Compact Formulae in Sparse Elimination

[Abstract of Invited talk at ACM ISSAC, July 2016]

Ioannis Z. Emiris[*]

National and Kapodistrian University of Athens, and Athena Research Center, Greece
emiris@di.uoa.gr

May 13, 2019

Abstract

It has by now become a standard approach to use the theory of sparse (or toric) elimination, based on the Newton polytope of a polynomial, in order to reveal and exploit the structure of algebraic systems. This talk surveys compact formulae, including older and recent results, in sparse elimination.

We start with root bounds and juxtapose two recent formulae: a generating function of the m-Bézout bound and a closed-form expression for the mixed volume by means of a matrix permanent. For the sparse resultant, a bevy of results have established determinantal or rational formulae for a large class of systems, starting with Macaulay. The discriminant is closely related to the resultant but admits no compact formula except for very simple cases. We offer a new determinantal formula for the discriminant of a sparse multilinear system arising in computing Nash equilibria. We introduce an alternative notion of compact formula, namely the Newton polytope of the unknown polynomial. It is possible to compute it efficiently for sparse resultants, discriminants, as well as the implicit equation of a parameterized variety. This leads us to consider implicit matrix representations of geometric objects.

Keywords. Sparse elimination; permanent; generating function; resultant matrix; discriminant; matrix representation

1 Introduction

It is today a standard approach to use the theory of sparse (or toric) elimination, based on the Newton polytope of a polynomial, in order to reveal and exploit the structure of algebraic systems. Consider Laurent polynomials in N variables:

$$f_i = \sum_{j=1}^{k_i} c_{ij} x^{a_{ij}}, \quad c_{ij} \neq 0,$$

where $x = (x_1, \ldots, x_N)$, $x^e = \prod_i x_i^{e_i}$, and $\{a_{i1}, \ldots, a_{ik_i}\} \subset \mathbb{Z}^N$ is the *support* of f_i. Let Q_i be the *Newton polytope* of f_i defined as the convex hull of the support.

A more constrained form of structure corresponds to multihomogeneous systems, which are defined on a product

$$\mathbb{P}^{n_1} \times \mathbb{P}^{n_2} \times \cdots \times \mathbb{P}^{n_S} \tag{1}$$

of projective spaces: the variables are partitioned into S subsets, or blocks, so that each equation is homogeneous in each block of $n_j + 1$ homogeneous variables. In contrast, semi-mixed systems are those whose polynomials can be partitioned into subsets, or blocks, with fixed monomial support. In sparse elimination, each subset of polynomials has the same Newton polytope. In the case of semi-mixed multihomogeneous systems, the polynomials in each block are homogeneous of the same degree per block of variables.

An interesting example of a semi-mixed multihomogeneous system appears in game theory and was studied in [20, 26]. Consider a game of S players, each with m_i options or strategies. The j-th player plays a totally mixed strategy when the corresponding strategy is chosen randomly with nonzero probabilities $p_1^{(j)}, \ldots, p_{m_j}^{(j)}$. A totally mixed Nash equilibrium (TMNE) is a combination of such strategies so that no player can improve his payoff by unilaterally choosing other (pure or mixed) strategy. We express payoff P_j of the j-th player choosing the i-th strategy by the polynomial

$$\sum_{k_1,\ldots,k_{j-1},k_{j+1},\ldots,k_S} a_{\ldots,k_{j-1},i,k_{j+1},\ldots}^{(j)} p_{k_1}^{(1)} \cdots p_{k_{j-1}}^{(j-1)} p_{k_{j+1}}^{(j+1)} \cdots p_{k_S}^{(S)}. \tag{2}$$

Here $a_{k_1,k_2,\ldots,k_S}^{(j)}$ denotes the predefined payoff of the player, assuming each player chooses the pure option $k_\ell \in \{1, \ldots, m_\ell\}$. The equations imply that the payoff P_j does not depend on own strategy i, as long as other players do not change their strategies. Eliminating the P_j's leads to a multilinear system with $n_j = m_j - 1$ in space (1). Actual probabilities are determined from the normalizing conditions $\sum_{i=1}^{m_j} p_i^{(j)} = 1$, for $j \in \{1, \ldots, S\}$.

2 Root bounds

Bounds on the number of common roots of an algebraic system admit different types of compact formulae. We start with classic bounds named after Bézout, tight for (multi)homogeneous systems. The most general such bound is the following m-Bézout bound, stated for dehomogenized system.

Theorem 2.1 (m-Bézout bound). *Consider a system of N equations in N affine variables, partitioned into S subsets so that the j-th subset includes n_j affine variables, and $N = n_1 + \cdots + n_S$. Let d_{ij} be the degree of the i-th equation in the j-th variable subset, for $i = 1, \ldots, N$ and $j = 1, \ldots, S$. If the system has finitely many complex roots in space (1), the coefficient of $x_1^{n_1} \cdots x_S^{n_S}$ in*

$$\prod_{i=1}^{N} (d_{i1}x_1 + \cdots + d_{iS}x_S) \tag{3}$$

bounds the number of roots. For generic coefficients this bound is tight.

We shall give a generating function for the m-Bézout number by extending results in [26, 27, 32]. MacMahon's Master theorem [25] is a powerful combinatorial result.

Theorem 2.2 (MacMahon's Master Theorem). *[25] Let $A = (a_{ij})$ be a complex $S \times S$ matrix, $x_1, \ldots, x_S$ be formal variables, and V denote the diagonal matrix with nonzero entries $x_1, \ldots, x_S$. The coefficient of $x_1^{n_1} \cdots x_S^{n_S}$ in*

$$\prod_{j=1}^{S} (a_{j1}x_1 + \cdots + a_{jS}x_S)^{n_j} \tag{4}$$

equals the coefficient of $x_1^{n_1} \cdots x_S^{n_S}$ in the multivariate Taylor expansion of

$$f(x_1, \ldots, x_S) = \frac{1}{\det(I - VA)} \tag{5}$$

around $(x_1, \ldots, x_S) = (0, 0, \ldots, 0)$.

Let us consider semi-mixed multihomogeneous systems, where the number of equation blocks equals the number S of variable blocks, and we have a varying number of equations per block; more importantly, the degree per block is arbitrary. Generically, this system has a finite number of solutions. MacMahon's Theorem yields a multivariate generating function for the m-Bézout bound of these systems, which could be helpful when one seeks root counts for a family of systems.

Theorem 2.3. [20] *Consider a multihomogeneous system on space (1) of $N = n_1 + \cdots + n_S$ equations, where the equations are partitioned into S subsets of exactly $n_1, \ldots, n_S$ equations. We assume that for any $i, j \in \{1, 2, \ldots, S\}$, the polynomial equations in the i-th subset have degree a_{ij} in the variables of the j-th variable subset. Let A be the $S \times S$ matrix defined by the a_{ij}'s. Then the m-Bézout bound for the multihomogeneous system equals the coefficient of $x_1^{n_1} \cdots x_S^{n_S}$ in the multivariate Taylor expansion of*

$$1/\det(I - VA)$$

around $(x_1, \ldots, x_S) = (0, 0, \ldots, 0)$.

For TMNE system (2), we have $a_{ii} = 0$, and $a_{ij} = 1$ for $i \neq j$. Then

$$I - VA = \begin{bmatrix} 1 & -x_1 & -x_1 & \cdots & -x_1 \\ -x_2 & 1 & -x_2 & \cdots & -x_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -x_S & -x_S & \cdots & -x_S & 1 \end{bmatrix},$$

and

$$1/\det(I - VA) = \frac{1}{1 - \sigma_2 - 2\sigma_3 - \cdots - (S-1)\sigma_S},$$

where σ_j (for $j = 2, \ldots, S$) denotes the j-th elementary symmetric polynomial in $x_1, \ldots, x_S$.

Now we pass to the most general root bound in sparse elimination theory, namely mixed volume: there are efficient algorithms but, in certain cases, the mixed volume computation is reduced to a matrix permanent. The toric root bound is named after Bernstein, Khovanskii, and Kushnirenko (BKK) [1]:

Theorem 2.4. *For $f_1, \ldots f_N \in \mathbb{C}[x_1^{\pm 1}, \ldots, x_N^{\pm 1}]$ with Newton polytopes $Q_1, \ldots, Q_N$, the number of common isolated solutions, multiplicities counted, in the corresponding toric variety, which contains $(\mathbb{C}^*)^N$ as a dense subset, does not exceed $\mathrm{MV}(Q_1, \ldots, Q_N)$, independently of the toric variety's dimension.*

It is known that mixed volume is related to the permanent. In fact, by reduction to the latter one shows that mixed volume is #P-complete. For a multihomogeneous system on space (1) of $N = n_1 + \cdots + n_S$ equations, assume the i-th equation has degree a_{ij} in the j-th variable block. Let $A = (a_{ij})$ be an $N \times N$ matrix with the columns repeated n_j times for each $j \in \{1, 2, \ldots, S\}$. Then the mixed volume equals [27, Thm 2]

$$\frac{1}{n_1! \cdots n_S!} \, \mathrm{perm} \, A.$$

A generalization to arbitrary (non-homogeneous) systems follows. For each block of variables there is a n_j-dimensional polytope Γ_j, each in a separate complementary space, $1 \leq j \leq S$. Assume the Newton polytopes are direct products of scalar multiples (by a_{ij}) of the Γ_j's. We obtain an algebraic system on the product of toric varieties corresponding to Γ_j, of dimension n_j.

Theorem 2.5. [20] *Let A be the matrix of a_{ij}'s with columns repeated n_j times. If there are $N = \sum_{j=1}^{S} n_j$ equations, then*

$$\mathrm{MV}(Q_1, \ldots, Q_N) = \mathrm{perm}\, A \prod_{j=1}^{S} \mathrm{vol}(\Gamma_j), \quad Q_i = \prod_{j=1}^{S} a_{ij} \Gamma_j.$$

3 Resultant formulae

For the sparse resultant, a bevy of results have established matrix and, more particularly, determinantal formulae for a large class of systems. An introductory survey can be found in [28].

Resultants provide efficient ways for studying and solving polynomial systems by means of matrices whose determinant is a non-trivial multiple of the resultant. They are most efficiently expressed by a generically non-singular matrix, whose determinant is the resultant polynomial or, when this is impossible, by also specifying a minor M' of resultant matrix M which divides the determinant of the latter so as to yield a rational formula for the resultant R:

$$R = \det M / \det M'.$$

Macaulay's classical result [24] establishes such rational formula for arbitrary dense systems, and D'Andrea, 100 years later [7], for arbitrary systems in the context of sparse elimination. His recursive construction has been simplified in certain cases [17].

For two univariate polynomials there are matrix formulae named after Sylvester and Bézout, whose determinant is equal to the resultant; we refer to them as determinantal formulae. The largest family admitting determinantal formulae is a class of multihomogeneous systems studied in [31, 33], see also [21, Sect.13.2], where, each block of variables contains a single dehomogenized variable or is of linear total degree. These systems are moreover unmixed, in other words all polynomials have the same Newton polytope; they are called multigraded.

The multigraded resultant matrices were made explicit in [10, 18], which offered determinantal formulae of Sylvester, Bézout and hybrid types, the latter containing blocks of both Sylvester and Bézout types. In [10], it is shown that there exists a determinantal pure Bézout-type resultant formula if and only if there exists such a Sylvester-type formula. The Bézout-type matrices generalize those identified in [6]. In [18] was proven the existence of determinantal formulae for systems whose Newton polytopes are scaled copies of one polytope.

In [8], hybrid formulae were proposed in the mixed homogeneous case. The most complete study of hybrid formulae is [12] where resultants are given as a matrix determinant or a Pfaffian, i.e. a determinant square root, for unmixed systems of 3 dense polynomials, for up to 5 polynomials of degree up to 4, 6, or 8, and for 6 quadratic polynomials.

Some studies have focused on three polynomials. In the unmixed case a determinantal formula is established in [23], more direct than [12] and generalizing the Macaulay-style formula obtained for the dense case in [22], where the numerator matrix has one row corresponding to the affine Jacobian. For Newton polygons which are scaled copies of a single one, the smallest Sylvester-type matrices are obtained by using a row containing the coefficients of the toric Jacobian [9]. In [34] they construct determinantal Sylvester-type formulae for unmixed systems whose Newton polygon is a rectangle from which smaller rectangles have been removed at the corners.

4 Discriminant formulae

Discriminants are crucial in studying well-constrained algebraic systems, the system's zero set, its singularities, and other properties of the system. The discriminant is closely related to the resultant but admits no compact formula except for very simple cases. Below we offer a new determinantal formula for a specific system. Recent work has shed light on the degree of mixed discriminants [5, 11] as a function of the Newton polytopes of the corresponding polynomials. An important open question is to obtain an explicit formula for multilinear systems [30].

In the context of sparse elimination the discriminant is defined for a set of polynomials F_i with fixed supports A_i, hence it is denoted $\Delta_{A_1,\ldots,A_n}(F_1,\ldots,F_n)$ or $\Delta(F_1,\ldots,F_n)$.

Definition 4.1. *The mixed discriminant $\Delta(F_1,\ldots,F_n)$ of n polynomials in n variables with fixed supports $A_1,\ldots,A_n$ in $\mathbb{Z}^n$ is the irreducible polynomial (with integer coprime coefficients, defined up to sign) in the coefficients of the F_i which vanishes whenever the system $F_1 = \cdots = F_n = 0$ has a multiple root (i.e., a root which is not simple) with nonzero coordinates, provided his discriminantal variety is a hypersurface; otherwise, the mixed discriminant is equal to the constant 1.*

Clearly, the zero locus of the mixed discriminant is the variety of ill-posed systems, i.e., systems with at least one multiple root.

We focus on systems over $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ where each equation is missing one variable, motivated by the study of TMNE in games of 3 players with 2 pure strategies each, see polynomials (2). The bilinear system is as follows:

$$
\begin{aligned}
F_1 : & & a_0 y_1 z_1 + a_1 y_1 z_0 + a_2 y_0 z_1 + a_4 y_0 x_0 &= 0, \\
F_2 : & & b_0 x_1 z_1 + b_1 x_1 z_0 + b_3 x_0 z_1 + b_4 x_0 z_0 &= 0, \\
F_3 : & & c_0 x_1 y_1 + c_2 x_1 y_0 + c_3 x_0 y_1 + c_4 x_0 y_0 &= 0.
\end{aligned}
$$

The projective coordinates $(x_1 : x_0)$, $(y_1 : y_0)$, $(z_1 : z_0)$ represent the unknown probability ratios for the two options of each player; the coefficients a_i equal the differences in the given payoffs (between the two options) of the first player, for the 4 combinations of pure choices of the other two players; similarly, b_i, c_i equal the differences in the given payoffs (between their two options) of the other two players.

The following theorem gives a 6×6 determinantal expression for the discriminant. The formula is of Sylvester-type because the matrix entries are either zero or equal to a polynomial coefficient.

Theorem 4.2. [19] *The discriminant $\Delta(F_1, F_2, F_3)$ equals*

$$
\det \begin{bmatrix}
0 & 0 & c_0 & c_1 & b_0 & b_1 \\
0 & 0 & c_2 & c_4 & b_3 & b_4 \\
c_0 & c_2 & 0 & 0 & a_0 & a_2 \\
c_1 & c_4 & 0 & 0 & a_3 & a_4 \\
b_0 & b_3 & a_0 & a_3 & 0 & 0 \\
b_1 & b_4 & a_2 & a_4 & 0 & 0
\end{bmatrix}.
$$

5 Sparse implicitization

We argue that in sparse elimination it makes sense to consider an alternative notion of compact formula, namely the Newton polytope of the unknown polynomial. It typically allows computation of the polynomial by interpolation of the coefficients. It is possible to compute it efficiently for sparse resultants [13], discriminants, as well as the implicit equation of a parameterized variety.

This section focuses on the latter case. Given (a superset of) the Newton polytope of the implicit equation, one constructs an interpolation matrix, which allows for various geometric operations to be efficiently computed on this matrix without need of developing the monomial representation of the implicit equation. The representation by interpolation matrices has been developed for parametric plane curves, surfaces, and hypersurfaces in the context of sparse elimination [15, 16]. It generalizes to space curves and, generally, objects of codimension higher than 1 by means of the Chow form, which is a generalization of resultants to N-variate systems with more than $N + 1$ polynomials [14].

Another powerful matrix representation of geometric objects is based on syzygies. The corresponding theory, including μ-bases, has been developed for parametric models [3, 4, 29]. The computation of syzygies can exploit the sparse elimination setting [2] and, moreover, can be achieved for point cloud models [14]. Roughly, syzygies of a sufficiently high degree define moving curves or surfaces that "follow" the parametric object. One forms the matrix expressing these moving curves and/or surfaces, which represents the geometric object because its rank at some point p drops if and only if p belongs to the algebraic closure of the image of the parameterization [3].

References

[1] D. Bernstein. The number of roots of a system of equations. *Funct. Anal. and Appl.*, 9(2):183–185, 1975.

[2] N. Botbol and A. Dickenstein. Implicitization of rational hypersurfaces via linear syzygies: A practical overview. *J. Symbolic Computation*, 74:493–512, 2016.

[3] L. Busé and T. Luu Ba. Matrix-based implicit representations of rational algebraic curves and applications. *Computer-Aided Geometric Design*, 27(9):681–699, 2010.

[4] L. Busé, D. Cox, and C. D'Andrea. Implicitization of surfaces in P^3 in the presence of base points, *J. Algebra Appl.*, 2(2):189-214, 2003.

[5] E. Cattani, M. Cueto, A. Dickenstein, S. Di Rocco, and B. Sturmfels. Mixed discriminants. *Math. Z.*, 274:761–778, 2013.

[6] A. Chtcherba and D. Kapur. Conditions for exact resultants using the Dixon formulation. In *Proc. ACM Intern. Symp. on Symbolic & Algebraic Comput.*, pages 62–70, 2000.

[7] C. D'Andrea. Macaulay-style formulas for the sparse resultant. *Trans. AMS*, 354:2595–2629, 2002.

[8] C. D'Andrea and A. Dickenstein. Explicit formulas for the multivariate resultant. *J. Pure & Applied Algebra*, 164(1-2):59–86, 2001.

[9] C. D'Andrea and I. Emiris. Hybrid sparse resultant matrices for bivariate polynomials. *J. Symbolic Computation, Special Issue*, 33(5):587–608, 2002.

[10] A. Dickenstein and I. Emiris. Multihomogeneous resultant formulae by means of complexes. *J. Symbolic Computation, Special issue*, 36(3-4):317–342, 2003.

[11] A. Dickenstein, I. Emiris, and A. Karasoulou. Plane mixed discriminants and toric jacobians. In *SAGA: Advances in ShApes, Geometry, and Algebra*, volume 10 of *Geometry and Computing*, pages 105–121. Springer, 2014.

[12] D. Eisenbud and F.-O. Schreyer. Resultants and Chow forms via exterior syzygies. *J. AMS*, 16:534–579, 2003.

[13] I. Emiris, V. Fisikopoulos, C. Konaxis, and L. Peñaranda. An oracle-based, output-sensitive algorithm for projections of resultant polytopes. *Intern. J. Computational Geom. Appl., Special Issue*, 23:397–423, 2014.

[14] I. Emiris, K. Gavriil, and C. Konaxis. Matrix representations by means of interpolation. Manuscript, 2016.

[15] I. Emiris, T. Kalinka, and C. Konaxis. Geometric operations using sparse interpolation matrices. *Graphical Models*, 82:99–109, Nov. 2015.

[16] I. Emiris, T. Kalinka, C. Konaxis, and T. Luu Ba. Sparse implicitization by interpolation: Characterizing non-exactness, and an application to computing discriminants. *Computer-Aided Design, Special Issue*, 45:252–261, 2013.

[17] I. Emiris and C. Konaxis. Single-lifting Macaulay-type formulae of generalized unmixed sparse resultants. *J. Symbolic Computation*, 46(8):919–942, 2011.

[18] I. Emiris and A. Mantzaflaris. Multihomogeneous resultant matrices for systems with scaled support. In *Proc. ACM Intern. Symp. on Symbolic & Algebraic Comput.*, pages 143–150, 2009.

[19] I. Emiris and R. Vidunas. Discriminants of multilinear systems. Manuscript, 2015.

[20] I. Emiris and R. Vidunas. Root counts of semi-mixed systems, and an application to counting Nash equilibria. In *Proc. ACM Intern. Symp. on Symbolic & Algebraic Comput.*, pages 154–161, Kobe, Japan, 2014.

[21] I. Gelfand, M. Kapranov, and A. Zelevinsky. *Discriminants, Resultants and Multidimensional Determinants*. Birkhäuser, Boston, 1994.

[22] J.-P. Jouanolou. Formes d'inertie et résultant : Un formulaire. *Adv. in Math.*, 126:119–250, 1997.

[23] A. Khetan. The resultant of an unmixed bivariate system. *J. Symbolic Computation*, 36:425–442, 2003.

[24] F. Macaulay. Some formulae in elimination. *Proc. London Math. Soc.*, 1(33):3–27, 1902.

[25] P. MacMahon. *Combinatory analysis*. Cambridge Univ. Press, 1916.

[26] R. McKelvey and A. McLennan. The maximal number of regular totally mixed Nash equilibria. *J. Economic Theory*, 72:411–425, 1997.

[27] A. McLennan. The maximum number of real roots of a multihomogeneous system of polynomial equations. *Beitrage zur Algebra und Geom.*, 40:343–350, 1999.

[28] A. Morozov and S. Shakirov. New and old results in resultant theory. *Theoretical & Math. Physics*, 163(2):587–617, 2010.

[29] T. Sederberg and F. Chen. Implicitization using moving curves and surfaces. In R. Cook, editor, *Proc. SIGGRAPH*, pages 301–308. Addison Wesley, 1995.

[30] B. Sturmfels. The Hurwitz form of a projective variety. Presented at MEGA 2015. Submitted to journal, 2016.

[31] B. Sturmfels and A. Zelevinsky. Multigraded resultants of Sylvester type. *J. Algebra*, 163(1):115–127, 1994.

[32] R. Vidunas. Counting derangements and Nash equilibria. arxiv.org/1401.5400, 2014.

[33] J. Weyman and A. Zelevinsky. Multigraded formulae for multigraded resultants. *J. Algebr. Geometry*, 3(4):569–597, 1994.

[34] M. Zhang and R. Goldman. Rectangular corner cutting and Sylvester A-resultants. In *Proc. ACM Intern. Symp. on Symbolic & Algebraic Comput.*, pages 301–308, 2000.