Cayley-Dixon construction of Resultants of Multi-Univariate Composed Polynomials‡

Arthur D. Chtcherba*
University of Texas – Pan Am
Dept. of Computer Science
Edinburg, TX, USA
cherba@cs.panam.edu

Deepak Kapur*
University of New Mexico
Dept. of Computer Science
Albuquerque, NM, USA
kapur@cs.unm.edu

Manfred Minimair*†
Seton Hall University
Dept. of Mathematics and Computer Science
South Orange, NJ, USA
manfred@minimair.org

University of New Mexico, Dept. of Computer Science
Technical report **TR-CS-2005-15**

April, 2005

Abstract

The Cayley-Dixon formulation for multivariate resultants have been shown to be efficient (both experimentally and theoretically) for computing resultants by simultaneously eliminating many variables from a polynomial system. In this paper, the behavior of Cayley-Dixon resultant construction and the structure of Dixon matrices is analyzed for composed polynomial systems constructed from a multivariate system in which each variable is substituted by a univariate polynomial in a distinct variable. It is shown that Dixon projection operator (multiple of resultant) of the composed system can be expressed as a power of the resultant of the *outer* polynomial system multiplied by powers of the leading coefficients of the univariate polynomials substituted for variables in the outer system. A new resultant formula is derived for systems where it is known that the Cayley-Dixon construction does not contain extraneous factors. The derivation of the resultant formula for the composed system unifies all the known related results in the literature. Furthermore, it demonstrates that the resultant of a composed system can be effectively calculated by considering only the resultant of the outer system. Since the complexity of resultant computation is typically determined by the degree (and support) of the polynomial system, resultants of a composed system can be computed much faster by focussing only on the outer system.

1 Introduction

Problems in many application domains, including engineering and design, graphics, CAD-CAM, geometric modeling, etc. can be modelled using polynomial systems [SG86, Hof89, Mor87, KL92, Chi90, Zha00, BGW88, PK92, KSY94]. Often a polynomial system arising

*Supported by NSF grant no. CCR-0203051 and a grant from the Computer Science Research Institute at the Sandia National Labs.

†Also supported by NSF grant CCF 0430741.

‡Submitted for publication to CASC'05.

from an application has a structure. Particularly in engineering and design applications and in geometric modeling, a polynomial system can be expressed as a composition of two distinct polynomial systems, each of which is of much lower degree in comparison to the resulting system. Furthermore, if the structure of given polynomials is not known a priori, one can efficiently check if they can be decomposed [Rub00].

This paper addresses the resultant computation for such composed polynomial systems [CMW95, Jou91, Min03a, Min02, Min01, Min03b, Min03c, Min04, HM02, KS97]. The resultant of a polynomial system with symbolic parameters is a necessary and sufficient condition on its parameters for the polynomial system to have a common solution[1]. Resultant computations have been found useful in many application domains including engineering and design, robotics, inverse kinematics, manufacturing, design and analysis of nano devices in nanotechnology, image understanding, graphics, solid modeling, implicitization, CAD-CAM design, geometric construction, drug-design, and control theory.

The focus in this paper is on the Cayley-Dixon formulation for multivariate resultants which have been shown to be efficient (both experimentally and theoretically) for computing resultants by simultaneously eliminating many variables from a polynomial system. The behavior of Cayley-Dixon resultant construction and structure of Dixon matrices is analyzed for composed polynomial systems constructed from a multivariate system in which each variable is substituted by a univariate polynomial in a distinct variable, referred as *multi-univariate* composition in [Rub00]. Moreover it is shown that for n-degree polynomials the resultant of the composed system can be expressed as a power of the resultant of the *outer* polynomial system, multiplied by powers of leading coefficients of univariate polynomials substituted for variables in the outer system. It is important to point out that the techniques used for deriving resultant formulas in the current paper are different from the techniques used in previous works [CMW95, Jou91, Min03a, Min02, Min01, Min03b, Min03c, Min04, HM02, KS97]. Previous techniques seemed not applicable.

A new resultant formula is derived for multi-univariate composed polynomials where it is known that Cayley-Dixon resultant formulation does not produce any extraneous factors for outer system. The derivation unifies all known related results in the literature [KS97, MW89]. Such systems include n-degree, [KSY94] as well as bivariate corner cut [ZG00] and generalized corner cut systems [Cht03]. Even when extraneous factors are present a similar formula is derived, showing that extraneous factor of outer system will be "amplified" in the extraneous factor of composed system. Hence exploiting composed structure of polynomial system can reduce extraneous factors. Furthermore, it demonstrates that the resultant of a composed system can be effectively calculated by considering only the resultant of the outer system. For practical applications, that is what is needed. Since the complexity of resultant computation is typically determined by the degree (and support) of the polynomial system, resultants of composed systems can be computed much faster by focussing only on the outer system.

Particularly, in case, a resultant matrix is singular for an outer system, the resultant matrix of the composed system is also singular. However, the rank submatrix construction used in [KSY94] (see also [BEM00]) works on the resultant matrix of the composed system as well, giving a projection operator. The rank of the resultant matrix of the composed system can be shown to be $\prod_{i=1}^{n} k_i$ times the rank of the outer polynomial system, where k_i is the degree of the univariate polynomial substituted for the respective variable in the

[1] Resultant depends on algebraic set in consideration in which solutions are sought, [BEM00].

outer system. Furthermore, the extraneous factor arising from the gcd of determinant of the maximal minors of the Dixon matrix of the outer system appears as an extraneous factor in the determinant of the maximal minor of the Dixon matrix of the composed system (but raised to the power $\prod_{i=1}^{n} k_i$). Since the Dixon matrix of the composed system can be larger, there can be additional extraneous factors as well arising from each maximal minor of the Dixon matrix of outer system.

In [CK03, Cht03, FC04], conditions on the support of generic unmixed polynomial systems have been identified for which the Cayley-Dixon formulation generates resultants exactly (without any extraneous factor). The class of polynomial systems for which resultants can be computed exactly can be broadened by composing polynomial systems. More interestingly it can be shown that the composed system of mixed supports can be generated from a unmixed outer system when univariate substitutions are made for distinct variables, thus establishing a class of mixed supports for which Dixon-Cayley construction yields resultants (without extraneous factors). This result about computing resultants of mixed systems without extraneous factors appears to be the first of its kind. Furthermore, it is also possible to compute resultants exactly for other outer polynomial systems obtained by functional decomposition of composed systems whose resultants can be computed exactly. This construction is illustrated using an example. Such an approach for identifying polynomial systems for which resultants can be computed exactly is novel and seems promising.

Below, we first state the main results of the paper. This is followed by a section on preliminaries and notation; the generalized Cayley-Dixon formulation as proposed by Kapur, Saxena and Yang [KSY94] is briefly reviewed. Since the Cayley-Dixon formulation involves two disjoint sets of variables, the bilinear form representation of a polynomial in disjoint sets of variables is useful. In section 2, we discuss how bilinear forms are affected by polynomial operations, particularly when two polynomials are multiplied, a polynomial is composed with other polynomials by substituting variables by polynomials etc. To express these relations among bilinear forms, a series of matrix operations is introduced.

Section 2.3 discusses in detail the case of univariate composed system, i.e., how the Cayley-Dixon resultant computation of a composed system obtained by composing two polynomials in a single variable with another univariate polynomial is related to the various polynomials appearing in the composed system. This construction is later stated in general terms for the multivariate case. The case when the Dixon matrix is singular or non-square is analyzed. Then, in section 3.2 a new resultant formula is derived for n-degree polynomials systems composed in multi-univariate manner. This is followed by a brief section where the example of a mixed composed system is discussed whose resultant can be computed exactly.

We assume that the reader is familiar with the notion of resultant with respect to a given variety (see for example [BEM00]). This notion includes classic resultants like the projective (Macaulay) resultant where the variety is projective space and more recent generalizations like toric resultants where the varieties are suitable toric varieties.

1.1 Main Result

Consider a polynomial system $F = (f_0, f_1, \ldots, f_n)$ with symbolic coefficients, where $F \subset \mathbb{K}[\mathbf{c}][y_1, \ldots, y_n]$ and

$$f_i = \sum_{\alpha \in \mathcal{F}_i} c_{i,\alpha} \mathbf{y}^{\alpha} \qquad \text{for} \quad i = 0, \ldots, n,$$

where $\mathbf{y}^\alpha = y_1^{\alpha_1}, \ldots, y_n^{\alpha_n}$ and $\mathcal{F}_i$ is the set of exponent vectors corresponding to the terms appearing in f_i, also called the *support* of f_i. The list $\mathbf{c}$ consists "other" variables in terms of which polynomial coefficients $c_{i,\alpha} \in \mathbb{K}[\mathbf{c}]$ are defined. They are also sometimes referred as parameters of the polynomial system. A polynomial system is called *generic* if there is no algebraic relation among coefficients $c_{i,\alpha}$ of F.

Let $G = (g_1, \ldots, g_n)$ be another polynomial system in which each g_j, $j = 1, \ldots, n$, is a *univariate* polynomial in x_j, i.e.,

$$g_j = d_{j,k_j} x_j^{k_j} + d_{j,k_j-1} d_j^{k_j-1} + \cdots + d_{j,0}.$$

Let $k = (k_1, \ldots, k_n)$ be the degree vector of G.

It is possible to construct another polynomial system by **composing** F with G, written as $F \circ G$, which is the list of polynomials obtained from the list F of polynomials by replacing each y_j by g_j respectively. The operator $\circ$ is called functional composition on polynomial systems.

The main results of this paper are:

(i) The Dixon matrix $\Theta_{F\circ G}$ of a composed system $F \circ G$ is shown to be a product of three matrices:

$$\Theta_{F\circ G} = A_{\mathrm{L}} \times \mathrm{Diag}_{k_1 \cdots k_n}(\Theta_F) \times A_{\mathrm{R}},$$

where Θ_F is the Dixon matrix of the outer system F and matrix A_{L} as well as A_{R} have triangular shape and contains only polynomials in terms of the coefficients of the polynomials in G. The matrix $\mathrm{Diag}_{k_1 \cdots k_n}(\Theta_F)$ is block diagonal, where Θ_F is repeated $k_1 \cdots k_n$ times along the diagonal.

(ii) If F is a polynomial system for which the determinant of Dixon matrix is $\mathrm{Res}(F)$, then

$$\mathrm{Res}(F \circ G) = d_{1,k_1}^{\epsilon_1} \cdots d_{n,k_n}^{\epsilon_n} \mathrm{Res}(F)^\delta,$$

where ϵ_j's depend on the degrees of G as well as F but δ depends only on the degrees of G.

(iii) It is shown that even if Θ_F is not square or is singular, rank submatrix construction introduced in [KSY94] (see also [BEM00]) also works for composed systems. Then, the projection operator extracted from Θ_F is a factor of the projection operator extracted from $\Theta_{F\circ G}$ raised to the appropriate power; in addition to the leading coefficients d_{j,k_j} of polynomials in G, there are also additional factors introduced in projection operator extracted from $\Theta_{F\circ G}$.

(iv) The resultant of composed n-degree system, with degrees $(m_1, \ldots, m_n)$, is

$$\mathrm{Res}(F \circ G) = \left(d_{1,k_1}^{m_1} \cdots d_{n,k_n}^{m_n} \right)^{\frac{(n+1)!}{2} m_1 \cdots m_2 \, k_1 \cdots k_n} \mathrm{Res}(F)^{k_1 \cdots k_n}.$$

2 Cayley-Dixon Formulation and Bilinear Form

2.1 The Cayley-Dixon Formulation

In [Dix08], Dixon extended Bezout-Cayley's construction for computing the resultant of two univariate polynomials to the bivariate case for three polynomials. Kapur, Saxena and Yang [KSY94] generalized this construction to the multivariate case. The concepts of a Dixon polynomial and a Dixon matrix were introduced. Below, the generalized multivariate Dixon formulation for simultaneously eliminating many variables from a polynomial system and computing its resultant is briefly reviewed. Let $\pi_i(\mathbf{y}^\alpha) = \overline{y}_1^{\alpha_1} \cdots \overline{y}_i^{\alpha_i} y_{i+1}^{\alpha_{i+1}} \cdots y_n^{\alpha_n}$, where $i \in \{0, 1, \ldots, n\}$, and $\overline{y}_i$'s are new variables;[2] $\pi_0(\mathbf{y}^\alpha) = \mathbf{y}^\alpha$. π_i is extended to polynomials in a natural way as: $\pi_i(f_j(y_1, \ldots, y_n)) = f_j(\overline{y}_1, \ldots, \overline{y}_i, y_{i+1}, \ldots, y_n)$.

Definition 2.1 *Given a n-variate polynomial system $F = (f_0, f_1, \ldots, f_n)$, where polynomial $f_i \subset \mathbb{K}[\mathbf{c}][y_1, \ldots, y_n]$, define its* ***Dixon polynomial*** *as*

$$\theta(F) = \prod_{i=1}^{n} \frac{1}{\overline{y}_i - y_i} \det \begin{pmatrix} \pi_0(f_0) & \pi_0(f_1) & \cdots & \pi_0(f_n) \\ \pi_1(f_0) & \pi_1(f_1) & \cdots & \pi_1(f_n) \\ \vdots & \vdots & \ddots & \vdots \\ \pi_n(f_0) & \pi_n(f_1) & \cdots & \pi_n(f_n) \end{pmatrix}.$$

Hence, $\theta(f_0, f_1, \ldots, f_n) \in \mathbb{K}[\mathbf{c}][y_1, \ldots, y_n, \overline{y}_1, \ldots, \overline{y}_n]$, where $\overline{y}_1, \ldots, \overline{y}_n$ are new variables. The matrix above is called the cancellation *matrix.*

The order in which original variables in $y_1, \ldots, y_n$ are replaced by new variables in $\overline{y}_1, \ldots, \overline{y}_n$ is significant in the sense that the computed Dixon polynomial can be different for two different orderings. See [Dix08, KSY94, Sax97, Cht03, BEM00].

Definition 2.2 *A Dixon polynomial $\theta(f_0, f_1, \ldots, f_n)$ can be written in the bilinear form as*

$$\theta(F) \;=\; \overline{Y}^T \times \Theta_F \times Y,$$

where $\overline{Y} = [\overline{\mathbf{y}}^{\beta_1}, \ldots, \overline{\mathbf{y}}^{\beta_k}]$ and $Y = [\mathbf{y}^{\alpha_1}, \ldots, \mathbf{y}^{\alpha_l}]$ are column vectors. The $k \times l$ matrix Θ is called the ***Dixon matrix****.*

Each entry in Θ is a polynomial in the coefficients of the polynomials in F.

As shown in [KSY94] and [BEM00], Θ is a resultant matrix. However, it can be singular especially for nongeneric polynomial systems. In such a case, the resultant is extracted from the determinant of some maximal minor of Θ; this determinant is called a *projection operator* [KSY94, Sax97, BEM00].

2.2 Operations on Bilinear Forms

It is easy to see that a multivariate polynomial in terms of two disjoint sets of variables, e.g., the Dixon polynomial above, can be represented in a *bilinear form*. For analyzing how the functional composition of two polynomial systems affects the Dixon polynomials and Dixon

[2] For n-tuple $\mathbf{x} = (x_1, \ldots, x_n)$, the notation $\mathbf{x}^\alpha = (x_1, \ldots, x_n)^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ for $\alpha \in \mathbb{N}^n$.

matrices of the polynomial systems, bilinear form representations turn out to be useful. Below, we discuss various polynomial operations and their effect on bilinear forms.

A bilinear form of a polynomial p in two disjoint sets of variables is expressed as a matrix, post and pre-multiplied by monomial vectors. That is

$$p(x_1, \ldots, x_k, \overline{x}_1, \ldots, \overline{x}_l) = \sum_{\alpha,\beta} p_{\alpha,\beta}\, \overline{\mathbf{x}}^{\alpha} \mathbf{x}^{\beta} \;=\; \overline{X_p}^{\,T} \times M_p \times X_p,$$

where $\overline{X_p}$ and X_p are vectors with entries being monomials in terms of variables $\{\overline{x}_1, \ldots, \overline{x}_l\}$ and $\{x_1, \ldots, x_k\}$, respectively. M_p is a matrix with the coefficients $p_{\alpha,\beta}$ of power products in p as its entries.

For example, let $p = \overline{xy}x^2 + 2\overline{xy}y - 3x^2$, then

$$p = \overline{X_p}^{\,T} \times M_p \times X_p = [\ \overline{xy}\ \ 1\ ] \times \begin{bmatrix} 1 & 2 \\ -3 & 0 \end{bmatrix} \times \begin{bmatrix} x^2 \\ y \end{bmatrix}. \tag{1}$$

The matrix M_p in the above definition depends on the monomial ordering used. We will assume a total degree ordering on power products, and state explicitly if it is otherwise. Also, implicit in the above definition of matrix M_p are the row labels $\overline{X}_p$, and column labels X_p.

Let $\mathcal{P}$ be the ordered set of the exponent vectors corresponding to X_p; $\mathcal{P}$ is also called the **support** of the polynomial p w.r.t variables $\{x_1, \ldots, x_k\}$. Similarly, let $\overline{\mathcal{P}}$ be the support of p w.r.t. variables $\{\overline{x}_1, \ldots, \overline{x}_l\}$ ($\overline{\mathcal{P}}$ is also the ordered set of the exponent vectors corresponding to $\overline{X}_p$). For the above example, $\mathcal{P} = [(0,1),(2,0)]$ and $\overline{\mathcal{P}} = [(0,0),(1,1)]$.

As stated above, the Dixon polynomial can be conveniently represented in bilinear form using the original variables and the new variables, highlighting the Dixon matrix. Let Δ_F be the support of polynomial $\theta(F)$ in terms of variables $\mathbf{y}$ and $\overline{\Delta}_F$ in terms of variables $\overline{\mathbf{y}}$.

Consider the following matrix construction operators, $\mathrm{RowStack}_{\alpha\in\mathcal{C}}(N_\alpha)$, $\mathrm{ColStack}_{\alpha\in\mathcal{C}}(N_\alpha)$ and $\mathrm{Diag}_{\alpha\in\mathcal{C}}(N_\alpha)$ respectively denoting the (block)-row and -column vector/matrix with its (block) indices taken from support $\mathcal{C}$ and the block-diagonal matrix with as many blocks as elements in the support $\mathcal{C}$. See figure 1.

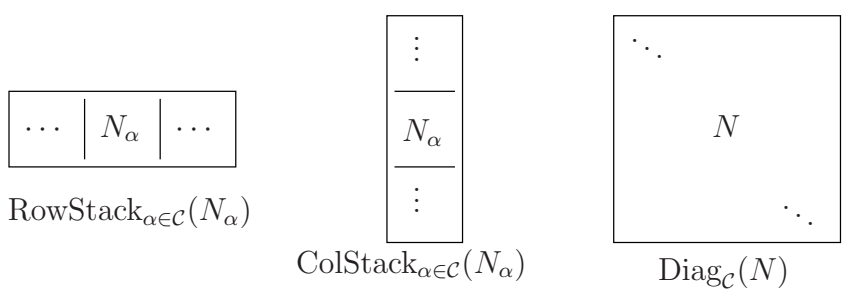

Figure 1: Matrix constructors.

We will express bilinear polynomial product and composition in terms of above operators.

Subsequently consider two polynomials p and q in bilinear forms along with their associated power products. I.e.,

$$p = \overline{X_p}^{\,T} \times M_p \times X_p, \quad \text{and} \quad q = \overline{X_q}^{\,T} \times M_q \times X_q.$$

The respective supports of p in $\mathbf{x}$ and $\overline{\mathbf{x}}$ are $\mathcal{P}, \overline{\mathcal{P}}$; similarly, the respective supports of q are $\mathcal{Q}, \overline{\mathcal{Q}}$. Let $\mathcal{P}+\mathcal{Q}$ stand for the Minkowski sum of supports $\mathcal{P}$ and $\mathcal{Q}$.

2.2.1 Polynomial Product in Terms of Bilinear Forms

Subsequently, we derive bilinear matrix form of the polynomial product of two polynomials in terms of their bilinear matrix forms. For this purpose first define so-called "left" and "right" operators on bilinear forms.

Definition 2.3 *Given two polynomials p and q admitting bilinear form, consider the following polynomial products*

$$\begin{array}{rcl} p' & = & p \cdot \sum_{\overline{e}_q \in \overline{\mathcal{Q}}} \overline{\mathbf{x}}^{\overline{e}_q} \mathbf{z}^{\overline{e}_q} = \overline{X}_{p'} \times M_{p'} \times X_{p'}, \\ q' & = & q \cdot \sum_{e_p \in \mathcal{P}} \overline{\mathbf{z}}^{e_p} \mathbf{x}^{e_p} = \overline{X}_{q'} \times M_{q'} \times X_{q'}, \end{array} \qquad \textit{and}$$

where $\overline{z}_1, \ldots, \overline{z}_n$ and $z_1, \ldots, z_n$ are new variables. Define two matrix operators

$$\mathrm{L}_{\overline{\mathcal{Q}}}(M_p) = M_{p'} \qquad \textit{and} \qquad \mathrm{R}_{\mathcal{P}}(M_q) = M_{q'},$$

where columns of $M_{p'}$ are ordered first by $\{z_1, \ldots, z_n\}$ and then by $\{x_1, \ldots, x_n\}$, and rows of $M_{q'}$ first by $\{\overline{x}_1, \ldots, \overline{x}_n\}$ and then by $\{\overline{z}_1, \ldots, \overline{z}_n\}$, where orders used, w.r.t. $\{z_1, \ldots, z_n\}$ and $\{\overline{x}_1, \ldots, \overline{x}_n\}$ as well as $\{\overline{x}_1, \ldots, \overline{x}_n\}$ and $\{\overline{z}_1, \ldots, \overline{z}_n\}$ are same monomial orderings.

The above matrix operators are defined in such a way that matrix multiplication would coincide with polynomial multiplication. New variables $\overline{z}_1, \ldots, \overline{z}_n$ and $z_1, \ldots, z_n$ are auxiliary variables for creating block matrix structure, as well as ensuring that resulting matrix rows and columns are in matching order. Notice that row indices of $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$ are $\overline{\mathcal{P}}+\overline{\mathcal{Q}}$ and column indices are $\overline{\mathcal{Q}} \times \mathcal{P}$, coming from monomials $\mathbf{z}^{\overline{e}_q}\mathbf{x}^{e_p}$ for $\overline{e_q} \in \overline{\mathcal{Q}}$ and $e_p \in \mathcal{P}$. Similarly, row indices of $\mathrm{R}_{\mathcal{P}}(M_q)$ are $\overline{\mathcal{Q}} \times \mathcal{P}$ (coming from monomials $\overline{\mathbf{x}}^{e_q}\overline{\mathbf{z}}^{e_p}$) and column indices are $\mathcal{P}+\mathcal{Q}$.

Matrix $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$ is quite sparse and its entries are either 0 or coefficients of polynomial p. In fact the entry of $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$ indexed by row $\overline{\mathbf{x}}^{\overline{e}_p+\overline{e}_q}$ and column $\mathbf{z}^{\overline{e}_q}\mathbf{x}^{e_p+e_q}$ is equal to $p_{\overline{e}_p,e_p}$. All other entries are 0. Also it has block matrix structure

$$\mathrm{L}_{\overline{\mathcal{Q}}}(M_p) = \mathrm{RowStack}_{\alpha \in \overline{\mathcal{Q}}}(N_\alpha \times M_p)\,,$$

where N_α is a matrix which adds zero rows to M_p (depending on α, $\overline{\mathcal{Q}}$ and $\overline{\mathcal{P}}$). $\mathrm{R}_{\mathcal{P}}(M_q)$ also admits a similar block decomposition.

Example 2.1 [Left and right multiplication operators] Let $p = a_1\,\overline{xy}x^2 + a_2\,\overline{xy}y + a_3\,x^2$ and $q = b_1\,\overline{xy}x^3y + b_2\,\overline{x}x^3y + b_3x^3y$, then

$$p = \begin{pmatrix} \overline{xy} \\ 1 \end{pmatrix}^T \times \begin{pmatrix} a_1 & a_2 \\ a_3 & 0 \end{pmatrix} \times \begin{pmatrix} x^2 \\ y \end{pmatrix} \quad \text{and} \quad q = \begin{pmatrix} \overline{xy} \\ \overline{x} \\ 1 \end{pmatrix}^T \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \times \left(x^3y\right).$$

Bilinear supports of these polynomials are $\overline{\mathcal{P}} = [(1,1),(0,0)]$, $\mathcal{P} = [(2,0),(0,1)]$, $\overline{\mathcal{Q}} = [(1,1),(1,0),(0,0)]$ and $\mathcal{Q} = [(3,1)]$. The left operator on M_p is

$$p' = p \cdot \sum_{\overline{e}_q \in \overline{\mathcal{Q}}} \overline{\mathbf{x}}^{\overline{e}_q} \mathbf{z}^{\overline{e}_q} = \begin{pmatrix} \overline{x}^2\overline{y}^2 \\ \overline{x}^2\overline{y} \\ \overline{xy} \\ \overline{x} \\ 1 \end{pmatrix}^T \times \underbrace{\left(\begin{array}{cc|cc|cc} a_1 & a_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_1 & a_2 & 0 & 0 \\ a_3 & 0 & 0 & 0 & a_1 & a_2 \\ 0 & 0 & a_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_3 & 0 \end{array}\right)}_{\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)} \times \begin{pmatrix} z_1 z_2\, x^2 \\ z_1 z_2\, y \\ z_1\, x^2 \\ z_1\, y \\ x^2 \\ y \end{pmatrix},$$

and

$$q' = q \cdot \sum_{e_p \in \mathcal{P}} \overline{\mathbf{z}}^{e_p}\, \overline{\mathbf{x}}^{e_p} = \begin{pmatrix} xy\, z_1^2 \\ xy\ \ z_2 \\ x\ \ z_1^2 \\ x\ \ z_2 \\ 1\ \ z_1^2 \\ 1\ \ z_2 \end{pmatrix}^T \times \underbrace{\begin{pmatrix} b_1 & 0 \\ 0 & b_1 \\ b_2 & 0 \\ 0 & b_2 \\ b_3 & 0 \\ 0 & b_3 \end{pmatrix}}_{\mathrm{R}_{\mathcal{P}}(M_q)} \times \begin{pmatrix} x^5 y \\ x^3 y^2 \end{pmatrix}.$$

□

Using the above operators, we can express the bilinear form of a polynomial product as matrix multiplication, as shown in Figure 2.

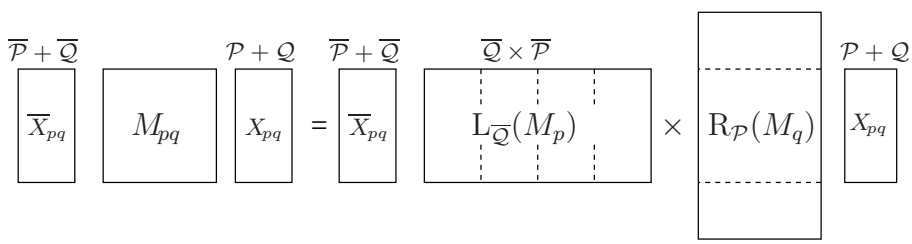

Figure 2: Left and right multiplication operators.

Lemma 2.1

$$M_{p\,q} = \mathrm{L}_{\overline{\mathcal{Q}}}(M_p) \times \mathrm{R}_{\mathcal{P}}(M_q)\,.$$

Proof: Directly from the polynomial product of polynomials p and q,

$$(M_{p\,q})_{\alpha,\beta} = \sum_{\substack{\alpha = \overline{e}_p + \overline{e}_q, \\ \beta = e_p + e_q}} p_{\overline{e}_p, e_p} q_{\overline{e}_q, e_q}$$

for $\overline{e}_p \in \overline{\mathcal{P}}$, $\overline{e}_q \in \overline{\mathcal{Q}}$, $e_p \in \mathcal{P}$ and $e_q \in \mathcal{Q}$. On the other hand

$$\begin{aligned} \left(\mathrm{L}_{\overline{\mathcal{Q}}}(M_p) \times \mathrm{R}_{\mathcal{P}}(M_q)\right)_{\alpha,\beta} &= \mathrm{Row}_\alpha\left(\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)\right) \cdot \mathrm{Col}_\beta\left(\mathrm{R}_{\mathcal{P}}(M_q)\right) \\ &= \sum_{\substack{\overline{e}_q \in \overline{\mathcal{Q}}, \\ e_p \in \mathcal{P}}} \mathrm{coeff}_{\overline{\mathbf{x}}^\alpha \mathbf{z}^{\overline{e}_q} \mathbf{x}^{e_p}}(p') \cdot \mathrm{coeff}_{\mathbf{x}^\beta \overline{\mathbf{z}}^{e_p} \overline{\mathbf{x}}^{\overline{e}_q}}(q'), \end{aligned}$$

but

$$\mathrm{coeff}_{\overline{\mathbf{x}}^{\alpha}\mathbf{z}^{\overline{e}_q}\mathbf{x}^{e_p}}(p') = \begin{cases} p_{\overline{e}_p,e_p} & \text{if } \alpha = \overline{e}_p + \overline{e}_q \\ 0 & \text{otherwise} \end{cases},$$

and

$$\mathrm{coeff}_{\mathbf{x}^{\beta}\overline{\mathbf{z}}^{e_p}\overline{\mathbf{x}}^{\overline{e}_q}}(q') = \begin{cases} q_{\overline{e}_q,e_q} & \text{if } \beta = e_p + e_q \\ 0 & \text{otherwise} \end{cases}.$$

Therefore

$$\sum_{\substack{\overline{e}_q\in\overline{\mathcal{Q}},\\ e_p\in\mathcal{P}}} \mathrm{coeff}_{\overline{\mathbf{x}}^{\alpha}\mathbf{z}^{\overline{e}_q}\mathbf{x}^{e_p}}(p') \cdot \mathrm{coeff}_{\mathbf{x}^{\beta}\overline{\mathbf{z}}^{e_p}\overline{\mathbf{x}}^{\overline{e}_q}}(q')$$

$$= \sum_{\substack{\alpha=\overline{e}_p+\overline{e}_q,\\ \beta=e_p+e_q}} \mathrm{coeff}_{\overline{\mathbf{x}}^{\alpha}\mathbf{z}^{\overline{e}_q}\mathbf{x}^{e_p}}(p') \cdot \mathrm{coeff}_{\mathbf{x}^{\beta}\overline{\mathbf{z}}^{e_p}\overline{\mathbf{x}}^{\overline{e}_q}}(q') = \sum_{\substack{\alpha=\overline{e}_p+\overline{e}_q,\\ \beta=e_p+e_q}} p_{\overline{e}_p,e_p} q_{\overline{e}_q,e_q},$$

where exponents are chosen $e_p \in \mathcal{P}$, $\overline{e}_p \in \overline{\mathcal{P}}$, $q \in \overline{\mathcal{Q}}$ and $\overline{e}_q \in \overline{\mathcal{Q}}$ □

One useful property of L operator is that application on matrix product results application on one of the matrices times a block diagonal matrix of the other factor.

Lemma 2.2 *Given a product of two matrices $A \times B$*

$$\mathrm{L}_{\mathcal{P}}(A \times B) \quad = \quad \mathrm{L}_{\mathcal{P}}(A) \times \mathrm{Diag}_{\mathcal{P}}(B)\,,$$

where the product of above matrices is assumed to conform to row and column labels.

Same holds for operator R, but one has to take into account difference in the row order. Proof: By definition,

$$\begin{aligned} \mathrm{L}_{\mathcal{P}}(A \times B) &= \mathrm{RowStack}_{\alpha\in\mathcal{P}}(N_\alpha \times (A \times B)) = \mathrm{RowStack}_{\alpha\in\mathcal{P}}((N_\alpha \times A) \times B) \\ &= \mathrm{RowStack}_{\alpha\in\mathcal{P}}(N_\alpha \times A) \times \mathrm{Diag}_{\mathcal{P}}(B) = \mathrm{L}_{\mathcal{P}}(A) \times \mathrm{Diag}_{\mathcal{P}}(B)\,. \end{aligned}$$

□

The following is simple but useful observation is used in proving main result.

Proposition 2.1 *For polynomials $p(\overline{x}_1, \ldots, \overline{x}_k, x_1, \ldots, x_l)$ and $q(\overline{y}_1, \ldots, \overline{y}_k, x_1, \ldots, x_l)$ which are defined in terms of different sets of variables,*

$$\mathrm{L}_{\overline{\mathcal{Q}}}(M_p) = \mathrm{Diag}_{\overline{\mathcal{Q}}}(M_p)\,.$$

Proof: By definition

$$p' = p \cdot \sum_{\overline{e}_q\in\overline{\mathcal{Q}}} \mathbf{z}^{\overline{e}_q}\overline{\mathbf{y}}^{\overline{e}_q} = \overline{X}_{p'} \times \mathrm{L}_{\overline{\mathcal{Q}}}(M_p) \times X_{p'}.$$

Since polynomial p does not have terms in variables $\overline{y}_1, \ldots, \overline{y}_k, z_1, \ldots, z_k$, the bilinear form of p, that is matrix M_p is repeated $|\overline{\mathcal{Q}}|$ times along the diagonal in $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$. □

2.2.2 Bilinear Form under Composition with Univariate Polynomials

To express the effect of substituting a univariate polynomial g_i in x_i for y_i in f_j, the following operator is needed. This operator is then used below to express how bilinear forms are affected by functional composition of two polynomial systems.

Definition 2.4 *Given a support $\mathcal{P}$ and the set of univariate polynomials $G = (g_1, \ldots, g_n)$, where each g_i is in x_i, let*

$$s = \sum_{\alpha \in \mathcal{P}} \overline{\mathbf{x}}^{\alpha} G^{\alpha} = \overline{X}_s \times M_s \times X_s = \overline{X}_s \times \mathrm{S}_{\mathcal{P}}(G) \times X_s,$$

where $G^{\alpha} = \prod_{i=1}^{n} g_i^{\alpha_i}$ and define operator $\mathrm{S}_{\mathcal{P}}(G) = M_s$.

$\mathrm{S}_{\mathcal{P}}(G)$ is thus the matrix whose rows are indexed by $\mathcal{P}$ and whose columns are indexed by the union over $\alpha \in \mathcal{P}$ of the supports of $\prod_{j=1}^{n} g_j^{\alpha_j}$. Note that monomial vector, with support $\mathcal{P}$ composed with G can be expressed as

$$Y_p \circ G = \mathrm{S}_{\mathcal{P}}(G) \times X_s,$$

where X_s is union of all monomials in G^{α} for all $\alpha \in \mathcal{P}$. Matrix $\mathrm{S}_{\mathcal{P}}(G)$ is also very sparse. More specifically

$$\left(\mathrm{S}_{\overline{\mathcal{P}}}(G)\right)_{\overline{e}_s, e_s} = \begin{cases} (d_{1,k_1}, \ldots, d_{n,k_n})^{\overline{e}_s} & \text{if } (e_s)_i = k_i(\overline{e}_s)_i, \ \forall i, \\ 0 & \text{if } \exists\, i \text{ s.t. } k_i(\overline{e}_s)_i < (e_s)_i, \\ s_{\overline{e}_s, e_s} & \text{otherwise, i.e. if } \forall\, i, \ k_i(\overline{e}_s)_i > (e_s)_i. \end{cases} \tag{2}$$

In particular, in univariate case if $\overline{\mathcal{P}} = [m-1, \ldots, 0]$ then support of X_s is $[(m-1)k, \ldots, 0]$, and

$$\left(\mathrm{S}_{\overline{\mathcal{P}}}(G)\right)_{i,j} = \begin{cases} d_k^i & \text{if } j = k \cdot i, \\ 0 & \text{if } j > k \cdot i, \\ s_{i,j} & \text{otherwise,} \end{cases} \tag{3}$$

for $i \in [m-1, \ldots, 0]$ and $j \in [(m-1)k, \ldots, 0]$.

Next we illustrate the operator $\mathrm{S}_{\mathcal{P}}(G)$ in the bivariate setting.

Example 2.2 [Operator $\mathrm{S}_{\mathcal{P}}(G)$] Let $\mathcal{P} = [(2,0), (1,1), (0,1), (0,0)]$, $g_1 = a_2 x_1^2 + a_1 x_1 + a_0$ and $g_2 = b_1 x_2 + b_0$. Then

$$\mathrm{S}_{\mathcal{P}}(G) = \begin{pmatrix} g_1^2 \\ g_1 g_2 \\ g_2 \\ 1 \end{pmatrix} = \begin{array}{c} \begin{array}{cccccccc} (4,0) & (3,0) & (2,1) & (2,0) & (1,1) & (1,0) & (0,1) & (0,0) \end{array} \\ \begin{pmatrix} a_2^2 & 2a_2a_1 & 0 & 2a_2a_0 + a_1^2 & 0 & 2a_1a_0 & 0 & a_0^2 \\ 0 & 0 & a_2b_1 & 0 & a_1b_1 & a_1b_0 & a_0b_1 & a_0b_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_1 & b_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array},$$

where top row shows support of the X_s, i.e. column labels of $\mathrm{S}_{\mathcal{P}}(G)$. $\square$

The following lemma states how the bilinear form of a polynomial p in variables sets $\mathbf{y}$ and $\overline{\mathbf{y}}$ is affected when for each $i = 1, \ldots, n$, y_i and $\overline{y}_i$ are, respectively, substituted by g_i and $\overline{g}_i$, where g_i is a univariate polynomial in x_i and $\overline{g}_i$ is the univariate polynomial g_i in which x_i is uniformly replaced by $\overline{x}_i$. Also let $\overline{G} = (\overline{g}_1, \ldots, \overline{g}_n)$.

Lemma 2.3 *Let p be a polynomial in the variables $\overline{\mathbf{y}}$, $\mathbf{y}$, and G a set of univariate polynomials g_i in variable x_i, for $i = 1, \ldots, n$. Then*

$$M_{p \circ (\overline{G},\, G)} \;=\; \mathrm{S}_{\overline{\mathcal{P}}}\big(\overline{G}\big)^{\mathrm{T}} \;\times\; M_p \;\times\; \mathrm{S}_{\mathcal{P}}(G)\,,$$

where $\overline{G} = (\overline{g}_1, \ldots, \overline{g}_n)$, and $\overline{g}_i = g_i(\overline{x}_i)$.

Proof: Since $p = \overline{Y}_p \times M_p \times Y_p$, we have

$$p \circ (\overline{G},\, G) \;=\; (\overline{Y}_p^T \circ \overline{G}) \times M_p \times (Y_p \circ G)$$

and $Y_p \circ G = \mathrm{S}_{\mathcal{P}}(G) \times X_s$ by definition. □

2.2.3 Properties of operators L, R and S

Consider the following slight generalization of triangular matrices to non-square matrices.

Definition 2.5 *For a $k \times l$ matrix M, let t_i be the column index of first non-zero entry in row i. M is said to be (upper) step-triangular if $t_{i+1} > t_i$ for all $i = 1, \ldots, k-1$. The first non-zero entries in each row are called diagonal entries, which make up the step diagonal of matrix M.*

Note that if matrix M is square and step-triangular then it is triangular. A matrix is lower step triangular if its transpose is upper step triangular.

It is not hard to see that for any $\mathcal{P} \subset \mathbb{N}^d$, matrix $\mathrm{S}_{\mathcal{P}}(G)$ is upper step triangular, by description of matrix entries in equation (2).

Very useful property of operators L and S is that in combination they also produce step-triangular matrices, for a special support $\overline{\mathcal{Q}}$.

Proposition 2.2 *For polynomial $q = \prod_{i=1}^{n}(g_i - \overline{g}_i)/(x_i - \overline{x}_i)$ the bilinear form matrix M_q is (anti) triangular of size $k \times k$, and more over diagonal entries are $d_{1,k_1} \cdots d_{n,k_n}$.*

Proof: In the polynomial

$$\frac{g_i - \overline{g}_i}{x_i - \overline{x}_i} = \sum_{j=0}^{k_i} d_{i,j}\, \frac{x_i^j - \overline{x}_i^j}{x_i - \overline{x}_i} \;=\; \sum_{j=1}^{k_i} d_{i,j} \sum_{l=0}^{j-1} \overline{x}_i^l\, x_i^{j-l-1},$$

monomials $\overline{x}^j\, x^l$ for $j + l \geq k_i$ are not present. Moreover when $j + l = k_i - 1$, the coefficient of $\overline{x}_i^j\, x_i^l$ is just d_{i,k_i}.

Since q is a product of such polynomials, which are defined in terms of different variables, we can characterize coefficients of q and hence entries of M_q as

$$q_{\overline{e}_q, e_q} = \begin{cases} d_{1,k_1} \cdots d_{n,k_n} & \text{if } \overline{e}_q + e_q = k - 1 \\ 0 & \text{if } \exists\, i \text{ s.t. } (\overline{e}_q)_i + (e_q)_i > k_i - 1 \end{cases}$$

It can be seen that under lexicographical order on variables $x_1, \ldots, x_n$ and $\overline{x}_1, \ldots, \overline{x}_n$ the matrix M_q will be anti-triangular, i.e. 0 above the anti-diagonal.□

Notice that the support $\overline{\mathcal{Q}}$ of q in the above proposition, in terms of variables $\overline{x}_1, \ldots, \overline{x}_n$ is

$$\overline{e}_q \in \overline{\mathcal{Q}} \qquad \text{iff} \qquad 0 \leq (\overline{e}_q)_i < k_i \text{ for all } i = 1, \ldots, n.$$

Using above notions we can show that operators L and S in combination produce step triangular matrices, an important property used in derivation of main results.

Proposition 2.3 *Let $\overline{\mathcal{Q}}$ be a support of $\prod_{i=1}^{n} \frac{g_i - \overline{g}_i}{x_i - \overline{x}_i}$, then for any support $\mathcal{P}$, the matrix $\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\mathcal{P}}\big(\overline{G}\big)^{\mathrm{T}}\Big)$ is (lower) step triangular (after column reordering), moreover entry in column $\overline{e}_q e_p$ and row α is*

$$\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\mathcal{P}}\big(\overline{G}\big)^{\mathrm{T}}\Big)_{\alpha, \overline{e}_q . e_p} = \begin{cases} (d_{1,k_1}, \ldots, d_{n,k_n})^{e_p} & \text{if } \alpha = \overline{e}_s + \overline{e}_q \text{ and } (\overline{e}_s)_i = k_i (e_p)_i, \\ \Big(\mathrm{S}_{\mathcal{P}}\big(\overline{G}\big)^{\mathrm{T}}\Big)_{\overline{e}_s, e_p} & \text{if } \alpha = \overline{e}_s + \overline{e}_q \text{ and } \forall i,\ (\overline{e}_s)_i < k_i (e_p)_i, \\ 0 & \text{otherwise} \end{cases}$$

i.e. in every column first non-zero entry is product of leading coefficients of G, and all these leading non-zero entries are in different rows.

Proof: Note that columns of $\mathrm{S}_{\mathcal{P}}\big(\overline{G}\big)^{\mathrm{T}}$ are labelled by $\mathcal{P}$ and rows by $\overline{X}_s$ which is set of all monomials in $\overline{G}^{\alpha} = \overline{g}_1^{\alpha_1} \cdots \overline{g}_n^{\alpha_n}$ for all $\alpha \in \mathcal{P}$. (Since we are considering transpose of $\mathrm{S}_{\mathcal{P}}\big(\overline{G}\big)^{\mathrm{T}}$, the $\overline{X}_s$ and X_s are switched as in the definition.)

Consider the following polynomial

$$s = \overline{X}_s \times \mathrm{S}_{\mathcal{P}}\big(\overline{G}\big)^{\mathrm{T}} \times X_s, \qquad \text{and} \qquad s' = s \cdot \sum_{\overline{e}_q \in \overline{\mathcal{Q}}} \mathbf{z}^{\overline{e}_q}\, \overline{\mathbf{x}}^{\overline{e}_q},$$

as in definition 2.3 of $\mathrm{L}_{\overline{\mathcal{Q}}}(M_p)$. We already now that

$$\mathrm{coeff}_{\overline{\mathbf{x}}^{\alpha} \mathbf{z}^{\overline{e}_q} \mathbf{x}^{e_s}}(s') = \begin{cases} s_{\overline{e}_s, e_s} & \text{if } \alpha = \overline{e}_s + \overline{e}_q, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Since support of s is $\mathcal{P}$, we will use labels e_p instead of e_s. Putting equations (2) and (4) together we get

$$\mathrm{coeff}_{\overline{\mathbf{x}}^{\alpha} \mathbf{z}^{\overline{e}_q} \mathbf{x}^{e_p}}(s') = \begin{cases} (d_{1,k_1}, \ldots, d_{n,k_n})^{e_p} & \text{if } \alpha = \overline{e}_s + \overline{e}_q \text{ and } (\overline{e}_s)_i = k_i (e_p)_i, \\ s_{\overline{e}_s, e_p} & \text{if } \alpha = \overline{e}_s + \overline{e}_q \text{ and } \forall i,\ (\overline{e}_s)_i < k_i (e_p)_i, \\ 0 & \text{otherwise.} \end{cases}$$

□

Next we illustrate how the left multiplication operator interacts with the operator S.

Example 2.3 [L and S interaction] Consider a bivariate support composed with univariate system G from example 2.2, where $\mathrm{S}_{\overline{\mathcal{P}}}\big(\overline{G}\big)$ is shown for $\overline{\mathcal{P}} = [(2,0), (1,1), (0,1), (0,0)]$. Let

$\overline{\mathcal{Q}} = [(1,0),(0,0)]$, by previous proposition. Then $\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\mathcal{P}}}\big(\overline{G}\big)^{\mathrm{T}}\Big)$ is a matrix with rows and columns of the following table.

	$\overline{\mathcal{Q}}$	$(1,0)$				$(0,0)$			
	$\overline{\mathcal{P}}$	$(2,0)$	$(1,1)$	$(0,1)$	$(0,0)$	$(2,0)$	$(1,1)$	$(0,1)$	$(0,0)$
	$(5,0)$	a_2^2	0	0	0	0	0	0	0
	$(4,0)$	$2a_2a_1$	0	0	0	a_2^2	0	0	0
$\mathcal{P}$	$(3,1)$	0	a_2b_1	0	0	0	0	0	0
	$(3,0)$	$2a_2a_0+a_1^2$	0	0	0	$2a_2a_1$	0	0	0
$+$	$(2,1)$	0	a_1b_1	0	0	0	a_2b_1	0	0
	$(2,0)$	$2a_1a_0$	a_1b_0	0	0	$2a_2a_0+a_1^2$	0	0	0
$\overline{\mathcal{Q}}$	$(1,1)$	0	a_0b_1	b_1	0	0	a_1b_1	0	0
	$(1,0)$	a_0^2	a_0b_0	b_0	1	$2a_1a_0$	a_1b_0	0	0
	$(0,1)$	0	0	0	0	0	a_0b_1	b_1	0
	$(0,0)$	0	0	0	0	a_0^2	a_0b_0	b_0	1

Note that there is an order on the columns of the matrix, so that matrix is step triangular. Columns of $\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\mathcal{P}}}\big(\overline{G}\big)^{\mathrm{T}}\Big)$ are ordered by $\{z_1,\dots,z_n\}$ and then by $\{x_1,\dots,x_n\}$. The order which makes above matrix step triangular is lexicographical order $[x_1,z_1,x_2,z_2,\dots,x_n,z_n]$. $\square$

In particular, in univariate case, when $G=(g)$, g of degree k and $\mathcal{P}=[0,\dots,m-1]$ then $\overline{\mathcal{Q}}=[0,\dots,k-1]$, the row support of polynomial s is $[0,\dots,k(m-1)]$, then matrix $\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\mathcal{P}}\big(\overline{G}\big)^{\mathrm{T}}\Big)$ has km rows labelled by $[0,\dots,k(m-1)]+[0,\dots,k-1]$ and km columns labelled by $[0,\dots,k-1]\times[0,\dots,m-1]$. That is matrix is square, and more over

$$\Big(\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\mathcal{P}}\big(\overline{G}\big)^{\mathrm{T}}\Big)\Big)_{i,j.l} = \begin{cases} 0 & \text{if } i<j, \\ d_k^l & \text{if } i-j=kl, \\ \mathrm{S}_{\mathcal{P}}\big(\overline{G}\big)^{\mathrm{T}}_{i-j,l} & \text{if } i-j<kl, \\ 0 & \text{if } i-j>kl, \end{cases} \tag{5}$$

for $i\in[0,\dots,km-1]$, $j\in[0,\dots,k-1]$ and $l\in[0,\dots,m-1]$. It is easy to see that for fixed l, we get lower triangular submatrix of size $k\times k$. In fact running indices in $(i,l.j)$ order will result in a triangular matrix, with diagonal entries d_k^l.

In the rest of the paper we use the above operators in expressing the manipulations of bilinear forms of various polynomials arising in the Cayley-Dixon construction, to show that Dixon matrix of composed system can be decomposed as a matrix product.

Particularly, the next section illustrates when outer system F consists of two univariate polynomials in y and G consists of a single univariate polynomial in x.

2.3 Case Study: Cayley-Dixon construction for Univariate composed System

The purpose of this section is to illustrate the use of the operators L, R and S and to show in great detail how they can be used to derive a resultant formula for the composed polynomials $F\circ G$ in a special case. The special case considered in this section is when F consists of two

univariate polynomials. As the reader will see, this derivation proceeds by relating the Dixon matrix of $F \circ G$ to the Dixon matrix of F.

Consider a general univariate polynomial system $F = (f_0, f_1)$, where

$$f_0 = a_{m_0} y^{m_0} + \cdots + a_1 y + a_0, \qquad \text{and} \qquad f_1 = b_{m_1} y^{m_1} + \cdots + b_1 y + b_0,$$

and let $m = \max(m_0, m_1)$. Let $G = (g)$, where

$$g = d_k x^k + d_{k-1} x^{k-1} + \cdots + d_2 x^2 + d_1 x + d_0.$$

McKay and Wang [MW89] showed that the resultant of the composed polynomials $F \circ G$, can be factored as follows:

$$\operatorname{Res}(f_0 \circ g,\ f_1 \circ g) = d_k^{m_0 m_1 k} \operatorname{Res}(f_0,\ f_1)^k. \tag{6}$$

Subsequently we compute the same formula using the matrix techniques introduced in the current paper. This computation is much longer than the short proof by McKay and Wang. However, as the reader will see in the next section, this computation can naturally be generalized to study Dixon matrices for multivariate polynomials. It seems that the techniques used by McKay and Wang cannot be generalized for this purpose.

The **Beźout-Cayley Construction** for the composed polynomials $f_0 \circ g$ and $f_1 \circ g$ is done as follows. Let $\overline{g}$ denote the polynomial obtained from g by replacing x with $\overline{x}$. We get the Bézout polynomial of the composed system

$$\begin{aligned} \theta\,(F \circ G) &= \frac{\det \begin{pmatrix} f_0 \circ g & f_1 \circ g \\ f_0 \circ \overline{g} & f_1 \circ \overline{g} \end{pmatrix}}{x - \overline{x}} = \frac{\det \begin{pmatrix} f_0 \circ g & f_1 \circ g \\ f_0 \circ \overline{g} & f_1 \circ \overline{g} \end{pmatrix}}{g - \overline{g}} \cdot \frac{g - \overline{g}}{x - \overline{x}} \\ &= (\theta(F) \circ (\overline{g},\, g)) \cdot \frac{g - \overline{g}}{x - \overline{x}}. \end{aligned}$$

By Lemma 2.1, which factors the bilinear form of a product of two polynomials, we have

$$\theta(F \circ G) \;=\; \mathrm{L}_{\overline{\mathcal{Q}}}(M_p) \times \mathrm{R}_{\mathcal{P}}(M_q)\,,$$

where $p = \theta(F) \circ (\overline{g}, g)$ and $q = (g - \overline{g})/(x - \overline{x})$, moreover $\mathcal{P} = \{0, \ldots, (m-1)k\}$, $\overline{\mathcal{Q}} = \{0, \ldots, k-1\}$.

By Lemma 2.3, which factors the bilinear form of a composed polynomial,

$$M_p \;=\; \mathrm{S}_{\overline{\Delta}_F}(\overline{g})^{\mathrm{T}} \times \Theta_F \times \mathrm{S}_{\Delta_F}(g)\,.$$

By Lemma 2.2, which relates operator L applied to product of matrices, we can decompose left side as

$$\mathrm{L}_{\overline{\mathcal{Q}}}(M_p) \;=\; \mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\Delta}_F}(\overline{g})^{\mathrm{T}}\Big) \times \mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F) \times \mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(g))\,.$$

Therefore

$$\Theta_{F \circ G} = \mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\Delta}_F}(g)^{\mathrm{T}}\Big) \;\times\; \mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F) \times \big(\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(g)) \times \mathrm{R}_{\mathcal{P}}(M_q)\big)\,.$$

This factorization can also be found in Figure 3.

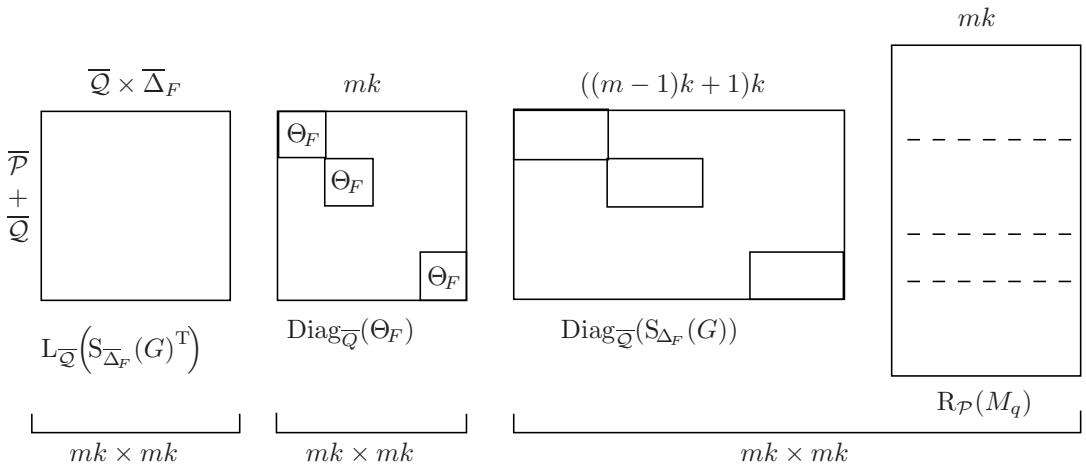

Figure 3: Decomposition of $\Theta_{F\circ G}$ in univariate setting.

Next we compute the determinant of the Bézout matrix $\Theta_{F\circ G}$ of the composed polynomials. Notice that the above factors

$$\mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F)\,, \quad \mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\Delta}_F}(G)^{\mathrm{T}}\Big) \qquad \text{and} \qquad \mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)$$

are all square matrices of size mk.

Note that $\Delta_F = \overline{\Delta}_F = \{0,\ldots,m-1\}$, that is Θ_F is square of size m and since $\overline{\mathcal{Q}} = \{0,\ldots,k-1\}$, that is $|\overline{\mathcal{Q}}| = k$, the determinant of $\mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F)$ is $\det(\Theta_F)^k$.

By proposition 2.3, the columns of matrix $\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\Delta}_F}(G)^{\mathrm{T}}\Big)$ can permuted to make the matrix step triangular. Moreover in the univariate case it is square and hence triangular with entries d_k^i for $i \in \{0,\ldots,m-1\}$. Its determinant is (see equation (5))

$$\det\Big[\mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\Delta}_F}(G)^{\mathrm{T}}\Big)\Big] \;=\; \sum_{i=0}^{m-1} d_k^{k\,i} = (d_k)^{km(m-1)/2}.$$

Next consider the matrix $\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)$. Note that, by the proposition 2.2, the matrix M_q is anti-triangular of size $k \times k$, with anti-diagonal entries d_k.

Proposition 2.4 *The matrix* $\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)$ *is* triangular *with diagonal entries* d_k^{i+1} *in row labelled by* $i.l$ *for all* $i \in \Delta_F = \{0,\ldots,m-1\}$ *and* $l \in \overline{\mathcal{Q}} = \{0,\ldots,k-1\}$. *Therefore*

$$\det\big[\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)\big] = \prod_{i=0}^{m-1}\prod_{l=0}^{k-1} d_k^{i+1} = d_k^{km(m+1)/2}.$$

Proof: Let $s = \overline{Z}_s \times \mathrm{S}_{\Delta_F}(G) \times X_s$, and set $M_s = \mathrm{S}_{\Delta_F}(G)$, where support of $\overline{Z}_s$ is Δ_F, in terms of new variable $\overline{z}$. Since

$$\mathrm{L}_{\overline{\mathcal{Q}}}(M_s) = \mathrm{Diag}_{\overline{\mathcal{Q}}}(M_s)$$

whenever Δ_F and $\overline{\mathcal{Q}}$ are supports in terms of different variables, by proposition 2.1 we have

$$\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q) \;=\; \mathrm{L}_{\overline{\mathcal{Q}}}(M_s) \times \mathrm{R}_{\mathcal{P}}(M_q) = M_{sq}.$$

The polynomial product

$$s \cdot q = \sum_{\substack{\overline{e}_s \in \Delta_F \\ \overline{e}_q \in \overline{\mathcal{Q}}}} \overline{z}^{\overline{e}_s}\, \overline{x}^{\overline{e}_q} \sum_{\beta = e_s + e_q} s_{\overline{e}_s, e_s}\, q_{\overline{e}_q, e_q}\; x^{\beta}.$$

Combining the descriptions of the coefficients of polynomials q and s, which are

$$q_{\overline{e}_q, e_q} = \begin{cases} d_k & \text{if } \overline{e}_q + e_q = k-1, \\ 0 & \text{if } \overline{e}_q + e_q > k-1, \end{cases} \qquad \text{by Proposition 2.2}$$

and also

$$s_{\overline{e}_s, e_s} = \begin{cases} d_k^{\overline{e}_s} & \text{if } e_s = k\overline{e}_s, \\ 0 & \text{if } e_s > k\overline{e}_s, \end{cases} \qquad \text{by equation (3)}$$

we get that

$$(s \cdot q)_{\overline{e}_s \overline{e}_q, \beta} = \sum_{\beta = e_s + e_q} s_{\overline{e}_s, e_s}\, q_{\overline{e}_q, e_q}\; x^{\beta} = \begin{cases} d_k^{\overline{e}_s + 1} & \text{if } e_s = k\overline{e}_s \text{ and } \overline{e}_q + e_q = k-1, \\ 0 & \text{if } e_s > k\overline{e}_s \text{ or } \overline{e}_q + e_q > k-1, \end{cases}$$

where $\beta = e_s + e_q$.

Since $\overline{e}_s \in \Delta_F = [m-1, \ldots, 0]$, $e_s \in [(m-1)k, \ldots, 0]$, $\overline{e}_q \in [k-1, \ldots, 0]$ and $e_s \in [k-1, \ldots, 0]$, we can rewrite above as

$$(s \cdot q)_{i.l,j} = \begin{cases} d_k^{i+1} & \text{if } j = ik + l, \\ 0 & \text{if } j < ik + l. \end{cases}$$

It is easy to see that M_{sq} is lower triangular matrix if rows of are indexed by $[m-1, \ldots, 0] \times [k-1, \ldots, 0]$ and columns indexed by $[km-1, \ldots, 0]$. $\square$

Hence

$$\begin{aligned} \det(\Theta_{F \circ G}) &= \det\left[\mathrm{L}_{\overline{\mathcal{Q}}}\left(\mathrm{S}_{\overline{\Delta}_F}(G)^{\mathrm{T}}\right)\right] \times \det\left[\mathrm{Diag}_k(\Theta_F)\right] \times \det\left[\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)\right] \\ &= (d_k)^{km(m-1)/2}\, (\det(\Theta_F))^k\, (d_k)^{km(m+1)/2} \\ &= (d_k)^{km^2} \det(\Theta_F)^k. \end{aligned}$$

It is well-known that in the case of $m_0 > m_1$, the determinant of the Bézout matrix constructed for F has $a_{m_0}^{(m_0 - m_1)}$ as an extraneous factor. For the composed system $F \circ G$,

$$\begin{aligned} \det(\Theta_{F \circ G}) &= d_k^{km^2}\, a_{m_0}^{(m_0 - m_1)k}\, \mathrm{Res}(f_0,\, f_1)^k \\ &= \left(d_k^{m_0} a_{m_0}\right)^{k(m_0 - m_1)} \cdot \mathrm{Res}(F \circ G). \end{aligned}$$

In this case, the extraneous factor is $(d_k^{m_0} a_{m_0})^{k(m_0 - m_1)}$, which is the extraneous factor arising from F raised to the power $k(m_0 - m_1)$ in addition to another extraneous factor which is a power of d_k, the leading coefficient of g.

Most of the above reasoning carries to general multivariate case, with a few caveats. First, Dixon matrices are not guaranteed to be square or non-singular, and as the result determinant cannot be extracted. The technique introduced in [KSY94] for extracting multiple of the resultant from matrix minor can be extended to these cases. Second, extra care is required to show that matrices (or their minors) are triangular so that determinant can be extracted. Moreover the extraneous factors arising in multivariate setting are more complex.

3 Multivariate Case

Consider a polynomial system $F = (f_0, f_1, , \ldots, f_n)$, in variables $y_1, \ldots, y_n$. Let $G = (g_1, \ldots, g_n)$ be a list *univariate* polynomials defined as

$$g_i = d_{i,k_i}\, x_i^{k_i} \,+\, d_{i,k_i-1}\, x_i^{k_i-1} \,+ \cdots +\, d_{i,0}, \qquad \text{for} \quad i = 1, \ldots, n,$$

of degrees $k_1, \ldots, k_n$, respectively, and let $\overline{G} = (\overline{g}_1, \ldots, \overline{g}_n)$, where $\overline{g}_j$ is obtained from g_j by replacing x_j with $\overline{x}_j$.

The **Cayley-Dixon Construction** of the composed polynomials $F \circ G$ is a generalization of the Cayley-Bézout construction from the univariate case. The Dixon polynomial of the composed system

$$\begin{aligned}
\theta_{F\circ G} &= \frac{\det \begin{bmatrix} f_0 \circ (\pi_0(G)) & \ldots & f_n \circ (\pi_0(G)) \\ \vdots & \ddots & \vdots \\ f_0 \circ (\pi_n(G)) & \ldots & f_n \circ (\pi_n(G)) \end{bmatrix}}{\prod_{i=1}^n (x_i - \overline{x}_i)} \\
&= \frac{\det \begin{bmatrix} f_0 \circ (\pi_0(G)) & \ldots & f_n \circ (\pi_0(G)) \\ \vdots & \ddots & \vdots \\ f_0 \circ (\pi_n(G)) & \ldots & f_n \circ (\pi_n(G)) \end{bmatrix}}{\prod_{i=1}^n (g_i - \overline{g}_i)} \times \frac{\prod_{i=1}^n (g_i - \overline{g}_i)}{\prod_{i=1}^n (x_i - \overline{x}_i)} \\
&= \theta_F \circ (\overline{G}, G) \,\times\, \prod_{i=1}^n \frac{g_i - \overline{g}_i}{x_i - \overline{x}_i}.
\end{aligned}$$

Let

$$p \,=\, \theta_F \circ (\overline{G}, G), \qquad \text{and} \qquad q \,=\, \prod_{i=1}^n \frac{g_i - \overline{g}_i}{x_i - \overline{x}_i},$$

where $\mathcal{P}$ is the support of p with respect to the variables $x_1, \ldots, x_n$ and $\overline{\mathcal{Q}}$ is the support of q with respect to the variables $\overline{x}_1, \ldots, \overline{x}_n$.

Then with application of lemmas 2.1, 2.2 and 2.3 as in univariate case, we obtain the the factorization

$$\Theta_{F\circ G} \,=\, \mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\Delta}_F}(G)^{\mathrm{T}}\Big) \times \mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F) \times \big(\mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(M_q)\big)\,.$$

This proves following main theorem.

Theorem 3.1 *Let $F = (f_0, f_1, \ldots, f_n)$ and $G = (g_1, \ldots, g_n)$ be lists of generic polynomials. Then the Dixon matrix $\Theta_{F\circ G}$ is*

$$A_{\mathrm{L}} \times \mathrm{Diag}_{k_1 \cdots k_n}(\Theta_F) \times A_{\mathrm{R}},$$

and moreover matrices A_{L} and A_{R} are step-triangular matrices (up to row/column permutation), where

$$A_{\mathrm{L}} = \mathrm{L}_{\overline{\mathcal{Q}}}\Big(\mathrm{S}_{\overline{\Delta}_F}(\overline{G})^{\mathrm{T}}\Big)\,, \qquad \textit{and} \qquad A_{\mathrm{R}} = \mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(\mathrm{M}_q)\,,$$

where $q = \prod_{i=1}^n \frac{g_i - \overline{g}_i}{x_i - \overline{x}_i}$, and $\mathcal{Q}$, $\overline{\mathcal{Q}}$ are, respectively, the supports of q in the variables $\mathbf{x}$ and $\overline{\mathbf{x}}$.

In particular, for the generic n-degree polynomial system F and a generic system G of n polynomials used to substitute for variables $y_1, \cdots, y_n$ in F, the factors above can be proved to be square and non-singular matrices [Sax97]. We investigate this in the next section.

More generally, if the factors are square in the above theorem, then we can derive precise expression for the determinant of the Dixon matrix. The next lemma, uses the same notation as Theorem 3.1, and derives the determinants of the factors of the Dixon matrix of the composed system.

Lemma 3.1 *If* $|\overline{\Delta}_F| \cdot \prod_{j=1}^{n} k_j \;=\; |\overline{\Delta}_{F\circ G}|$, *i.e.,* A_{L} *is square, then*

$$\det(A_{\mathrm{L}}) \;=\; \pm\,(d_{1,\,k_1}, \ldots, d_{n,\,k_n})^{(\sum_{\alpha\in\overline{\Delta}_F}\alpha)\,k_1\cdots k_n}\,;$$

if $|\Delta_F| \;=\; |\overline{\Delta}_F|$, *i.e.,* Θ_F *is square, then*

$$\det\Big(\mathrm{Diag}_{|\overline{\mathcal{Q}}|}(\Theta_F)\Big) \;=\; (\det(\Theta_F))^{k_1\cdots k_n}\,;$$

and if $|\Delta_F| \cdot \prod_{j=1}^{n} k_j \;=\; |\Delta_{F\circ G}|$, *i.e.,* A_{R} *is square, then*

$$\det(A_{\mathrm{R}}) \;=\; \pm\,(d_{1,\,k_1}, \ldots, d_{n,k_n})^{(|\Delta_F| + \sum_{\beta\in\Delta_F}\beta)k_1\cdots k_n}\,.$$

The above results hold when the generic coefficients of f_i's and g_j's are specialized as long as the sizes of the matrices and its ranks do not degenerate.

Proof: When A_L is square it is triangular (up to column permutation), with diagonal entries

$$(A_L)_{\alpha,\overline{e}_q.e_p} = (d_{1,k_1}, \ldots, d_{n,k_n})^{e_p}$$

in column $\overline{e}_q e_p$, where $e_p \in \mathcal{P} = \overline{\Delta}_F$, by proposition 2.3. Note that the size of $\overline{Q}$ is $k_1 \cdots k_n$. Therefore

$$\det(A_L) = \prod_{\substack{e_p\in\overline{\Delta}_F\\ \overline{e}_q\in\overline{\mathcal{Q}}}} (d_{1,k_1}, \ldots, d_{n,k_n})^{e_p} = (d_{1,\,k_1}, \ldots, d_{n,k_n})^{(\sum_{\alpha\in\overline{\Delta}_F}\alpha)\,k_1\cdots k_n}\,.$$

Also, for $A_R = \mathrm{Diag}_{\overline{\mathcal{Q}}}(\mathrm{S}_{\Delta_F}(G)) \times \mathrm{R}_{\mathcal{P}}(\mathrm{M}_q)$, let $s = \overline{Z}_s \times \mathrm{S}_{\Delta_F}(G) \times X_s$, $A_R = M_{sq}$, as in the univariate case. By proposition 2.2, M_q is triangular, where

$$q_{\overline{e}_q,e_q} = \begin{cases} d_{1,k_1}\cdots d_{n,k_n} & \text{if } \forall i \;\text{ s.t. } (\overline{e}_q)_i + (e_q)_i = k_i - 1,\\ 0 & \text{if } \exists i \;\text{ s.t. } (\overline{e}_q)_i + (e_q)_i > k_i - 1 \end{cases}$$

and entries of $\mathrm{S}_{\Delta_F}(G)$ by equation 2 are

$$s_{\overline{e}_s,e_s} = \begin{cases} (d_{1,k_1}, \ldots, d_{n,k_n})^{\overline{e}_s} & \text{if } \forall\, i \text{ s.t. } (e_s)_i = k_i(\overline{e}_s)_i,\\ 0 & \text{if } \exists\, i \text{ s.t. } k_i(\overline{e}_s)_i < (e_s)_i, \end{cases}$$

for $\overline{e}_s \in \Delta_F$ and e_s in support of G^α for all $\alpha \in \Delta_F$. Therefore

$$(s\cdot q)_{\overline{e}_s\overline{e}_q,e_s+e_q} \;=\; \begin{cases} (d_{1,k_1}, \ldots, d_{n,k_n})^{\overline{e}_s+1} & \text{if } \forall\, i \text{ s.t. } (e_s)_i = k_i(\overline{e}_s)_i\\ & \text{and } \overline{e}_q + e_q = k-1,\\ 0 & \text{if } \exists\, i \text{ s.t. } (e_s)_i > k_i(\overline{e}_s)_i\\ & \text{or } (\overline{e}_q)_i + (e_q)_i > k_i - 1, \end{cases}$$

i.e., A_{R} is triangular, since first non-zero entry in each row is in different column.

In row $\overline{e}_s\overline{e}_q$ the diagonal element is $(d_{1,k_1},\ldots,d_{n,k_n})^{\overline{e}_s+1}$. Since $\overline{e}_s \in \Delta_F$ and $\overline{e}_q \in \overline{\mathcal{Q}}$, where $|\overline{\mathcal{Q}}| = k_1\cdots k_n$, we have the determinant of A_{R}

$$\det(A_R) = \prod_{\substack{\overline{e}_s\in\Delta_F\\ \overline{e}_q\in\overline{\mathcal{Q}}}} (d_{1,\,k_1},\ldots,d_{n,k_n})^{\overline{e}_s+1} = (d_{1,\,k_1},\ldots,d_{n,k_n})^{(|\Delta_F|+\sum_{\beta\in\Delta_F}\beta)\,k_1\ldots k_n}.$$

□

By Theorem 3.1 and the above proposition, we have another main results of the paper.

Theorem 3.2 *Let F be a polynomial system for which the Cayley-Dixon resultant formulation leads to a square and nonsingular resultant matrix Θ_F whose determinant is* $\mathrm{Res}(F)$. *Then under multi-univariate polynomial composition $F\circ G$,*

$$\mathrm{Res}(F\circ G) = (d_{1,\,k_1},\ldots,d_{n,\,k_n})^{(\sum_{\alpha\in\overline{\Delta}_F}\alpha+|\Delta_F|+\sum_{\beta\in\Delta_F}\beta)\,k_1\cdots k_n}\,\mathrm{Res}(F)^{k_1\cdots k_n}.$$

3.1 Rank Submatrix Construction

This subsection examines the cases when the Dixon matrix of the composed polynomials (or any of its factors in Lemma 3.1) is not square or when the Dixon matrix is rank deficient. In these cases one can extract a projection operator from the Dixon matrix by computing the determinant of any maximal minor [KSY94, BEM00]. Since the Dixon matrix $\Theta_{F\circ G}$ can be factored into a product one obtains a similar factorization of a maximal minor,

$$\begin{aligned}\det_{\max}&\left[A_{\mathrm{L}}\times\mathrm{Diag}_{k_1\cdots k_n}(\Theta_F)\times A_{\mathrm{R}}\right]\\ &= \det\left[\max_{\mathrm{row}}(A_{\mathrm{L}})\times\mathrm{Diag}_{k_1\cdots k_n}(\Theta_F)\times\max_{\mathrm{col}}(A_{\mathrm{R}})\right]\\ &= \det\left[M_{\mathrm{L}}\times\mathrm{Diag}_{k_1\cdots k_n}(\Theta_F)\times M_{\mathrm{R}}\right],\end{aligned}$$

by selecting appropriate rows M_{L} of A_{L} and columns M_{R} of A_{R}. Furthermore, the well-known Cauchy-Binet formula allows us to expand the determinant of the minor into a sum of products of the form $l\cdot s\cdot r$, where l ranges over determinants of minors of M_{L}, s ranges over determinants of minors of $\mathrm{Diag}_{k_1\cdots k_n}(\Theta_F)$ and r ranges over determinants of minors of M_{R}.

More formally, given a square matrix M of size $m\times m$, where $M = T_1\times D\times T_2$, and D is of size $s\times t$ for $m>s$ or $m>t$ then $\det(M)=0$; otherwise, when $m=s=t$ then $\det(M)=\det(T_1)\cdot\det(D)\cdot\det(T_2)$. If $m<s$ or $m<t$, by application of the Cauchy-Binet expansion of the determinant of the product of non-square matrices we get

$$\det(T_1\ \times\ D\ \times\ T_2)\quad=\sum_{\substack{\sigma\in\mathtt{C}^s_m,\\ \rho\in\mathtt{C}^t_m}}\det(\mathrm{cols}_\sigma(T_1))\ \cdot\ \det(\mathrm{submatrix}_{\rho,\sigma}(D))\ \cdot\ \det(\mathrm{cols}_\rho(T_2)),$$

where $\mathtt{C}^s_m$ is set of subsets of size m from set of $\{1,\ldots,s\}$.

The following elementary linear algebra result, guarantees that the gcd of the determinants of all maximal minors of matrix D will be a factor in any maximal minor determinant of M.

Proposition 3.1 *If $M = T_1 \times D \times T_2$ and the rank of T_1 equals the number of columns of T_1 and the rank of T_2 equals the number of rows of T_2, then* $\mathrm{rank}(M) = \mathrm{rank}(D)$.

That is, if $\mathrm{rank}(D) = m$ and matrices T_1 and T_2 are of rank s and t respectively, then $\gcd(\det_{\max}(D))$ is a factor in $\det(T_1 \times D \times T_2)$.

Using above we can compute the determinant of the maximal minor of Dixon matrix of composed system by considering maximal minors of A_L, Θ_F and A_R. Since A_L and A_R are step triangular they are of "full" rank, that is the rank is equal to the minimum number of rows and columns. This allows to obtain similar formula to non-square case.

Theorem 3.3 *For a polynomial system $F = (f_0, f_1, \ldots, f_n)$, composed with univariate polynomials $G = (g_1, \ldots, g_n)$,*

$$\det_{\max}(\Theta_{F\circ G}) = d_{1,k_1}^{\epsilon_1} \cdots d_{n,k_n}^{\epsilon_n} E \left(\gcd \det_{\max}(\Theta_F)\right)^{k_1 \cdots k_n},$$

where E is extraneous factor dependent not only on the coefficients of G but also F.

The above theorem establishes that whenever the resultant can be computed by Cayley-Dixon construction, the resultant is also decomposable as shown above.

It is an open question, what the values of $\epsilon_1, \ldots, \epsilon_n$ are in general and whether the factor E for resultant is constant in all cases.

To illustrate this further consider an example of a composed polynomial system for which the Dixon matrix is non-square and for which the determinants of the maximal minors has the structure described in the previous subsection.

Example 3.1 [Maximal minor construction] Let

$$\begin{array}{rcl} f_0 & = & y_1\,y_2 + a\,y_1 + b\,y_2 + a\,b, \\ f_1 & = & y_1\,y_2 + a\,y_1 + b\,y_2 + c, \\ f_2 & = & y_1\,y_2 + y_1 + b\,y_2 + a, \end{array} \quad \text{and} \quad \begin{array}{rcl} g_1 & = & d_{1,2}\,x_1^2 + d_{1,1}\,x_1 + d_{1,0}, \\ g_2 & = & d_{2,1}\,x_2 + d_{2,0}. \end{array}$$

For the composed polynomials $F \circ G$, the Dixon matrix $\Theta_{F\circ G}$ is a 4×2 matrix with rank 2. Then the determinant of any maximal minor of 4×2 matrix $\Theta_{F\circ G}$ is

$$-d_{21}^2 d_{12}^2 (a-1)^2 (ab-c)^2 (d_{12}b + d_{10}d_{12} - d_{11}^2),$$

By Theorem 3.1, $\Theta_{F\circ G} = A_{\mathrm{L}} \times \mathrm{Diag}_2(\Theta_F) \times A_{\mathrm{R}}$, where

$$A_{\mathrm{L}} = \begin{bmatrix} d_{1,2} & 0 & 0 & 0 \\ d_{1,1} & 0 & d_{1,2} & 0 \\ d_{1,0} & 1 & d_{1,1} & 0 \\ 0 & 0 & d_{1,0} & 1 \end{bmatrix}, \quad A_{\mathrm{R}} = \begin{bmatrix} 0 & d_{1,2}\,d_{2,1} \\ d_{1,2}\,d_{2,1} & d_{1,1}\,d_{2,1} \end{bmatrix},$$

$$\Theta_F = \begin{bmatrix} -a\,c + a^2\,b + c - a\,b \\ -a\,b\,c + b\,c - a\,b^2 + a^2\,b^2 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

The 2×2 minor of $\Theta_{F\circ G}$ consisting for instance of the second and the third rows and the first two columns (obtained by selecting the second and third rows of A_{L}) factorizes

$$M_L \times \mathrm{Diag}_2(\Theta_F) \times M_R = \begin{bmatrix} d_{1,1} & 0 & d_{1,2} & 0 \\ d_{1,0} & 1 & d_{1,1} & 0 \end{bmatrix} \times \begin{bmatrix} w_1 & 0 \\ w_2 & 0 \\ 0 & w_1 \\ 0 & w_2 \end{bmatrix} \times \begin{bmatrix} 0 & d_{1,2}\,d_{2,1} \\ d_{1,2}\,d_{2,1} & d_{1,1}\,d_{2,1} \end{bmatrix}.$$

By the Cauchy-Binet formula, the determinant of the above product is

$$\begin{aligned}
&\det(M_L \times \mathrm{Diag}_2(\Theta_F) \times M_R) \\
&\qquad = \det \begin{bmatrix} 0 & d_{1,2} \\ 1 & d_{1,1} \end{bmatrix} \cdot \det \begin{bmatrix} w_2 & 0 \\ 0 & w_1 \end{bmatrix} \cdot \det \begin{bmatrix} 0 & d_{1,2}\, d_{2,1} \\ d_{1,2}\, d_{2,1} & d_{1,1}\, d_{2,1} \end{bmatrix} \\
&\qquad + \det \begin{bmatrix} d_{1,1} & d_{1,2} \\ d_{1,0} & d_{1,1} \end{bmatrix} \cdot \det \begin{bmatrix} w_1 & 0 \\ 0 & w_1 \end{bmatrix} \cdot \det \begin{bmatrix} 0 & d_{1,2}\, d_{2,1} \\ d_{1,2}\, d_{2,1} & d_{1,1}\, d_{2,1} \end{bmatrix} \\
&\qquad\qquad = \gcd(\det_{\max}(\mathrm{Diag}_2(\Theta_F))) \cdot E_g = \gcd(\det_{\max}(\Theta_F))^2 \cdot E_g,
\end{aligned}$$

where $E_g = (d_{12}b + d_{10}d_{12} - d_{11}^2)(-d_{21}^2 d_{12}^2)$. Note that the determinants of two maximal minors of Θ_F are

$$(a-1)(ab-c) \quad \text{and} \quad b(a-1)(ab-c)$$

Note that the greatest common divisor of the determinants of the maximal minors of Θ_F is $(a-1)\,(ab-c)$, and the determinant of the maximal minor of $\Theta_{F\circ G}$ is

$$((a-1)\,(ab-c))^2\,(-d_{1,2}^2\, d_{2,1}^2)\,(-d_{1,2}\, b \,+\, d_{1,1}^2 - d_{1,2}d_{1,0}),$$

exhibiting that the greatest common divisor of the determinants of the maximal minors of Θ_F raised by 2 $(= k_1\, k_2)$ is a factor.

In general, the factor E_g will contain an extra factor for each maximal minor selected from Θ_F, in this case it is 1 and b.

In this example, the matrix A_{R} is square whereas in general that need not be the case. Even if A_{R} is not square, the Cauchy-Binet formula can be used to obtain a similar sum. □

3.2 Resultant of Composed n-degree polynomial system

In this section, we derive the McKay and Wang formula (6) (shown on page 14 in univariate setting) for n-degree polynomials systems.

Consider the $(m_1, \ldots, m_n)$-degree generic polynomials $f_0, f_1, \ldots, f_n$ where

$$f_j = \sum_{i_1=1}^{m_1} \cdots \sum_{i_n=1}^{m_n} c_{j,i_1,\ldots,i_n} y_1^{i_1} \cdots y_n^{i_n} \qquad \text{for} \qquad j = 0, 1, \ldots, n,$$

with generic coefficients $c_{j,i_1,\ldots,i_n}$ and variables $y_1, \ldots, y_n$.

It is easy to see that the composed polynomials $f_i \circ (g_1, \ldots, g_n)$, $0, \leq i \leq n$, are $(m_1k_1, \ldots, m_nk_n)$-degree as well.

The support of the Dixon polynomial for the n-degree polynomial system F is

$$\begin{aligned}
\overline{\Delta}_F &= \{\, \alpha \in \mathbb{N}^n \mid \alpha_i < (n-i+1)m_i \quad \text{for } i = 1, \ldots, n \,\}, \\
\Delta_F &= \{\, \alpha \in \mathbb{N}^n \mid \alpha_i < i m_i \quad \text{for } i = 1, \ldots, n \,\},
\end{aligned}$$

and therefore $|\overline{\Delta}_F| = |\Delta_F| = n!\, m_1 \cdots m_n$.

To apply Lemma 3.1, in the above support, sum of all points in the support, for a particular coordinate $i \in \{1, \ldots, n\}$ is

$$\begin{aligned}
\sum_{\alpha \in \overline{\Delta}_F} \alpha_i &= n m_1 (n-1) m_2 \cdots (n-i+2) m_{i-1} \left(\sum_{j=0}^{(n-i+1)m_i - 1} j \right) (n-i) m_{i+1} \cdots m_n \\
&= n!\, m_1 \cdots m_n \frac{(n-i+1)m_i - 1}{2}, \\
\sum_{\alpha \in \Delta_F} \alpha_i &= m_1\, 2 m_2 \cdots (i-1) m_{i-1} \left(\sum_{j=0}^{i m_i - 1} j \right) (i+1) m_{i+1} \cdots n m_n \\
&= n!\, m_1 \cdots m_n \frac{i m_i - 1}{2}.
\end{aligned}$$

Substituting into Lemma 3.1,

$$\begin{aligned}
\det \left[\mathrm{Diag}_{\overline{\mathcal{Q}}}(\Theta_F) \right] &= (\det(\Theta_F))^{k_1 \cdots k_n}, \\
\det [A_L] &= \prod_{i=1}^{n} d_{i,\, k_i}^{n!\, m_1 \cdots m_n \frac{(n-i+1)m_i - 1}{2}\, k_1 \cdots k_n}, \\
\det [A_R] &= \prod_{j=1}^{n} d_{j,\, n_j}^{(n!\, m_1 \cdots m_n + n!\, m_1 \cdots m_n \frac{i m_i - 1}{2})\, k_1 \cdots k_n}.
\end{aligned}$$

Note that, if F and G are generic, then the coefficients of $F \circ G$ will still not have any algebraic relations, and therefore the system $F \circ G$ is generic. By Theorem 3.2 and the fact that the Dixon matrix is exact for generic n-degree systems, we have another main result of the paper.

Theorem 3.4 *For the unmixed n-degree case,*

$$\mathrm{Res}(F \circ G) = \left(d_{1,k_1}^{m_1} \cdots d_{n,k_n}^{m_n} \right)^{\frac{(n+1)!}{2}\, m_1 \cdots m_2\, k_1 \cdots k_n} \mathrm{Res}(F)^{k_1 \cdots k_n}.$$

4 Exact Mixed Systems

As the following example illustrates, resultants of composed systems can be computed exactly under certain conditions using the Cayley-Dixon construction if the outermost system in a composed system is such that its resultant can be computed exactly using the Cayley-Dixon construction. In [CK03], we have identified a class of generic unmixed polynomial systems for which the resultant can be computed without extraneous factor using the Cayley-Dixon method. By composing such an unmixed system F with G, it is possible to compute resultants without extraneous factors of generic as well as specialized mixed systems as well. This opens up a promising area of research as very little is known about the subclasss of mixed or nongeneric systems for which resultants can be computed exactly.

Consider the following unmixed bivariate polynomial system:

$$\begin{aligned}
f_0 &= a_1 a_2 + a_1 y_1 + a_2 y_2 + y_1 y_2, \\
f_1 &= b_0 + b_1 y_1 + b_2 y_2 + b_3 y_1 y_2, \\
f_2 &= c_0 + c_1 y_1 + c_2 y_2 + c_3 y_1 y_2.
\end{aligned}$$

This system F has unmixed bidegree support; the determinant of the Dixon matrix of F is exactly its resultant,

$$\begin{aligned}\mathrm{Res}(f_0, f_1, f_2) = (&a_2^2 b_1 - a_2 b_0 - a_2^2 b_3 + c_2 b_0 + a_2 b_2 - c_0 b_2 + c_0 b_3 a_2 - c_2 b_1 a_2) \\ &(a_1 b_0 - b_0 + c_0 b_1 - c_2 a_1 b_1 - c_0 b_3 a_1 + b_3 c_2 a_1^2 + b_2 a_1 - b_2 a_1^2).\end{aligned}$$

Consider now a substitution $G = (g_1 = x_1^2 + r x_1 - a_2, g_2 = x_2 - a_1)$, where r is arbitrary symbolic parameter. The composed system is neither generic nor unmixed.

$$\begin{aligned}
f_0 &= x_1^2 x_2 + r x_1 x_2, \\
f_1 &= b_0 - b_2 a_1 - b_1 a_2 + b_3 a_1 a_2 + (b_1 - b_3 a_1) x_1^2 + (-b_3 a_2 + b_2) x_2 \\
&\quad + (b_1 r - b_3 a_1 r) x_1 + b_3 x_1^2 x_2 + b_3 r x_1 x_2, \\
f_2 &= c_0 - c_2 a_1 - c_1 a_2 + c_3 a_1 a_2 + (c_1 - c_3 a_1) x_1^2 + (-c_3 a_2 + c_2) x_2 \\
&\quad + (c_1 r - c_3 a_1 r) x_1 + c_3 x_1^2 x_2 + c_3 r x_1 x_2.
\end{aligned}$$

The determinant of the Dixon matrix of the above composed system is: $\mathrm{Res}(f_0, f_1, f_2)^2$, we get the exact resultant $F \circ G$ (modulo sign). However, the determinant of the Dixon matrix corresponding to the generic system whose supports is the same as that of $F \circ G$, i.e. the system

$$\begin{aligned}
f_0 &= a_1 x_1^2 x_2 + a_2 x_1 x_2, \\
f_1 &= b_1 + b_2 x_1^2 + b_3 x_2 + b_4 x_1 + b_5 x_1^2 x_2 + b_6 x_1 x_2, \\
f_2 &= c_1 + c_2 x_1^2 + c_3 x_2 + c_4 x_1 + c_5 x_1^2 x_2 + c_6 x_1 x_2,
\end{aligned}$$

has an extraneous factor

$$(b_4^2 c_1 c_2 + c_2^2 b_1^2 + c_1^2 b_2^2 + b_2 c_4^2 b_1 - b_1 b_4 c_2 c_4 - 2 c_1 b_2 b_1 c_2 - c_1 b_2 b_4 c_4)(c_1 b_3 - c_3 b_1)$$

along with the resultant.

Hence there exist a condition on the coefficients of the polynomial system to have exact Dixon matrices. So far researchers have investigated only conditions on the support of polynomial system to have Dixon-exact matrices.

This raises a question of whether for any non-exact system there is a transformation into exact (possibly non-generic) system.

5 Conclusion

This paper studied the Cayley-Dixon construction of resultants for multi-univariate composed polynomials. It gave a factorization of the Cayley-Dixon matrix induced by the structure of the composed polynomials and it showed how to efficiently extract the Dixon projection operator utilizing the factorization of the Cayley-Dixon matrix.

In a special case, when $g_i = x_i^k$, the composition problem in the context of Cayley-Dixon construction was analyzed in [KS97], where it was studied as support scaling. Under this setting the main result of that paper coincides with Theorem 3.2. Results presented here are thus strict generalizations.

A new resultant formula like in [MW89] has been derived for multi-univariate composition of n-degree systems.

This paper also highlighted a possible class of mixed or non-generic polynomial systems for which the resultant can be computed exactly because of given composition structures in the polynomial systems. This opens up a promising area of research because very little is known of the class of such polynomial systems.

References

[BEM00] Laurent Buse, Mohamed Elkadi, and Bernar Mourrain. Generalized resultants over unirational algebraic varieties. *J. Symbolic Computation*, 29:515–526, 2000.

[BGW88] C. Bajaj, T. Garrity, and J. Warren. On the application of multi-equational resultants. Technical Report CSD-TR-826, Dept. of Computer Science, Purdue University, Nov 1988.

[Chi90] E. Chionh. *Base points, resultants, and the implicit representation of rational Surfaces.* PhD dissertation, University of Waterloo, Department of Computer Science, 1990.

[Cht03] A. D. Chtcherba. *A new Sylvester-type Resultant Method based on the Dixon-Bézout Formulation.* PhD dissertation, University of New Mexico, Department of Computer Science, Aug 2003.

[CK03] A.D. Chtcherba and D. Kapur. Exact resultants for corner-cut unmixed multivariate polynomial systems using the Dixon formulation. *Journal of Symbolic Computation*, 36(3–4):289–315, 2003.

[CMW95] C. C. Cheng, J. H. McKay, and S. S. Wang. A chain rule for multivariable resultants. *Proceedings of the American Mathematical Society*, 123(4):1037–1047, April 1995.

[Dix08] A.L. Dixon. The eliminant of three quantics in two independent variables. *Proc. London Mathematical Society*, 6:468–478, 1908.

[FC04] Mao-Ching Foo and Eng-Wee Chionh. Corner edge cutting and dixon a-resultant quotients. *Journal of Symbolic Computation*, 37:101–119, 2004.

[HM02] H. Hong and M. Minimair. Sparse resultant of composed polynomials I. *J. Symbolic Computation*, 33:447–465, 2002.

[Hof89] C.M Hoffman. *Geometric and Solid modeling.* Morgan Kaufmann Publishers, Inc., San Mateo, California 94403, 1989.

[Jou91] J. P. Jouanolou. Le formalisme du résultant. *Adv. Math.*, 90(2):117–263, 1991.

[KL92] D. Kapur and Y. Lakshman. *Symbolic and Numeric Computation for Artificial Inteligence*, chapter Elimination Methods: an Introduction. Academic Press, 1992. Donald, Kapur and Mundy (eds.).

[KS97] D. Kapur and T. Saxena. Extraneous factors in the Dixon resultant formulation. In *ISSAC*, pages 141–147, Maui, Hawaii, USA, 1997.

[KSY94] D. Kapur, T. Saxena, and L. Yang. Algebraic and geometric reasoning using the Dixon resultants. In *ACM ISSAC 94*, pages 99–107, Oxford, England, July 1994.

[Min01] M. Minimair. *Sparse Resultants of Composed Polynomials*. PhD thesis, North Carolina State University, Raleigh, NC, USA, 2001.

[Min02] M. Minimair. Sparse resultant of composed polynomials II. *J. Symbolic Computation*, 33:467–478, 2002.

[Min03a] M. Minimair. Dense resultant of composed polynomials. *J. Symbolic Computation*, 36(6):825–834, December 2003.

[Min03b] M. Minimair. Factoring resultants of linearly combined polynomials. In J. R. Sendra, editor, *Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation*, pages 207–214, New York, NY, 2003. ACM. ISSAC 2003, Philadelphia, PA, USA, August 3-6, 2003.

[Min03c] M. Minimair. Sparse resultant under vanishing coefficients. *J. Algebraic Combinatorics*, 18(1):53–73, July 2003.

[Min04] M. Minimair. Computing resultants of partially composed polynomials. In V. G. Ganzha, E. W. Mayr, and E. V. Vorozhtsov, editors, *Computer Algebra in Scientific Computing. Proceedings of the CASC 2004 (St. Petersburg, Russia)*, pages 359–366. TUM München, 2004.

[Mor87] A.P. Morgan. *Solving polynomial systems using continuation for Scientific and Engineering problems*. Prentice-Hall, Englewood Cliffs, NJ, 1987.

[MW89] J. H. McKay and S. S. Wang. A chain rule for the resultant of two polynomials. *Arch. Math.*, 53(4):347–351, 1989.

[PK92] J. Ponce and D.J. Kriegman. *Symbolic and Numerical Computation for Artificial Intelligence*, chapter Elimination Theory and Computer Vision: Recognition and Positioning of Curved 3D Objects from Range. Academic Press, 1992. Donald, Kapur and Mundy (eds.).

[Rub00] R. Rubio. *Unirational Fields. Theorems, Algorithms and Applications*. PhD thesis, University of Cantabria, Santander, Spain, 2000.

[Sax97] T. Saxena. *Efficient variable elimination using resultants*. PhD thesis, Department of Computer Science, State Univeristy of New York, Albany, NY, 1997.

[SG86] T.W. Sederberg and R. Goldman. Algebraic geometry for computer-aided design. *IEEE Computer Graphics and Applications*, 6(6):52–59, 1986.

[ZG00] Ming Zhang and Ron Goldman. Rectangular corner cutting and sylvester $\mathcal{A}$-resultants. In Carlo Traverso, editor, *Proc. of the ISSAC*, pages 301–308, St. Andrews, Scotland, August 2000. ACM Press.

[Zha00] M. Zhang. *Topics in Resultants and Implicitization*. PhD thesis, Rice University, Dept. of Computer Science, May 2000.