0 > b_1 > b_2$ and fix the lexicographic term order. Compute the reduced Gröbner basis $\mathcal{G}$ for the ideal $\langle a_0 + a_1 x + a_2 x^2,\ b_0 + b_1 x + b_2 x^2 \rangle$. Then $\mathcal{G}$ consists of six elements, and the unique element in $\mathcal{G} \cap \mathbb{Q}[a_0, a_1, a_2, b_0, b_1 . b_2]$ is the Sylvester resultant (1.2).

We may apply the same procedure to eliminate $x$ from two general equations:

$$\begin{aligned} f &= a_0 + a_1 x + a_2 x^2 + \cdots + a_{m-1} x^{m-1} + a_m x^n, \\ g &= b_0 + b_1 x + b_2 x^2 + \cdots + b_{n-1} x^{n-1} + b_n x^n. \end{aligned} \tag{1.3}$$

1991 *Mathematics Subject Classification.* Primary 13P99, 14Q99; Secondary 68Q40.

The author is supported in part by an NSF National Young Investigator Fellowship and a David and Lucile Packard Fellowship.

Introduce a polynomial ring in $m+n+3$ variables $x > a_0 > a_1 > \cdots > a_m > b_0 > b_1 > \cdots > b_n$, fix the lexicographic term order, and compute the reduced Gröbner basis $G_{m,n}$ for the two equations in (1.3). Then $G_{m,n}$ contains a unique polynomial $R_{m,n}$ which is free of $x$. It has total degree $m+n$ and is called the *resultant* of $f$ and $g$. We encourage the reader to do this experiment for some small values of $m$ and $n$ using a fast Gröbner basis program (such as COCOA, MACAULAY or SINGULAR). For instance, in the case $m = 3, n = 4$, the Gröbner basis $G_{3,4}$ has 41 elements, one of which is $R_{3,4}$. Most elements of $G_{3,4}$ look like a mess, don't they ? Now try the same calculation for the next case $m = n = 4$. What happens ?

A central goal of classical elimination theory is to find explicit formulas for resultants, where the coefficients of the input polynomials are regarded as indeterminates. In what follows we present three examples of such formulas.

Our first example is the expression (1.2). It equals the following determinant:

$$R_{2,2} \quad = \quad det \begin{pmatrix} a_0 & 0 & b_0 & 0 \\ a_1 & a_0 & b_1 & b_0 \\ a_2 & a_1 & b_2 & b_1 \\ 0 & a_2 & 0 & b_2 \end{pmatrix}. \tag{1.4}$$

This is an instance of the *Sylvester formula*, which is presented in (2.1) below.

Our second example is a special case of Theorem 2.2 below. Let $m = n = 4$. Then *Bézout's formula* for the resultant of two univariate quartics states

$$R_{4,4} \quad = \quad det \begin{pmatrix} [04] & [03] & [02] & [01] \\ [14] & [04]+[13] & [03]+[12] & [02] \\ [24] & [14]+[23] & [04]+[13] & [03] \\ [34] & [24] & [14] & [04] \end{pmatrix}, \tag{1.5}$$

where $[ij] := a_i b_j - a_j b_i$. It is easy to generate the matrix (1.5) using MAPLE: the command `bezout` will do this. For a comparison try also the command `sylvester`.

Our third and last example is more sophisticated. We consider the problem of eliminating two variables $x$ and $y$ from a system of three quadratic equations

$$\begin{aligned} f &= a_0x^2 + a_1xy + a_2y^2 + a_3x + a_4y + a_5, \\ g &= b_0x^2 + b_1xy + b_2y^2 + b_3x + b_4y + b_5, \\ h &= c_0x^2 + c_1xy + c_2y^2 + c_3x + c_4y + c_5, \end{aligned} \tag{1.6}$$

We begin by homogenizing the given equations. Instead of (1.6) we consider

$$\begin{aligned} F &= a_0x^2 + a_1xy + a_2y^2 + a_3xz + a_4yz + a_5z^2, \\ G &= b_0x^2 + b_1xy + b_2y^2 + b_3xz + b_4yz + b_5z^2, \\ H &= c_0x^2 + c_1xy + c_2y^2 + c_3xz + c_4yz + c_5z^2, \end{aligned} \tag{1.6'}$$

where $z$ is a new variable. Next we compute the *Jacobian determinant*

$$J \quad := \quad det \begin{pmatrix} \partial F/\partial x & \partial F/\partial y & \partial F/\partial z \\ \partial G/\partial x & \partial G/\partial y & \partial G/\partial z \\ \partial H/\partial x & \partial H/\partial y & \partial H/\partial z \end{pmatrix},$$

The partial derivatives of $J$ can be written as quadratic polynomials in $x, y, z$:

$$\begin{aligned} \partial J/\partial x &= u_0 x^2 + u_1 xy + u_2 y^2 + u_3 xz + u_4 yz + u_5 z^2, \\ \partial J/\partial y &= v_0 x^2 + v_1 xy + v_2 y^2 + v_3 xz + v_4 yz + v_5 z^2, \\ \partial J/\partial z &= w_0 x^2 + w_1 xy + w_2 y^2 + w_3 xz + w_4 yz + w_5 z^2. \end{aligned}$$

Note that each of $u_i$, $v_j$, $w_k$ is a polynomial of degree 3 in the coefficients $a_i, b_j, c_k$ of the input equations (1.6). We now compute the following $6 \times 6$-determinant:

$$\mathcal{R} \quad = \quad det \begin{pmatrix} a_0 & b_0 & c_0 & u_0 & v_0 & w_0 \\ a_1 & b_1 & c_1 & u_1 & v_1 & w_1 \\ a_2 & b_2 & c_2 & u_2 & v_2 & w_2 \\ a_3 & b_3 & c_3 & u_3 & v_3 & w_3 \\ a_4 & b_4 & c_4 & u_4 & v_4 & w_4 \\ a_5 & b_5 & c_5 & u_5 & v_5 & w_5 \end{pmatrix} \tag{1.7}$$

This is a homogeneous polynomial of degree 12 in the 18 indeterminates $a_0, a_1, \ldots, a_5, b_0, b_1, \ldots, b_5, c_0, c_1, \ldots, c_5$. The polynomial $\mathcal{R}$ is called the *resultant* of the system (1.6). The full expansion of $\mathcal{R}$ has 21,894 terms. (Please check !)

The resultant $\mathcal{R}$ vanishes if and only if the equations (1.6') have a common solution $(x, y, z)$, where $x, y, z$ are complex numbers not all zero. It is possible, at least in principle, to find $\mathcal{R}$ by computing a lexicographic Gröbner basis of (1.6) with respect to 20 variables. However, this calculation is infeasible in practice.

Our experiments suggest that Gröbner bases are inefficient when it comes to computing resultants. The reason is that the output contains too many unwanted elements. As a result the Buchberger algorithm becomes hopelessly slow. This happens frequently if one is interested in eliminating $k$ indeterminates from a set of $k$ polynomials in more than $k$ variables. This task is handled best by classical elimination theory. In his excellent recent text book David Eisenbud writes that *the most effective current algorithms (for elimination theory) do not follow the older methods, but are based on Gröbner bases* [Eis, §14, p. 306]. I respectfully disagree with this view. In my opinion the old-fashioned approach is well, alive and effective.

For a comprehensive treatment of classical elimination theory from a modern perspective we recommend the book by Gel'fand, Kapranov and Zelevinsky [GKZ2]. The role of resultants and their matrices in (numerical methods for) solving polynomial equations will be discussed in the lecture of D. Manocha [Man, §3].

## 2. The Sylvester Resultant

Let $f$ and $g$ be two polynomials as in (1.3) whose coefficients are rational numbers with $a_m \neq 0$ and $b_n \neq 0$. The *Sylvester resultant* of $f$ and $g$ is the $(m+n) \times (m+n)$-

determinant

$$(2.1) \qquad Res_x(f,g) \;=\; det \begin{pmatrix} a_0 & & & & b_0 & & & \\ a_1 & a_0 & & & b_1 & b_0 & & \\ & a_1 & \ddots & & & b_1 & \ddots & \\ \vdots & & \ddots & a_0 & \vdots & & \ddots & b_0 \\ & \vdots & & a_1 & & \vdots & & b_1 \\ a_m & & & & b_n & & & \\ & a_m & & \vdots & & b_n & & \vdots \\ & & \ddots & & & & \ddots & \\ & & & a_m & & & & b_n \end{pmatrix},$$

where the blank spaces are filled with zeroes.

**Proposition 2.1.** *Three basic properties of the Sylvester resultant are*
*(a)* $Res_x(f,g)$ *is a polynomial in* $a_0,\ldots,a_m,b_0,\ldots,b_n$ *with integer coefficients;*
*(b)* $Res_x(f,g) = 0$ *if and only if* $f(x)$ *and* $g(x)$ *have a common factor in* $\mathbb{Q}[x]$*;*
*(c) There exist polynomials* $A, B \in \mathbb{Q}[x]$ *such that* $Af + Bg = Res_x(f,g)$.

This proposition is proved in [CLO1, §3.5]. There are many useful formulas for the Sylvester resultant. For instance, suppose that the roots of $f$ are $\xi_1,\ldots,\xi_m$ and the roots of $g$ are $\eta_1,\ldots,\eta_n$. Then we have the following product formulas:

$$Res_x(f,g) \;=\; a_0^n b_0^m \prod_{i=1}^{m}\prod_{j=1}^{n}(\xi_i - \eta_j) \;=\; a_0^n \prod_{i=1}^{m} g(\xi_i) \;=\; (-1)^{mn} b_0^m \prod_{j=1}^{n} f(\eta_j).$$

A completely explicit description of all terms appearing in the expansion of the Sylvester resultant is given in [GKZ1].

If $f$ and $g$ are polynomials of the same degree $m = n$, then one can evaluate their resultant as follows. Consider the following polynomial in two variables

$$\mathcal{B}(x,y) \;=\; \frac{f(x)g(y) - f(y)g(x)}{x-y} \;=\; \sum_{i,j=0}^{m-1} c_{ij} x^i y^j.$$

Form the symmetric $m \times m$-matrix $C = (c_{ij})$. Its entries $c_{ij}$ are quadratic polynomials in the coefficients $a_0,\ldots,a_m,b_0,\ldots,b_n$. The case $m = 4$ appears in (1.5).

**Theorem 2.2.** (Bézout resultant) *The determinant of* $C$ *equals* $\pm Res_x(f,g)$.

Proof. The resultant $Res_x(f,g)$ is an irreducible polynomial of degree $2m$ in $a_0,\ldots,a_m,b_0,\ldots,b_n$. The irreducibility property will be proved more generally in Theorem 3.1 (b) below. The determinant of $C$ is also a polynomial of degree $2m$. We will show that the zero set of $Res_x(f,g)$ is contained in the zero set of $det(C)$. This implies that the two polynomials are equal up to a constant. Looking at leading terms one finds the constant to be either 1 or $-1$.

If $(a_0, \ldots, a_m, b_0, \ldots, b_n)$ is in the zero set of $Res_x(f,g)$ then the corresponding polynomials (1.3) have a common complex root $x_0$. Then $\mathcal{B}(x_0, y)$ is identically zero as a polynomial in $y$. This implies that the non-zero complex vector $(1, x_0, x_0^2, \ldots, x_0^{m-1})$ lies in the kernel of $C$, and therefore $det(C) = 0$. $\square$

Practical experience shows that the Bézout matrix is superior to the Sylvester matrix with regard to numerical stability. So, if one has two polynomials of the same (large) degree with numerical coefficients, then it is best to use the Bézout matrix. Evidence supporting this view is given in [Sed, §3.3] and in [GMS, §2.2].

Computing the resultant of two polynomials with respect to a specified variable $x$ is a standard feature in many computer algebra systems. These implementations make use of either the Sylvester formula or the Bézout formula, and they work quite fast. For instance, in MAPLE one uses the command `resultant(f,g,x)`. We present three concrete examples of how to make use of this feature.

**Example 2.3.** *(Intersecting two algebraic curves in the real plane)*
Consider two polynomials in two variables, say,

$$\begin{aligned} f &= x^4 + y^4 - 1 \\ g &= x^5y^2 - 4x^3y^3 + x^2y^5 - 1. \end{aligned} \tag{2.2}$$

They define two curves $\{f = 0\}$ and $\{g = 0\}$ in the real plane $\mathbb{R}^2$. We wish to compute their intersection, that is, all points $(x, y) \in \mathbb{R}^2$ with $f(x, y) = g(x, y) = 0$. To this end we evaluate the resultant with respect to one of the variables,

$$\begin{aligned} Res_x(f,g) &= 2y^{28} - 16y^{27} + 32y^{26} + 249y^{24} + 48y^{23} - 128y^{22} + 4y^{21} - 757y^{20} \\ &\quad - 112y^{19} + 192y^{18} - 12y^{17} + 758y^{16} + 144y^{15} - 126y^{14} \\ &\quad + 28y^{13} - 251y^{12} - 64y^{11} + 30y^{10} - 36y^9 - y^8 + 16y^5 + 1. \end{aligned}$$

This is an irreducible polynomial in $\mathbb{Q}[y]$. It has precisely four real roots

$$y \sim -0.924209668, \quad y \sim -0.597428987, \quad y \sim 0.721113386, \quad y \sim 0.966506297.$$

Hence the two curves have four intersection points, having these $y$-coordinates. By symmetry in (2.2), the same values are also the possible $x$-coordinates. By trying out (numerically) all 16 conceivable $x$-$y$-combinations, we find that the following four pairs are the real solutions to our given equations $f(x, y) = g(x, y) = 0$:

$$\begin{aligned} &(x, y) \sim (-0.924209668, 0.721113386), \quad (x, y) \sim (0.721113386, -0.924209668), \\ &(x, y) \sim (-0.597428987, 0.966506297), \quad (x, y) \sim (0.966506297, -0.597428987). \end{aligned}$$

**Example 2.4.** *(Implicitization of a rational curve in the plane)*
Consider a plane curve which is given to us parametrically:

$$\mathcal{C} \quad = \quad \left\{ \left(\frac{a(t)}{b(t)}, \frac{c(t)}{d(t)}\right) \in \mathbb{R}^2 \quad : \quad t \in \mathbb{R} \right\},$$

where $a(t), b(t), c(t), d(t)$ are polynomials in $\mathbb{Q}[t]$. The goal is to find the defining equation of $\mathcal{C}$, that is, the unique irreducible polynomial $f \in \mathbb{Q}[x,y]$ which vanishes on $\mathcal{C}$. We may find $f$ by the general Gröbner basis approach explained in [CLO1, pp. 97]. It is more efficient, however, to use the following explicit formula:

$$f(x,y) \quad = \quad Res_t\big(\, b(t)\cdot x - a(t),\; d(t)\cdot y - c(t)\,\big).$$

Here is a numerical example:
If $a(t) = t^3-1$, $b(t) = t^2-5$, $c(t) = t^4-3$ and $d(t) = t^3-7$ then

$$\begin{aligned} f(x,y) \quad = \quad & 76\,x^4y^2 - 76\,x^3y^3 - 308\,x^4y + 758\,x^3y^2 - 450\,x^2y^3 - 484\,x^4 \\ & - 226\,x^3y + 870\,x^2y^2 - 540\,xy^3 + 440\,x^3 - 704\,x^2y + 378\,xy^2 \\ & - 216\,y^3 + 36\,x^2 + 18\,xy + 324\,y^2 - 16\,x - 162\,y + 26. \end{aligned}$$

See the lecture of T. Sederberg [Sed, Section 3] for details and applications. □

**Example 2.5.** *(Computation with algebraic numbers)*
Let $\alpha$ and $\beta$ be algebraic numbers over $\mathbb{Q}$. They are represented by their *minimal polynomials* $f, g \in \mathbb{Q}[x]$. These are the unique (up to scaling) irreducible polynomials satisfying $f(\alpha) = 0$ and $g(\beta) = 0$. Our problem is to find the minimal polynomials $p$ and $q$ for their sum $\alpha+\beta$ and their product $\alpha\cdot\beta$ respectively. The answer is given by the following two formulas involving the Sylvester resultant:

$$p(z) \;=\; Res_x\big(\, f(x),\, g(z-x)\,\big) \quad \text{and} \quad q(z) \;=\; Res_x\big(\, f(x),\, g(z/x)\cdot x^{deg(g)}\,\big).$$

It is easy to check the identities $p(\alpha+\beta) = 0$ and $q(\alpha\cdot\beta) = 0$.

As an example we consider two algebraic numbers given in terms of radicals:

$$\alpha \;=\; \sqrt[5]{2}, \qquad \beta \;=\; \sqrt[3]{-7/2 - 1/18\sqrt{3981}} \;+\; \sqrt[3]{-7/2 + 1/18\sqrt{3981}}.$$

Their minimal polynomials are $\alpha^5 - 2$ and $\beta^3+\beta+7$ respectively. Using the above formulas, we find that the minimal polynomial for their sum $\alpha+\beta$ equals

$$\begin{aligned} p(z) \quad = \quad & z^{15} + 5\,z^{13} + 35\,z^{12} + 10\,z^{11} + 134\,z^{10} + 500\,z^9 + 240\,z^8 + 2735\,z^7 \\ & + 3530\,z^6 + 1273\,z^5 - 6355\,z^4 + 12695\,z^3 + 1320\,z^2 + 22405\,z + 16167, \end{aligned}$$

and the minimal polynomial for their product $\alpha\cdot\beta$ equals

$$q(z) \quad = \quad z^{15} - 70\,z^{10} + 984\,z^5 + 134456.$$

## 3. The classical multivariate resultant

Consider a system of $n$ homogeneous polynomials in $n$ indeterminates

$$f_1(x_1,\dots,x_n) \quad = \quad \cdots \quad = \quad f_n(x_1,\dots,x_n) \quad = \quad 0. \tag{3.1}$$

We assume that the $i$-th equation is homogeneous of degree $d_i > 0$, that is,

$$f_i \quad = \quad \sum_{j_1+\cdots+j_n=d_i} c^{(i)}_{j_1,\dots,j_n}\, x_1^{j_1}\cdots x_n^{j_n}, \tag{3.2}$$

where the sum is over all $\binom{n+d_i-1}{d_i}$ monomials of degree $d_i$ in $x_1,\dots,x_n$. Note that the zero vector $(0,0,\dots,0)$ is always a solution of (3.1). Our basic question is to determine under which condition there is a non-zero solution.

As a simple example let us consider the special case $n = 3, d_1 = d_2 = d_3 = 1$:

$$\begin{aligned} f_1 &= c^1_{100}x_1 + c^1_{010}x_2 + c^1_{001}x_3 \\ f_2 &= c^2_{100}x_1 + c^2_{010}x_2 + c^2_{001}x_3 \\ f_3 &= c^3_{100}x_1 + c^3_{010}x_2 + c^3_{001}x_3 \end{aligned}$$

This linear system has a non-zero solution if and only if the determinant is zero:

$$det\begin{pmatrix} c^1_{100} & c^1_{010} & c^1_{001} \\ c^2_{100} & c^2_{010} & c^2_{001} \\ c^3_{100} & c^3_{010} & c^3_{001} \end{pmatrix} \quad = \quad 0. \tag{3.3}$$

The (non-trivial) case $n = 3, d_1 = d_2 = d_3 = 2$ is discussed in the end of Section 1.

Returning to the general case we introduce a variable $u^{(i)}_{j_1,\dots,j_n}$ for each coefficient $c^{(i)}_{j_1,\dots,j_n}$ of $f_i$. Let $\mathbb{Z}[u]$ be the ring of polynomials with integer coefficients in these variables. The total number of variables in $\mathbb{Z}[u]$ equals $N = \sum_{i=1}^n \binom{n+d_i-1}{d_i}$. Given a polynomial $P$ in $\mathbb{Z}[u]$ we let $P(f_1,\dots,f_n)$ denote the number obtained by replacing each variable $u^{(i)}_{j_1,\dots,j_n}$ by the corresponding coefficient $c^{(i)}_{j_1,\dots,j_n}$. For instance, the $3\times 3$-determinant in the example above may be regarded as a cubic polynomial $\det \in \mathbb{Z}[u]$, and we write $det(f_1,f_2,f_3)$ for the specific value on the left hand side of (3.3). We can now answer our basic question.

**Theorem 3.1.** *If we fix positive degrees $d_1,\dots,d_n$, then there exists a unique (up to sign) irreducible polynomial $Res \in \mathbb{Z}[u]$ which has the following properties:*

*(a) If $f_1,\dots,f_n \in \mathbb{C}[x_1,\dots,x_n]$ are homogeneous of degrees $d_1,\dots,d_n$ then the equations (3.1) have a non-zero solution in $\mathbb{C}^n$ if and only if $Res(f_1,\dots,f_n) = 0$.*

*(b) $Res$ is irreducible, even when regarded as a polynomial in $\mathbb{C}[u]$.*

The polynomial $Res$ is called the *(multivariate) resultant* of the system (3.1). In the end of Section 1 we have given a recipe for computing the resultant for $n = 3, d_1 = d_2 = d_3 = 2$. It can be extended to the case $n = 3, d_1 = d_2 = d_3 > 2$; this is carried out in [GKZ2, Section 3.4.D]. It is a non-trivial matter to prove the correctness of this classical recipe; see the comment after [GKZ2, Theorem 3.4.10].

We shall present a proof of Theorem 3.1, which assumes some familiarity with basic notions in algebraic geometry. It is independent of the subsequent material and may be skipped. The advantage of the geometric proof technique is that it is widely applicable to many problems in elimination theory.

*Proof.* The elements of $\mathbb{C}[u]$ are polynomial functions on the affine space $\mathbb{C}^N$. We regard $x = (x_1, \ldots, x_n)$ as homogeneous coordinates for the complex projective space $\mathbb{P}^{n-1}$. Thus $(u, x)$ are the coordinates on the product variety $\mathbb{C}^N \times \mathbb{P}^{n-1}$.

Let $\mathcal{I}$ denote the subvariety of $\mathbb{C}^N \times \mathbb{P}^{n-1}$ defined by the equations

$$\sum_{j_1+\cdots+j_n=d_i} u^{(i)}_{j_1,\ldots,j_n}\, x_1^{j_1} \cdots x_n^{j_n} \quad = \quad 0 \qquad \text{for } \ i = 1, 2, \ldots, n.$$

The set $\mathcal{I}$ is called the *incidence variety*. Note that $\mathcal{I}$ is defined over $\mathbb{Q}$. Consider the projection $\phi : \mathbb{C}^N \times \mathbb{P}^{n-1} \to \mathbb{P}^{n-1}$, $(u, x) \mapsto x$. Then $\phi(\mathcal{I}) = \mathbb{P}^{n-1}$. More precisely, the preimage $\phi^{-1}(x)$ of any point $x \in \mathbb{P}^{n-1}$ can be identified with the set $\{\, u \in \mathbb{C}^N \,:\, (u, x) \in \mathcal{I} \,\}$. This is a linear subspace of codimension $n$ in $\mathbb{C}^N$. To this situation we apply Theorem 8 in Section I.6.3 of [Sha]. It shows that $\mathcal{I}$ is a closed and irreducible subvariety of codimension $n$ in $\mathbb{C}^N \times \mathbb{P}^{n-1}$. Hence $dim(\mathcal{I}) = N - 1$.

Consider the projection $\psi : \mathbb{C}^N \times \mathbb{P}^{n-1} \to \mathbb{C}^N$, $(u, x) \mapsto u$. It follows from the *Main Theorem of Elimination Theory* [Eis, Thm. 14.1] that $\psi(\mathcal{I})$ is an irreducible subvariety of $\mathbb{C}^N$ which is defined over $\mathbb{Q}$ as well. Every point $c$ in $\mathbb{C}^N$ can be identified with a polynomial system $(f_1, \ldots, f_n)$ as in (3.1). That system has a nonzero root if and only if $c$ lies in the subvariety $\psi(\mathcal{I})$. For every such $c$ we have

$$(3.4) \quad dim(\psi(\mathcal{I})) \quad \leq \quad dim(\mathcal{I}) \quad = \quad N - 1 \quad \leq \quad dim(\psi^{-1}(c)) \,+\, dim(\psi(\mathcal{I})).$$

The two inequalities in (3.4) follow respectively from parts (2) and (1) of Theorem 7 in Section I.6.3 of [Sha]. We now choose $c = (f_1, \ldots, f_n)$ as follows. Let $f_1, \ldots, f_{n-1}$ be any equations as in (3.2) which have only finitely many zeros in $\mathbb{P}^{n-1}$. Then choose $f_n$ which vanishes at exactly one of these zeros, say $y \in \mathbb{P}^{n-1}$. Hence $\psi^{-1}(c) = \{(c, y)\}$, a zero-dimensional variety. For this particular choice of $c$ both inequalities in (3.4) hold with equality. This implies $dim(\psi(\mathcal{I})) = N - 1$.

We have shown that the image of $\mathcal{I}$ under $\psi$ is an irreducible hypersurface in $\mathbb{C}^N$, which is defined over $\mathbb{Z}$. Hence there exists an irreducible polynomial $Res \in \mathbb{Z}[u]$, unique up to sign, whose zero set equals $\psi(\mathcal{I})$. By construction, this polynomial $Res(u)$ satisfies properties (a) and (b) of Theorem 3.1. $\square$

We next give the degree of the resultant. If any of the equations $f_i$ in the system (3.1) is replaced by a scalar multiple $\lambda f_i$ then the solvability of (3.1) remains unchanged. This implies that, for each fixed $i$, the resultant $Res(u)$ is homogeneous in the coefficients $(u^{(i)}_a \,:\, |a| = d_i)$ of the $i$-th input equation $f_i$.

**Theorem 3.2.** *The resultant Res is homogeneous of degree* $d_1 \cdots d_{i-1} d_{i+1} \cdots d_n$ *in the coefficients* $(u^{(i)}_a \,:\, |a| = d_i)$ *for each fixed* $i \in \{1, \ldots, n\}$.

This theorem is derived from Bézout's Theorem [Sha, §IV.2.1] and the product formula [CLO2]. We omit the proof. For more general versions of Theorem 3.2 and of the product formula we refer to [Jo1] and to Theorem 4.2 below.

We next discuss a general approach for computing the resultant. In what follows we abbreviate a monomial $x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$ by $x^a$, and we write $|a| = a_1+a_2+\cdots+a_n$ for its degree. Let $f_1,\ldots,f_n$ be general forms of degrees $d_1,\ldots,d_n$. We call the integer $d := d_1+d_2+\cdots+d_n-n+1$ the *critical degree*. This is the smallest integer which has the following property: Every monomial $x^a$ of degree $d=|a|$ is divisible by $x_i^{d_i}$ for at least one index $i$. This suggests the following definition:

$$S_i \quad := \quad \{\, x^a \,:\, |a| = d \text{ and } i \text{ is the smallest index such that } x_i^{d_i} \text{ divides } x^a \,\}.$$

We obtain a partition of the set of monomials of critical degree into disjoint subsets:

$$S_1 \cup S_2 \cup \cdots \cup S_n \quad = \quad \{\, x^a \,:\, |a| = d \,\}.$$

For instance, $S_1$ is the set of all monomials $x^a$ with $|a|=d$ and $a_1 \geq d_1$, while $S_n$ is the set of all monomials $x_1^{a_1}\cdots x_{n-1}^{a_{n-1}}x_n^{d-a_1-\cdots-a_{n-1}}$ where $0 \leq a_1 < d_1,\ldots,0 \leq a_{n-1} < d_{n-1}$. It is easy to count these monomials:

**Lemma 3.3.** *The cardinality of $S_n$ equals $d_1 d_2 \cdots d_{n-1}$.*

We now consider the following system of $\binom{d+n-1}{d}$ equations:

$$\begin{aligned}
\frac{x^a}{x_1^{d_1}}\cdot f_1 \quad &= \quad 0 \qquad \text{for all } x^a \in S_1 \\
\vdots \qquad &\vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \qquad . \\
\frac{x^a}{x_n^{d_n}}\cdot f_n \quad &= \quad 0 \qquad \text{for all } x^a \in S_n
\end{aligned} \tag{3.5}$$

Each of these equations can be written as a linear combination of the $\binom{d+n-1}{d}$ monomials of degree $d$. The coefficients of these linear combinations form a square matrix of size $\binom{d+n-1}{d} \times \binom{d+n-1}{d}$. Each non-zero entry of this matrix is one of the indeterminates $u^{(i)}_{j_1,\ldots,j_n}$. Let $D_n$ denote the determinant of this square matrix.

**Theorem 3.4.** *The determinant $D_n$ is a non-zero polynomial in $\mathbb{Z}[u]$, homogeneous of degree $|S_i|$ in the coefficients of $f_i$, and the resultant Res divides $D_n$.*

*Proof.* Clearly, $D_n$ is a polynomial in $\mathbb{Z}[u]$. If we replace $f_i$ by $x_i^{d_i}$ then the above $\binom{d+n-1}{d} \times \binom{d+n-1}{d}$-matrix is the identity matrix. Therefore $D_n(x_1^{d_1},\ldots,x_n^{d_n}) = 1$, and, in particular, $D_n$ is not the zero polynomial. It follows immediately from the construction that $D_n$ is homogeneous of degree $|S_i|$ in the coefficients of $f_i$.

We next show that *Res* divides $D_n$. Any zero $u$ of *Res* corresponds to a system $f_1(y) = \cdots = f_n(y) = 0$ which has a solution $y \in \mathbb{P}^{n-1}$. Substituting $y$ for $x$ in (3.5), and assuming $u \neq 0$, we see that the coefficient matrix of (3.5) has a non-zero vector in its kernel. This implies $D_n(u) = 0$. Therefore $D_n$ vanishes whenever *Res* vanishes. Since *Res* is irreducible, we conclude that *Res* divides $D_n$. $\square$

**Example 3.5.** We illustrate this construction of the resultant for the case $n = 3, d_1 = d_2 = d_3 = 2$ using the notation in (1.6'). Here the critical degree is $d = 4$. There are fifteen monomials of critical degree. They are partitioned into three sets:

$$S_1 = \{x^3z, x^2z^2, x^2y^2, x^2yz, x^4, x^3y\}, \quad S_2 = \{y^3z, y^4, y^2z^2, xy^3, xy^2z\},$$
$$S_3 = \{yz^3, xyz^2, xz^3, z^4\}.$$

The equations (3.5) translate into the following $15 \times 15$-matrix:

$$\begin{array}{c} \\ xzF \\ z^2F \\ y^2F \\ yzF \\ x^2F \\ xyF \\ yzG \\ y^2G \\ z^2G \\ xyG \\ xzG \\ yzH \\ xyH \\ xzH \\ z^2H \end{array} \begin{array}{c} \begin{array}{ccccccccccccccc} x^4 & x^3y & x^3z & x^2y^2 & x^2yz & x^2z^2 & xy^3 & xy^2z & xyz^2 & xz^3 & y^4 & y^3z & y^2z^2 & yz^3 & z^4 \end{array} \\ \begin{pmatrix} 0 & 0 & a_0 & 0 & a_1 & a_2 & 0 & a_3 & a_4 & a_5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_0 & 0 & 0 & a_1 & a_2 & 0 & 0 & a_3 & a_4 & a_5 \\ 0 & 0 & 0 & a_0 & 0 & 0 & a_1 & a_2 & 0 & 0 & a_3 & a_4 & a_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_0 & 0 & 0 & a_1 & a_2 & 0 & 0 & a_3 & a_4 & a_5 & 0 \\ a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_0 & 0 & a_1 & a_2 & 0 & a_3 & a_4 & a_5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_0 & 0 & 0 & b_1 & b_2 & 0 & 0 & b_3 & b_4 & b_5 & 0 \\ 0 & 0 & 0 & b_0 & 0 & 0 & b_1 & b_2 & 0 & 0 & b_3 & b_4 & b_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & b_0 & 0 & 0 & b_1 & b_2 & 0 & 0 & b_3 & b_4 & b_5 \\ 0 & b_0 & 0 & b_1 & b_2 & 0 & b_3 & b_4 & b_5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_0 & 0 & b_1 & b_2 & 0 & b_3 & b_4 & b_5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & c_0 & 0 & 0 & c_1 & c_2 & 0 & 0 & c_3 & c_4 & c_5 & 0 \\ 0 & c_0 & 0 & c_1 & c_2 & 0 & c_3 & c_4 & c_5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_0 & 0 & c_1 & c_2 & 0 & c_3 & c_4 & c_5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & c_0 & 0 & 0 & c_1 & c_2 & 0 & 0 & c_3 & c_4 & c_5 \end{pmatrix} \end{array}$$

The columns are indexed by all monomials of critical degree, and the rows are indexed by multiples of the input equations $F, G, H$. The determinant of this matrix is equal to $a_0(a_0b_3 - a_3b_0)$ times the resultant.

Verifying the claim in the previous sentence is a highly recommended exercise. This exercise is non-trivial since your computer algebra system may crash when you try to expand this $15 \times 15$-determinant symbolically. □

Naturally, the construction in Theorem 3.4 makes sense for any ordering of the variables. In particular, any variable $x_i$ can play the role of the last variable $x_n$. For any $i$ choose an ordering which has $x_i$ as the last variable and let $D_i$ denote the corresponding determinant.

**Corollary 3.6.** *The resultant Res is the greatest common divisor of* $D_1, \ldots, D_n$.

*Proof.* Let $G$ be the greatest common divisor of $D_1, \ldots, D_n$. By Lemma 3.3 the degree of $D_i$ in the coefficients of $f_i$ equals $d_1 \cdots d_{i-1} d_{i+1} \cdots d_n$. Therefore $G$ has degree at most $d_1 \cdots d_{i-1} d_{i+1} \cdots d_n$ in the coefficients of $f_i$, for each $i \in \{1, \ldots, n\}$. On the other hand, *Res* has degree exactly $d_1 \cdots d_{i-1} d_{i+1} \cdots d_n$ in the coefficients of $f_i$, by Lemma 3.3 and Theorem 3.4. Also by Theorem 3.4, *Res* divides $G$. Therefore *Res* and $G$ are equal up to a rational constant. □

In order to apply the formulas in Theorem 3.4 and Corollary 3.6 for computing the resultant it is useful to know the extraneous factor $D_n/Res$ a priori. This information is provided by work of Macaulay [Mac]: he shows that the polynomial $D_n/Res$ equals the determinant of a certain explicit submatrix of the system arising from (3.5). This was exploited by Canny and his collaborators ([Can],[MC], [CE]) in the design of interpolation-based algorithms for evaluating the resultant *Res*. The numerical evaluation of the resultant for a system (3.1) with rational coefficients involves some clever perturbation techniques introduced in [Can]. We close this

section with a reference to [Jo1] which contains a systematic algebraic development of the theory of the classical multivariate resultant.

## 4. The sparse resultant

Most systems of polynomial equations encountered in real world applications are *sparse* in the sense that only few monomials appear with non-zero coefficient. The definition of the resultant given in Section 3 is not well suited to this situation. As an example consider the following system of three quadratic equations:

$$f \,=\, a_0 x + a_1 y + a_2 xy, \quad g \,=\, b_0 + b_1 xy + b_2 y^2, \quad h \,=\, c_0 + c_1 xy + c_2 x^2. \tag{4.1}$$

If we substitute the coefficients of $f, g$ and $h$ into the resultant (1.7), i.e., if we set the nine missing coefficients to zero in (1.7), then the resulting expression vanishes identically. This is consistent with Theorem 3.1 because the homogeneous equations

$$F \,=\, a_0 xz + a_1 yz + a_2 xy, \quad G \,=\, b_0 z^2 + b_1 xy + b_2 y^2, \quad H \,=\, c_0 z^2 + c_1 xy + c_2 y^2$$

always have the common root $(1,0,0)$, regardless of what the coefficients $a_i, b_j, c_k$ are. In other words, the three given quadrics always intersect in the projective plane $\mathbb{P}^2$. But they generally do not intersect in the affine plane $\mathbb{C}^2$. In order for this to happen, the following polynomial in the coefficients must vanish:

$$\begin{aligned}
&a_1^2 b_2 b_1^2 c_0^2 c_1 - 2a_1^2 b_2 b_1 b_0 c_0 c_1^2 + a_1^2 b_2 b_0^2 c_1^3 - a_1^2 b_1^3 c_0^2 c_2 + 2a_1^2 b_1^2 b_0 c_0 c_1 c_2 \\
&- a_1^2 b_1 b_0^2 c_1^2 c_2 - 2a_1 a_0 b_2^2 b_1 c_0^2 c_1 + 2a_1 a_0 b_2^2 b_0 c_0 c_1^2 + 2a_1 a_0 b_2 b_1^2 c_0^2 c_2 \\
&- 2a_1 a_0 b_2 b_0^2 c_1^2 c_2 - 2a_1 a_0 b_1^2 b_0 c_0 c_2^2 + 2a_1 a_0 b_1 b_0^2 c_1 c_2^2 + a_0^2 b_2^3 c_0^2 c_1 - a_0^2 b_2^2 b_1 c_0^2 c_2 \\
&- 2a_0^2 b_2^2 b_0 c_0 c_1 c_2 + 2a_0^2 b_2 b_1 b_0 c_0 c_2^2 + a_0^2 b_2 b_0^2 c_1 c_2^2 - a_0^2 b_1 b_0^2 c_2^3 - a_2^2 b_2^2 b_1 c_0^3 \\
&+ a_2^2 b_2^2 b_0 c_0^2 c_1 + 2a_2^2 b_2 b_1 b_0 c_0^2 c_2 - 2a_2^2 b_2 b_0^2 c_0 c_1 c_2 - a_2^2 b_1 b_0^2 c_0 c_2^2 + a_2^2 b_0^3 c_1 c_2^2.
\end{aligned} \tag{4.2}$$

The expression (4.2) is called the *sparse resultant* of (4.1). This resultant is custom-tailored to the specific monomials appearing in the given input equations.

In this section we introduce the basic set-up of "sparse elimination theory". In particular, we present the precise definition of the sparse resultant (caveat: it is subtle). For details and further reading we refer to [GKZ2], [EC], [PS], [St1], [St2].

Let $\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_n$ be finite subsets of $\mathbb{Z}^n$. Set $m_i := card(\mathcal{A}_i)$. Consider a system of $n+1$ Laurent polynomials in $n$ variables $x = (x_1, \ldots, x_n)$ of the following form

$$f_i(x) \quad = \quad \sum_{a \in \mathcal{A}_i} c_{ia}\, x^a \qquad\qquad (i = 0, 1, \ldots, n). \tag{4.3}$$

Here $x^a = x_1^{a_1} \cdots x_n^{a_n}$ for $a = (a_1, \ldots, a_n) \in \mathbb{Z}^n$. We say that $\mathcal{A}_i$ is the *support* of the polynomial $f_i(x)$. For instance, in the example (4.1) we have $n = 2$, $m_1 = m_2 = m_3 = 3$, $\mathcal{A}_0 = \{\,(1,0),\,(0,1),\,(1,1)\,\}$ and $\mathcal{A}_1 = \mathcal{A}_2 = \{\,(0,0),\,(1,1),\,(0,2)\}$.

For any subset $J \subseteq \{0,\ldots,n\}$ we consider the affine lattice spanned by $\sum_{j\in J} \mathcal{A}_j$,

$$\mathcal{L}_J \quad := \quad \{ \sum_{j\in J} \lambda_j a^{(j)} \;\mid\; a^{(j)} \in \mathcal{A}_j,\, \lambda_j \in \mathbb{Z} \text{ for all } j \in J \quad \text{and} \quad \sum_{j\in J} \lambda_j \,=\, 1 \,\}.$$

In what follows we shall assume that $\mathcal{L}_{\{0,1,\ldots,n\}} = \mathbb{Z}^n$. Let $rank(J)$ denote the rank of the lattice $\mathcal{L}_J$. A subcollection of supports $\{\mathcal{A}_i\}_{i\in I}$ is said to be *essential* if

$$rank(I) \,=\, card(I)-1 \quad \text{and} \quad rank(J) \,\geq\, card(J) \quad \text{for each proper subset } J \text{ of } I.$$

The vector of all coefficients $c_{ia}$ in (4.3) defines a point in the product of complex projective spaces $\mathbb{P}^{m_0-1} \times \cdots \times \mathbb{P}^{m_n-1}$. Let $Z$ denote the subset of those systems (4.3) which have a solution $x$ in $(\mathbb{C}^*)^n$, where $\mathbb{C}^* := \mathbb{C}\backslash\{0\}$. Let $\bar{Z}$ be the closure of $Z$ in $\mathbb{P}^{m_0-1} \times \cdots \times \mathbb{P}^{m_n-1}$. (Here it is equivalent to take the strong closure or the Zariski closure). The following lemma is proved similarly to Theorem 3.1 (b).

**Lemma 4.1.** *The projective variety $\bar{Z}$ is irreducible and defined over $\mathbb{Q}$.*

We now define the *sparse resultant* $\mathcal{R}$. If $codim(\bar{Z}) = 1$ then $\mathcal{R}$ is the unique (up to sign) irreducible polynomial in $\mathbb{Z}[\ldots, c_{ia}, \ldots]$ which vanishes on the hypersurface $\bar{Z}$. If $codim(\bar{Z}) \geq 2$ then $\mathcal{R}$ is defined to be the constant 1. Using Bernstein's Theorem [GKZ2, §6.2.D] one can derive the following result (see [PS] for details). Theorem 4.2 is a generalization of Theorem 3.2.

**Theorem 4.2.** *Suppose that $\{\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_n\}$ is essential, and let $Q_i$ denote the convex hull of $\mathcal{A}_i$. For all $i \in \{0,\ldots,n\}$ the degree of $\mathcal{R}$ in the $i$-th group of variables $\{c_{ia},\, a \in \mathcal{A}_i\}$ is a positive integer, equal to the mixed volume*

$$(4.4) \quad \mathcal{M}(Q_0,\ldots,Q_{i-1},Q_{i+1}\cdots,Q_n) \;=\; \sum_{J\subseteq\{0,\ldots,i-1,i+1\ldots,n\}} (-1)^{card(J)}\, vol\,\big(\sum_{j\in J} Q_j\big).$$

We refer to [GKZ2, §6.2.D] or [PS] for other definitions of the *mixed volume*. See [EC] for a practical algorithm for computing mixed volumes, that is, for evaluating the right hand side of (4.4). We next determine the codimension of the variety $\bar{Z}$ of solvable systems (4.3). The following two results appear in [St2, §1]:

**Theorem 4.3.** *The codimension of $\bar{Z}$ in $\mathbb{P}^{m_0-1} \times \cdots \times \mathbb{P}^{m_n-1}$ equals the maximum of the numbers $card(I) - rank(I)$, where $I$ runs over all subsets of $\{0,1,\ldots,n\}$.*

Next we give a combinatorial criterion for the existence of a non-trivial sparse resultant. Note that, if each $\mathcal{A}_i$ is $n$-dimensional, then $I = \{0,1,\ldots,n\}$ is essential.

**Corollary 4.4.** *The variety $\bar{Z}$ has codimension 1 if and only if there exists a unique subset $\{\mathcal{A}_i\}_{i\in I}$ which is essential. In this case the sparse resultant $\mathcal{R}$ coincides with the sparse resultant of the equations $\{f_i : i \in I\}$.*

**Example 4.5.** For the linear system

$$(4.5) \qquad c_{00}x + c_{01}y \quad = \quad c_{10}x + c_{11}y \quad = \quad c_{20}x + c_{21}y + c_{22} \quad = \quad 0,$$

the variety $\bar{Z}$ has codimension 1 in the coefficient space $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^2$. The unique essential subset consists of the first two equations. Hence the sparse resultant of (4.5) is *not* the $3 \times 3$-determinant (which would be reducible). The sparse resultant of (4.5) equals the $2 \times 2$-determinant $\mathcal{R} \;=\; c_{00}c_{11} \,-\, c_{10}c_{01}$. $\square$

We illustrate Theorem 4.2 with the little example in (4.1). Clearly, the triple of support sets $\{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3\}$ is essential, since all three *Newton polygons* $Q_i = conv(\mathcal{A}_i)$ are triangles. The mixed volume of two polygons in the plane can be computed by the formula (4.4) for $n = 2$:

$$\mathcal{M}(Q_i, Q_j) \quad = \quad area(Q_i + Q_j) - area(Q_i) - area(Q_j).$$

In our example the triangles $Q_2$ and $Q_3$ coincide, and we have

$$area(Q_1) = 1/2, \;\; area(Q_2) = 1, \;\; area(Q_1 + Q_2) = 9/2, \;\; area(Q_2 + Q_3) = 4.$$

This implies

$$\mathcal{M}(Q_1, Q_2) \;=\; \mathcal{M}(Q_1, Q_3) \;=\; 3 \quad \text{and} \quad \mathcal{M}(Q_2, Q_3) \;=\; 2. \tag{4.6}$$

This explains why the sparse resultant (4.2) is quadratic in $(a_0, a_1, a_2)$ and homogeneous of degree 3 in $(b_0, b_1, b_2)$ and in $(c_0, c_1, c_2)$ respectively.

One of the central problems in elimination theory is to find "nice" determinantal formulas for resultants. The best one can hope for is a *Sylvester-type formula*, that is, a square matrix whose non-zero entries are the coefficients of the given equation and whose determinant equals precisely the resultant. The archetypical example of such a formula is (2.1). Sylvester-type formulas do exist for many important special cases; see [GKZ2, Propositions 13.1.6 and 13.2.2]. However, they do not exist in general, even for the classical multivariate resultant discussed in Section 3. The state of the art on these matters can be found in the work of Weyman and Zelevinsky [WZ].

If a Sylvester-type formula is not available or too hard to find, the next best thing is to construct a "reasonably small" square matrix whose determinant is a non-zero multiple of the resultant under consideration. For the sparse resultant such a construction was given by Canny and Emiris [EC]; see also [St2, §3]. We present an instance of the Canny-Emiris construction for our example (4.1):

$$\begin{array}{c} \\ yf \\ y^2f \\ xy^2f \\ y^2g \\ xy^2g \\ yg \\ xyg \\ xy^2h \\ yh \\ xyh \end{array}
\begin{array}{c}
\begin{array}{cccccccccc} y^2 & y^3 & xy^3 & y^4 & xy^4 & xy^2 & x^2y^2 & x^2y^3 & y & xy \end{array} \\
\left(\begin{array}{cccccccccc}
a_1 & 0 & 0 & 0 & 0 & a_2 & 0 & 0 & 0 & a_0 \\
0 & a_1 & a_2 & 0 & 0 & a_0 & 0 & 0 & 0 & 0 \\
0 & 0 & a_1 & 0 & 0 & 0 & a_0 & a_2 & 0 & 0 \\
b_0 & 0 & b_1 & b_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & b_2 & b_0 & 0 & b_1 & 0 & 0 \\
0 & b_2 & 0 & 0 & 0 & b_1 & 0 & 0 & b_0 & 0 \\
0 & 0 & b_2 & 0 & 0 & 0 & b_1 & 0 & 0 & b_0 \\
0 & 0 & 0 & 0 & c_2 & c_0 & 0 & c_1 & 0 & 0 \\
0 & c_2 & 0 & 0 & 0 & c_1 & 0 & 0 & c_0 & 0 \\
0 & 0 & c_2 & 0 & 0 & 0 & c_1 & 0 & 0 & c_0
\end{array}\right)
\end{array} \tag{4.7}$$

The determinant of this matrix equals $a_1 b_2$ times the sparse resultant (4.2).

The structure of the Canny-Emiris matrix (4.7) can be understood as follows. Form the product $fgh$ and expand it into monomials in $x$ and $y$. A certain combinatorial rule selects 10 out of the 15 monomials appearing in $fgh$. The columns of the matrix (4.7) are indexed by these 10 monomials. Suppose the $i$-th column is indexed by the monomial $x^j y^k$, Next there is a second combinatorial rule which selects a monomial multiple of one of the input equations $f$, $g$ or $h$ such that this multiple contains $x^i y^j$ in its expansion. The $i$-th row of (4.7) is indexed by that polynomial. Finally the $(i,j)$-entry contains the coefficient of the $j$-th column monomial in the $i$-th row polynomial. This construction implies that the matrix (4.7) has non-zero entries along the main diagonal.

The two combinatorial rules mentioned in the previous paragraph are based on a geometric construction called *mixed subdivision of the Newton polytypes*. We refer to the original paper [EC] for the description of this polyhedral construction.

The main difficulty overcome by the Canny-Emiris formula is this: If one sets up a matrix like (4.7) just by "playing around" then most likely its determinant will vanish (try it), unless there is a good reason why it shouldn't vanish. Now the key idea is this: a big unknown polynomial (such as $\mathcal{R}$) will be non-zero if one can ensure that its initial monomial (with respect to some term order) is non-zero.

Consider the lexicographic term order induced by the variable ordering $a_1 > a_0 > a_2 > b_2 > b_1 > b_0 > c_0 > c_1 > c_2$. The monomials in (4.2) are listed in this order. Consider all permutations in $S_{10}$ which contribute a (possible) non-zero term to the expansion of the determinant of (4.7). There will undoubtedly be some cancellation. However, the unique largest monomial (in the above term order) appears only once, namely, on the main diagonal. This guarantees that the determinant of (4.7) is a non-zero polynomial. Note that the product of the diagonal elements in (4.7) equals $a_1 b_2$ times the leading monomial in (4.2).

An explicit combinatorial construction for all possible initial monomials (with respect to any term order) of the sparse resultant of (4.3) is given in [St2]. Moreover, it is shown in [St2, §3] that for any such initial monomial there exists a Canny-Emiris matrix which has that monomial on its main diagonal.

We close with some pointers to contemporary research in elimination theory. For work on resultants from the perspective of commutative algebra see [Jo1],[Jo2],[SS]. If one has more equations than variables to be eliminated, then one needs *resultant systems* instead of resultants; see [Lyu]. Another important topic is *subresultants* [Ch2] which are a tool for detecting higher degrees of solvability for systems of algebraic equations. A central theme in the work on resultants by Gel'fand, Kapranov and Zelevinsky [GKZ2] is the determinant of the Cayley-Koszul complex. This allows to write the resultant as an alternating product of Sylvester-like determinants. The computational feasibility of this was explored by Chardin in [Ch1].

There are numerous articles dealing with practical use of resultants for problems in applied mathematics and engineering. Three such articles are [AGW], [GMS] and [Wei]. For research on resultants from the point of view of computational complexity theory see [LM] and the references given there.

One generalization of resultants is particularly important in algebraic geometry; this is the notion of the *Chow form* of a projective variety. See [GKZ2, Chapter 4] or [Sha, §I.6] for the basics, and [KSZ] and [PS] for additional results.

## References


[AGM] E.L. Allgower, K. Georg and R. Miranda, *The method of resultants for computing real solutions of polynomial systems*, SIAM J. Numerical Analysis **29** (1992), 831–844.

[Can] J. Canny, *Generalized characteristic polynomials*, Journal of Symbolic Computation **3** (1990), 241–250.

[Ch1] M. Chardin, *The resultant via a Koszul complex*, Computational algebraic geometry (Nice, 1992), 29–39, Progr. Math., 109, Birkhauser Boston, Boston, MA, 1993.

[Ch2] M. Chardin, *Multivariate subresultants.*, J. Pure Appl. Algebra **101** (1995), 129–138.

[Cox] D. Cox, *Introduction to Gröbner bases*, these lecture notes.

[CLO1] D. Cox, J. Little and D. O'Shea, *Ideals, Varieties, and Algorithms*, Springer-Verlag, New York-Berlin-Heidelberg, 1996.

[CLO2] D. Cox, J. Little and D. O'Shea, *Using Algebraic Geometry*, in preparation.

[Dix] A.L. Dixon, *The eliminant of three quantics in two independent variables*, Proceedings of the London Mathematical Society **7** (1908), 49–62.

[Eis] D. Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*, Springer-Verlag, New York-Berlin-Heidelberg, 1994.

[EC] I. Emiris and J. Canny, *Efficient incremental algorithms for the sparse resultant and the mixed volume*, Journal of Symbolic Computation **20** (1995), 117-149.

[GKZ1] I.M. Gel'fand, M. Kapranov, A. Zelevinsky, *Newton polytopes of the classical discriminant and resultant*, Advances in Mathematics **84** (1990), 237–254.

[GKZ2] I.M. Gel'fand, M. Kapranov, A. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston-Basel-Berlin, 1994.

[GMS] J. Guckenheimer, M. Myers and B. Sturmfels, *Computing Hopf bifurcations*, *SIAM J. Numerical Analysis*, to appear.

[Jo1] J.-P. Jouanolou, *Le formalisme du résultant*, Advances Math. **98** (1991), 117–263.

[Jo2] J.-P. Jouanolou, *Aspects invariants de l'élimination*, Advances Math. **114** (1995), 1–174.

[KSZ] M.M. Kapranov, B. Sturmfels and A. Zelevinsky, *Chow polytopes and general resultants*, Duke Mathematical Journal **67** (1992), 189–218.

[LM] T. Lickteig and K. Meer, *A note on testing the resultant*, Journal of Complexity **11** (1995), 344–351.

[Lyu] G. Lyubeznik, *Minimal resultant systems*, Journal of Algebra **177** (1995), 612–616.

[Mac] F.S. Macaulay, *On some formulae in elimination*, Proceedings of the London Mathematical Society (1902), 3-27.

[Man] D. Manocha, *Numerical Methods for Solving Polynomial Equations*, these lecture notes.

[MC] D. Manocha and J. Canny, *Multipolynomial resultant algorithms*, Journal of Symbolic Computation **15** (1993), 99–122.

[PS] P. Pedersen and B. Sturmfels, *Product formulas for resultants and Chow forms*, Mathematische Zeitschrift **214** (1993), 377-396.

[SS] G. Scheja and U. Storch, *On projective representations of finite algebras by graded complete intersections*, Rend. Sem. Mat. Univ. Politec. Torino **49** (1993), 1–28.

[Sed] T. Sederberg, *Applications to computer-aided geometric design*, these lecture notes.

[Sha] I. Shafarevich, *Basic Algebraic Geometry*, Springer, Berlin-Heidelberg-New York, 1977.

[St1] B. Sturmfels, *Sparse elimination theory*, in *"Computational Algebraic Geometry and Commutative Algebra"* [D. Eisenbud and L. Robbiano, eds.], Proceedings Cortona (June 1991), Cambridge University Press, 1993, pp. 264–298.

[St2] B. Sturmfels, *The Newton polytope of the resultant*, J. Alg. Combinatorics **3** (1994), 207-236.

[Wei] J. Weiss, *Resultant methods for the inverse kinematics problem*, Computational kinematics (Dagstuhl Castle, 1993), 41–52, Solid Mech. Appl., 28, Kluwer Acad. Publ., Dordrecht, 1993.

[WZ] J. Weyman and A. Zelevinsky, *Determinantal formulas for multigraded resultants*, Journal of Algebraic Geometry **3** (1994), 569-597.



Department of Mathematics, University of California, Berkeley, CA 94720
*E-mail address*: bernd@math.berkeley.edu