

# Resultants of Partially Composed Polynomials

Manfred Minimair<sup>1</sup>

*Department of Mathematics and Computer Science  
Seton Hall University, 400 South Orange Avenue  
South Orange, NJ 07079, USA*

---

## Abstract

We study the structure of resultants of two homogeneous *partially composed* polynomials. By two homogeneous partially composed polynomials we mean a pair of polynomials of which one does not have any given composition structure and the other one is obtained by composing a bivariate homogeneous polynomial with two bivariate homogeneous polynomials. The main contributions are two equivalent formulas, each representing the resultant of two partially composed polynomials as a certain iterated resultant of the component polynomials. Furthermore, in many cases, this iterated resultant can be computed with dramatically increased efficiency, as demonstrated by experiments.

*Key words:* resultant, composed polynomial, utilizing structure for efficient computation

---

---

*Email address:* [manfred@minimair.org](mailto:manfred@minimair.org) (Manfred Minimair).

*URL:* <http://minimair.org> (Manfred Minimair).

<sup>1</sup> Partially supported by the NSF grant CCF 0430741: RUI: Resultant Techniques for Composed Polynomials

## 1 Introduction

Resultants are fundamental in solving systems of polynomial equations and therefore have been extensively studied ([11], [23], [6], [4], [7], [14], [19], [24], [8], [20], [33], [22], [10], [2]). Recent research is focused on utilizing structure of polynomials, naturally occurring in real life problems, for example, sparsity ([35], [13], [12], [9], [5], [36], [3], [31]) as well as composition ([25], [19], [26], [8], [21], [18], [29], [28], [30]). This paper is part of the author's work on utilizing composition structures. (See [18] and [27] for more detailed motivations for composition structures.)

The focus of the current paper is entirely different from the one of the previous papers ([18], [29], [28], [30]) by the author. The previous papers considered “fully” composed polynomials, That is, multivariate composed polynomials such as  $h_1 = f_1 \circ (g_1, g_2, g_3)$ ,  $h_2 = f_2 \circ (g_1, g_2, g_3)$  and  $h_3 = f_3 \circ (g_1, g_2, g_3)$ , where each composed polynomial  $h_i$  is obtained from the polynomial  $f_i$  in the variables  $y_1, y_2, y_3$  by replacing  $y_j$  with the bivariate polynomial  $g_j$ . Note that each composed polynomial has the same inner components  $g_1, g_2, g_3$ . In contrast, the current paper considers “partially” composed polynomials where only one polynomial has a composition structure. In detail, by two partially composed polynomials  $h_1$  and  $h_2$ , we mean a bivariate homogeneous polynomial  $h_1$  that *does not have any composition structure* and a bivariate homogeneous *composed polynomial*  $h_2 = f_2 \circ (g_1, g_2)$  that is obtained from the homogeneous bivariate polynomial  $f_2$  in the variables  $y_1$  and  $y_2$  by replacing  $y_j$  with the bivariate homogeneous polynomial  $g_j$ . (Of course,  $g_1$  and  $g_2$  are required to have the same total degrees to ensure that  $h_2$  is homogeneous.) A preliminary version of this article has appeared in the CASC 2004 proceedings.

The findings of the current paper are also quite different from previous findings ([18], [29], [28], [30]). The previous papers have determined the irreducible factors of projective (Macaulay) or toric (sparse) resultants of *fully composed* polynomials. In contrast, the current paper finds that the projective (dense, Sylvester/Macaulay) resultant of two *partially composed* polynomials  $h_1$  and  $h_2$  is a certain iterated resultant. More precisely, it is the resultant of the polynomials  $f_1$  and  $f_2$ , where  $f_1$  is the resultant of certain polynomials derived from the component polynomials  $h_1$ ,  $g_1$  and  $g_2$ . Interestingly, we find *two different* natural formulas for  $f_1$ , one involving a projective (dense, Sylvester/Macaulay) resultant and another one involving a toric (sparse) resultant. Moreover, we show in experiments that for many cases this iterated resultant can be computed, over the integers modulo a prime, with dramatically improved efficiency.

The motivation for considering partially composed structures originates in the observation that composed structures (nested polynomial functions) are quite irregular in practice. The study of partially composed polynomials of the cur-

rent paper is an approach towards utilizing arbitrary composition structures for efficient computation, in contrast to regular ones studied in previous works ([18], [29], [28], [30])).

This work can also be considered as a completion of works ([25] and [26]) by McKay and Wang. In [25] they study resultants of two inhomogeneous composed polynomials as well as two inhomogeneous *partially composed* polynomials (in Theorem 7 of [25]). Additionally, in [26] they study the homogeneous generalization for the case of two composed polynomials. However, they *ignore* the case of two *homogeneous partially composed* polynomials. Furthermore, they do not address efficient computation of partially composed polynomials. In fact, their presentation of their result (Theorem 7 of [25]) does not allow an immediate computational application. Also note that Jouanolou's work [19] that considers resultants of composed polynomials in Section 5.12 ignores the partially composed case as well.

Note that the main theorem of the present paper (Theorem 1) can be considered a generalization (to the homogeneous case) of Theorem 7 of the work [25] by McKay and Wang. When applying Theorem 7 to dehomogenized special partially composed polynomials we get a result equivalent to the application of the main theorem of the present paper. Let us elaborate. First, we precisely state Theorem 7 of [25] in (1). (For the sake of a more uniform presentation, with respect to the current work and to the previous works [18], [29], [28], [30] of the current author, we use different symbols for the polynomials than in [25].) Let  $F_2$  be a univariate polynomial and  $G$  and  $H_1$  be univariate polynomials. Then, the projective (dense, Sylvester) resultant of  $H_1$  and  $H_2 = F_2 \circ G$  is the resultant of the polynomials  $F_1$  and  $F_2$ . Moreover the polynomial  $F_1$ , univariate in the variable  $y$ , is given by the following formula involving the roots of  $H_1$ . That is,

$$F_1 = H_1(0)^d \prod_{\alpha} (y - G(\alpha)), \quad (1)$$

where  $d$  is the degree of the polynomial  $G$ , which be univariate in the variable  $x$ , and  $\alpha$  ranges over the roots of  $H_1$ . (In (1) the value  $G(\alpha)$  denotes the evaluation of  $G$  in  $x = \alpha$ .) Now, note that the polynomials  $F_2$ ,  $G$  and  $H_1$  can indeed be considered as a sub-case of the homogeneous polynomials subject of the current paper. That is, they can be represented by appropriate homogeneous bivariate polynomials  $f_2, g_1, g_2$  and  $h_1$  where  $F_2 = f_2(g_1(x, 1), y)$ ,  $g_1(x, 1) = 1$ ,  $G = g_2(x, 1)$  and  $H_1 = h_1(x, 1)$ . The formula for  $F_1$  differs from the corresponding formulas for  $f_1$  in Theorem 1 of the current paper. Nevertheless these formulas are equivalent. We defer the explanation to Remark 3 after stating Theorem 1.

The reader might wonder whether one can utilize composition structures for other fundamental operations. In fact, this has already been done for some op-

erations. For examples, projective (Macaulay) resultant, Gröbner bases ([32]), SAGBI bases, subresultants and Galois groups of certain differential operators have been studied respectively in [29], [16] and [15], [34], [17] and [1] using various mathematical techniques. However, it seems that those techniques cannot be applied to the study of resultants. Therefore in this paper we use mathematical methods that are essentially different from those.

We outline the structure of the paper. Section 2 gives the main (theoretical) results of the paper and Section 3 proves them. Furthermore, Section 4 discusses the computational efficiency of the main results.

## 2 Main results

We assume the reader is familiar with the notions of projective (Sylvester, Macaulay, dense) resultant, toric (sparse) resultant and supports of sparse polynomials (see [9], [13], [35]).

Before we state the main theorem we fix a few notations. Let  $h_1$  be a bivariate homogeneous polynomial in the variables  $x_1, x_2$  of degree  $e_1$ . Let  $f_2$  be a homogeneous bivariate polynomial in the variables  $y_1, y_2$  of degree  $c_2$ . Let  $g_1$  and  $g_2$  be bivariate homogeneous polynomials in the variables  $x_1, x_2$  of equal total degrees, denoted by  $d$ . Let the composed polynomial  $h_2 = f_2 \circ (g_1, g_2)$  be obtained from the polynomial  $f_2$  by replacing  $y_j$  with  $g_j$ . Note that we had to assume that  $g_1$  and  $g_2$  have equal total degrees in order to ensure that  $h_2$  is homogeneous. Let  $\text{Res}_{c_1, c_2}$  and  $\text{Res}_{c_1, c_2, c_3}$  respectively denote the projective (dense, Sylvester/Macaulay) resultant of two bivariate homogeneous polynomials of respective total degrees  $c_1$  and  $c_2$ , and the toric (sparse) resultant of three not necessarily homogeneous polynomials with supports  $\mathcal{C}_1, \mathcal{C}_2$  and  $\mathcal{C}_3$ .

Now we are ready to state the main theorem.

### Theorem 1 (Main theorem)

$$\text{Res}_{e_1, e_2}(h_1, f_2 \circ (g_1, g_2)) = \text{Res}_{c_1, c_2}(f_1, f_2), \quad (2)$$

where  $f_1$  is given by both equalities:

$$f_1 = \text{Res}_{e_1, d}(h_1, y_2 g_1 - y_1 g_2), \quad \text{and} \quad (3)$$

$$f_1 = (-1)^{e_1} \text{Res}_{c_1, c_2, c_3}(h_1, y_1 - g_1, y_2 - g_2). \quad (4)$$

In the above formulas, we have  $e_2 = c_2 d$  and  $c_1 = e_1$ . Furthermore, the set  $\mathcal{C}_1$  is the support of a dense homogeneous bivariate polynomial of degree  $e_1$ . That is,  $\mathcal{C}_1 = \{(e_1, 0), (e_1 - 1, 1), \dots, (0, e_1)\}$ . Whereas the sets  $\mathcal{C}_2 = \mathcal{C}_3$

consist of the origin and the support of a dense homogeneous bivariate polynomial of degree  $d$ . That is,  $\mathcal{C}_2 = \mathcal{C}_3 = \{(0, 0), (d, 0), (d - 1, 1), \dots, (0, d)\}$ . Moreover, we normalize the sign of the resultant  $\text{Res}_{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3}$  such that we have  $\text{Res}_{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3}(x_1^{e_1}, x_2^d, 1) = 1$ .

**Remark 2** Note that the resultants in (3) and (4) eliminate that variables  $x_1, x_2$  rather than  $y_1, y_2$ .

**Remark 3** The formula in (3) can be viewed as a generalization of McKay's and Wang's formula of (1). That is, (1) implies that, using the notation of Section 1,

$$F_1 = \text{Res}_{e_1, d}(H_1, y - G) = \text{Res}_{e_1, d}(h_1, y_2 g_1 - y_1 g_2),$$

where  $y_2 = y, y_1 = 1, g_1(x, 1) = 1, g_2(x, 1) = G$  and  $h_1(x, 1) = H_1$ .

Also note that McKay's and Wang's formula in (1) cannot be easily used for computations because it involves the roots of the polynomial  $H_1$ . On the contrary to this, the formula in (3) does not involve roots and thus can be easily used for computations.

**Remark 4** Since this paper considers projective (dense, Sylvester/Macaulay) resultants of partially composed polynomials, the reader might find it surprising that the polynomial  $f_1$  is expressed in terms of a toric (sparse) resultant (see 4) and not in terms of a projective (dense, Macaulay) resultant. Indeed, one can show that  $f_1$  is also related to a projective resultant. That is, Corollary 5 of [31] implies that the power  $f_1^d$  is the projective (dense, Macaulay) resultant of  $h_1, y_1 - g_1$  and  $g_2 - g_2$  with respect to the total degrees  $e_1, d$  and  $d$  (see Lemma 10).

**Remark 5** Naturally, one asks how Theorem 1 is related to the well-known formula for resultants of composed polynomials derived by [26] in the homogeneous bivariate case. It turns out that one can rewrite resultants of composed polynomials in terms of resultants of linearly combined polynomials by applying Theorem 1 twice. However, it seems that one cannot derive the main result of [26] only by applying Theorem 1.

To illustrate the previous paragraph, in the following we apply Theorem 1 to resultants of homogeneous bivariate composed polynomials twice. Let  $f_1$  and  $f_2$  be homogeneous bivariate polynomial in the variables  $y_1, y_2$  of respective degrees  $c_1$  and  $c_2$ . Let  $g_1$  and  $g_2$  be bivariate homogeneous polynomials in the variables  $x_1, x_2$  of equal total degrees, denoted by  $d$ . Then, by Theorem 1,

$$\text{Res}_{c_1 d, c_2 d}(f_1 \circ (g_1, g_2), f_2 \circ (g_1, g_2)) = \text{Res}_{c_1 d, c_2 d}(p, f_2), \quad (5)$$

where  $p = \text{Res}_{c_1 d, d}(f_1 \circ (g_1, g_2), y_2 g_1 - y_1 g_2)$  which equals, by Corollary 5 of [26],  $(-1)^{c_1 d^2} \text{Res}_{c_1 d, d}(y_2 g_1 - y_1 g_2, f_1 \circ (g_1, g_2))$ . Furthermore, by Theorem 1,

$p = \text{Res}_{d,c_1}(q, f_1)$ , where  $q = \text{Res}_{d,d}(y_2g_1 - y_1g_2, z_2g_1 - z_1g_2)$ , where  $z_1$  and  $z_2$  are new distinct variables. Therefore, indeed, one can use Theorem 1 to rewrite the resultant of two composed polynomials in terms of the resultant of two linearly combined polynomials. If one factors  $q$  into  $(-y_2z_1 - y_1z_2)^d \text{Res}_{d,d}(g_1, g_2)$ , applying Lemma 7 of [26], and if one utilizes the bi-homogeneity of the resultant, one can simplify (5) to obtain McKay's and Wang's formula

$$\text{Res}_{c_1d, c_2d}(f_1 \circ (g_1, g_2), f_2 \circ (g_1, g_2)) = \text{Res}_{c_1, c_2}(f_1, f_2)^d \text{Res}_{d,d}(g_1, g_2)^{c_1c_2}$$

for resultants of two homogeneous bivariate composed polynomials ([26]).

**Remark 6** In the following subsection, ‘‘Computational application of the main theorem’’, we will use Theorem 1 for efficiently computing resultants of partially composed polynomials. The reader will notice that we will not utilize (4). It is important to point out that we have stated (4) because it is of independent theoretical interest. That is, it makes an explicit connection between projective (dense, Sylvester/Macaulay) resultants of two polynomials and bivariable toric (sparse) resultants of three polynomials.

### 3 Proof of the main theorem

In this section we prove Theorem 1. First we prove (2) and (3). Then we prove that the right-hand side of (3) equals the right-hand side of (4).

*Proof of (2) and (3) of Theorem 1*

We start with an auxiliary lemma.

**Lemma 7** *Suppose  $\text{Res}_{e_1,d}(h_1, g_2) \neq 0$ . Then the leading coefficient, with respect to the variable  $z$ , of the polynomial  $\text{Res}_{e_1,d}(h_1, g_1 - z g_2)$  equals the resultant  $\text{Res}_{e_1,d}(h_1, g_2)$  and the degree in  $z$  of the polynomial is  $e_1$ .*

**Proof:** Let  $p(z) = \text{Res}_{e_1,d}(h_1, g_1 - z g_2)$ . By the bi-homogeneity of the resultant, the degree of  $p$  is at most  $e_1$ . Therefore, if  $p^h(1, 0) \neq 0$ , where  $p^h(y_1, y_2) = y_2^{e_1} p(\frac{y_1}{y_2})$ , then the leading coefficient of  $p$  is  $p^h(1, 0)$  and its degree is  $e_1$ . Since  $p^h(1, 0) = \text{Res}_{e_1,d}(h_1, g_2) \neq 0$ , we have shown the lemma.  $\square$

Now we are ready for the next lemma, Lemma 8, which shows (2) and (3) of Theorem 1.

The proof of Lemma 8 extends and generalizes the proof of Theorem 7 of [25]. Note that there is an interesting difference between the two proofs. The proof of Lemma 8 in a first step shows the lemma for polynomials with symbolic (algebraically independent) coefficients and only in a second step it shows the lemma for polynomials with arbitrary coefficients. Whereas, the proof of Theorem 7 of [25] shows the theorem for polynomials with arbitrary coefficients without any first step dealing with symbolic coefficients (compare Remark 3). This approach allows avoiding case distinctions in the proof.

It is also important to point out that one can find a different extension of the proof of Theorem 7 of [25] in the literature. That is, in [26], McKay and Wang extend the techniques presented in [25] in order to derive a product formula for resultants of two homogeneous composed polynomials (see Remark 5). This extension is different from the one included in the proof of Lemma 8. Moreover, it seems not possible to utilize the extended proof techniques presented in [26] to prove Lemma 8 of the current paper.

Furthermore, note that the proof of Lemma 8 is different from the proofs of the results of other papers ([19], [8], [21], [18], [29], [28], [30]) deriving product formulas for various resultants of composed polynomials.

**Lemma 8** *We have*

$$\text{Res}_{e_1, e_2}(h_1, f_2 \circ (g_1, g_2)) = \text{Res}_{c_1, c_2}(f_1, f_2),$$

where

$$f_1 = \text{Res}_{e_1, d}(h_1, y_2 g_1 - y_1 g_2).$$

**Proof:** Let us first assume that all the polynomials  $h_1$ ,  $f_2$ ,  $g_1$  and  $g_2$  have distinct symbolic coefficients. Let  $x$  be a new variable. Then we have by well known properties of the resultant ([23]) that  $\text{Res}_{e_1, e_2}(h_1, f_2 \circ (g_1, g_2)) = \text{Res}_{e_1, e_2}(h_1(x, 1), f_2 \circ (g_1, g_2)(x, 1))$ . Note that the resultant on the left-hand side of this equality eliminates the variables  $x_1$  and  $x_2$  from two homogeneous polynomials. Whereas, on the right-hand side it eliminates the variable  $x$  from two univariate (not necessarily homogeneous) polynomials. Furthermore, let  $\alpha$  range over the roots of  $h_1(x, 1)$ . Then, since  $g_2(\alpha, 1) \neq 0$  and by well known

properties of the resultant (see [25], [23]), we have

$$\begin{aligned}
\text{Res}_{e_1, e_2}(h_1, f_2 \circ (g_1, g_2)) &= h_1(0, 1)^{c_2 d} \prod_{\alpha} f_2 \circ (g_1, g_2)(\alpha, 1) \\
&= h_1(0, 1)^{c_2 d} \prod_{\alpha} f_2(g_1(\alpha, 1), g_2(\alpha, 1)) \\
&= h_1(0, 1)^{c_2 d} \prod_{\alpha} g_1(\alpha, 1)^{c_2} \prod_{\alpha} f_2\left(\frac{g_1(\alpha, 1)}{g_2(\alpha, 1)}, 1\right) \\
&= (\text{Res}_{e_1, d}(h_1, g_2))^{c_2} \prod_{\alpha} f_2\left(\frac{g_1(\alpha, 1)}{g_2(\alpha, 1)}, 1\right).
\end{aligned}$$

Now, observe that  $\beta = \frac{g_1(\alpha, 1)}{g_2(\alpha, 1)}$  for some  $\alpha$  iff

$$\prod_{\alpha} (g_1(\alpha, 1) - \beta g_2(\alpha, 1)) = 0.$$

Since  $h_1(1, 0)$ , the leading coefficient of  $h_1(x, 1)$ , does not vanish, the latter is equivalent to

$$\text{Res}_{e_1, d}(h_1(x, 1), g_1(x, 1) - \beta g_2(x, 1)) = 0,$$

which is equivalent to  $\text{Res}_{e_1, d}(h_1, g_1 - \beta g_2) = 0$ . Therefore and by Lemma 7,

$$\begin{aligned}
\text{Res}_{e_1, e_2}(h_1, f_2 \circ (g_1, g_2)) &= \\
&(\text{Res}_{e_1, d}(h_1, g_2))^{c_2} \times \prod_{\substack{\beta \\ \text{Res}_{e_1, d}(h_1, g_1 - \beta g_2) = 0}} f_2(\beta, 1) = \\
&(\text{Res}_{e_1, d}(h_1, g_2))^{c_2} \times \frac{\text{Res}_{e_1, c_2}(\text{Res}_{e_1, d}(h_1, g_1 - y g_2), f_2(y, 1))}{(\text{Res}_{e_1, d}(h_1, g_2))^{c_2}} = \\
&\text{Res}_{e_1, c_2}(f_1, f_2).
\end{aligned}$$

Therefore we have shown Lemma 8 for polynomials with symbolic coefficients.

Next, observe that the formulas of Lemma 8 are stable under specialization. Therefore Lemma 8 also holds for polynomials with arbitrary coefficients.  $\square$

*Proof that the right-hand side of (3) of Theorem 1 equals the right-hand side of (4)*

We assume that the reader is familiar with the notions of integer lattice affinely generated by supports of polynomials, Newton polytope, support of a polynomial, normalized mixed volume and normalized volume of polytopes, normalized with respect to the elementary simplex of an integer lattice, and lattice index (see [9], [13], [35]).

Before we state some results, we fix some notation. Let  $\text{Res}_{d_1, d_2, d_3}(p_1, p_2, p_3)$  denote the projective (dense, Macaulay) resultant of three bivariate polynomials  $p_1, p_2$  and  $p_3$  in the variables  $x_1$  and  $x_2$  with total degrees  $d_1, d_2$  and  $d_3$ .

The next lemma relates the right-hand side of (3) of Theorem 1 to the bivariable projective (dense, Macaulay) resultant.

**Lemma 9**

$$\text{Res}_{e_1, d}(h_1, y_2 g_1 - y_1 g_2) = 0$$

*implies that*

$$\text{Res}_{e_1, d, d}(h_1, y_1 - g_1, y_2 - g_2) = 0.$$

**Proof:** Let  $y_1$  and  $y_2$  be such that  $\text{Res}_{e_1, d}(h_1, y_2 g_1 - y_1 g_2) = 0$ . Therefore there is  $(\alpha_1, \alpha_2) \neq (0, 0)$  such that

$$\begin{aligned} h_1(\alpha_1, \alpha_2) &= 0, \\ y_2 g_1(\alpha_1, \alpha_2) - y_1 g_2(\alpha_1, \alpha_2) &= 0. \end{aligned} \tag{6}$$

Fix such  $(\alpha_1, \alpha_2)$ .

First assume that  $g_1(\alpha_1, \alpha_2) = g_2(\alpha_1, \alpha_2) = 0$ . Therefore the leading forms of  $h_1, y_1 - g_1$  and  $y_2 - g_2$ , viewed as polynomials in the variables  $x_1$  and  $x_2$ , vanish if  $x_1$  and  $x_2$  are replaced by  $\alpha_1$  and  $\alpha_2$ . Therefore  $\text{Res}_{e_1, d, d}(h_1, y_1 - g_1, y_2 - g_2) = 0$ .

Next assume that  $(g_1(\alpha_1, \alpha_2), g_2(\alpha_1, \alpha_2)) \neq 0$ . Then (6) implies that  $(y_1, y_2) = \beta(g_1(\alpha_1, \alpha_2), g_2(\alpha_1, \alpha_2))$ , for some  $\beta$ . Since  $g_1$  and  $g_2$  are homogeneous of equal degree, we have that  $(y_1, y_2) = (g_1(\gamma\alpha_1, \gamma\alpha_2), g_2(\gamma\alpha_1, \gamma\alpha_2))$ , for some  $\gamma$ . Note that also  $h_1(\gamma\alpha_1, \gamma\alpha_2) = 0$  because  $h_1$  is homogeneous. Therefore  $\text{Res}_{e_1, d, d}(h_1, y_1 - g_1, y_2 - g_2) = 0$ .  $\square$

The next lemma shows how the right-hand side of (4) is related to the bivariable projective (dense, Macaulay) resultant.

**Lemma 10** *We have that*

$$(\text{Res}_{c_1, c_2, c_3}(h_1, y_1 - g_1, y_2 - g_2))^d = \text{Res}_{e_1, d, d}(h_1, y_1 - g_1, y_2 - g_2).$$

Note that the magnitude of the exponent on the resultant

$$\text{Res}_{c_1, c_2, c_3}(h_1, y_1 - g_1, y_2 - g_2)$$

in Lemma 10 is not relevant for the proof of Theorem 1. However, the exponent answers a natural question included in Remark 4. Therefore this exponent is determined in Lemma 10.

**Proof:** Let  $\mathcal{D}$  be the set of all integer points in the triangle with vertices  $(0, 0)$ ,  $(d, 0)$  and  $(0, d)$ . By Theorem 1 of [31],

$$\begin{aligned} \text{Res}_{e_1, d, d}(h_1, y_1 - g_1, y_2 - g_2) &= (\text{Res}_{\mathcal{C}_1, \mathcal{D}, \mathcal{D}}(h_1, y_1 - g_1, y_2 - g_2))^\delta \\ &\quad \times \left( \text{Res}_{\{(0,0)\}, \{(0,0)\}}(y_1, y_2) \right)^\epsilon, \end{aligned}$$

where  $\delta$  is the lattice index of the integer lattice generated by  $\mathcal{C}_1$  and  $\mathcal{D}$  in the integer lattice of all integer points  $(i, j)$ . Since the integer lattice generated by  $\mathcal{C}_1$  and  $\mathcal{D}$  equals the integer lattice of all integer points  $(i, j)$ , we have  $\delta = 1$ . Furthermore, by definition,  $\text{Res}_{\{(0,0)\}, \{(0,0)\}}(y_1, y_2) = 1$ . Thus

$$\text{Res}_{e_1, d, d}(h_1, y_1 - g_1, y_2 - g_2) = \text{Res}_{\mathcal{C}_1, \mathcal{D}, \mathcal{D}}(h_1, y_1 - g_1, y_2 - g_2).$$

Now, by applying Theorem 1 of [31] twice, similarly to the proof of Corollary 5 of [31], we get that

$$\text{Res}_{\mathcal{C}_1, \mathcal{D}, \mathcal{D}}(h_1, y_1 - g_1, y_2 - g_2) = (\text{Res}_{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3}(h_1, y_1 - g_1, y_2 - g_2))^\phi,$$

where  $\phi$  is the index of the integer lattice affinely generated by  $\mathcal{C}_1$ ,  $\mathcal{C}_2$  and  $\mathcal{C}_3$  in the lattice of all integer points  $(i, j)$ . Observe that the integer lattice affinely generated by  $\mathcal{C}_2 = \mathcal{C}_3$  is the set of all points  $k(-1, 1) + l(0, d)$ , where  $k$  and  $l$  range over the integers. This lattice includes the lattice of all points  $k(-1, 1)$  which is the lattice affinely generated by  $\mathcal{C}_1$ . Therefore the elementary simplex of the lattice generated by  $\mathcal{C}_1$ ,  $\mathcal{C}_2$  and  $\mathcal{C}_3$  is the triangle with the vertices  $(0, 0)$ ,  $(d - 1, 1)$  and  $(d, 0)$ . Its volume (area) is  $\frac{d}{2}$ . Next, observe that the volume (area) of the elementary simplex of the lattice of all integer points  $(i, j)$  is  $\frac{1}{2}$ . Therefore  $\phi = \frac{d/2}{1/2} = d$ .  $\square$

The following observation is important for showing that the right-hand side of (4) equals the right-hand side of (3).

**Observation 11** *If  $h_1$ ,  $g_1$  and  $g_2$  have distinct symbolic coefficients, distinct from the distinct symbols  $y_1$  and  $y_2$ , then  $\text{Res}_{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3}(h_1, y_1 - g_1, y_2 - g_2)$  is absolutely irreducible because  $\text{Res}_{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3}(h_1, y_1 + g_1, y_2 + g_2)$  is irreducible by definition.*

Even though the following lemma is well-known, it is presented here for the convenience of the reader. We also note that the argument provided by the lemma has already been used in another related publication ([18]).

**Lemma 12** *Let  $p$  and  $q$  be (multi-variate) non-constant polynomials over an algebraically closed field. Furthermore, let  $p = 0$  imply  $q = 0$ . Moreover, let  $q$  be irreducible. Then  $p = \lambda q^\delta$ , for a constant  $\lambda$  and a positive integer  $\delta$ .*

**Proof:** By Hilbert's Nullstellensatz, the polynomial  $q$  is in the radical generated by the polynomial  $p$ . That is,  $q^\epsilon = r p$ , for some integer  $\epsilon$  and a polynomial  $r$ . Since the polynomial  $q$  is irreducible,  $q^\epsilon$  is the irreducible factorization of the polynomial  $r p$ . Therefore, the only irreducible factor of the polynomial  $p$  is  $q$ . Thus,  $p = \lambda q^\delta$ , for a constant  $\lambda$  and a positive integer  $\delta$ .  $\square$

Now we are ready to show that the right-hand side of (4) equals the right-hand side of (3).

**Lemma 13** *We have that*

$$\text{Res}_{e_1,d}(h_1, y_2 g_1 - y_1 g_2) = (-1)^{e_1} \text{Res}_{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3}(h_1, y_1 - g_1, y_2 - g_2).$$

**Proof:** By Lemmas 9 and 10, for all  $h_1, g_1, g_2, y_1$  and  $h_2$ ,

$$\text{Res}_{e_1,d}(h_1, y_2 g_1 - y_1 g_2) = 0$$

implies that

$$\text{Res}_{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3}(h_1, y_1 - g_1, y_2 - g_2) = 0.$$

By Observation 11 and Lemma 12, we have that

$$\text{Res}_{e_1,d}(h_1, y_2 g_1 - y_1 g_2) = \lambda \times (\text{Res}_{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3}(h_1, y_1 - g_1, y_2 - g_2))^\delta,$$

where  $\lambda$  is a factor only depending on  $\mathcal{C}_1, \mathcal{C}_2$  and  $\mathcal{C}_3$ . In the following, we determine the constants  $\delta$  and  $\lambda$ .

In order to determine the exponent  $\delta$ , we compare the total degrees in the coefficients of  $h_1$  of the left-hand side and of the right-hand side of the above equality. The total degree of the left hand-side is  $d$ . Furthermore, the total degree of the right-hand side is  $\delta$  times the mixed volume, normalized with respect to the integer lattice affinely generated by  $\mathcal{C}_2$  and  $\mathcal{C}_3$ , of the Newton polytopes generated by  $\mathcal{C}_2$  and  $\mathcal{C}_3$ . Since  $\mathcal{C}_2 = \mathcal{C}_3$ , the total degree of the right-hand side is twice the volume, normalized with respect to the integer lattice affinely generated by  $\mathcal{C}_2$ , of the Newton polytope generated by  $\mathcal{C}_2$  (see [9]). The (not normalized) volume of the Newton polytope generated by  $\mathcal{C}_2$  is the area of the triangle with vertices  $(0, 0)$ ,  $(d, 0)$  and  $(0, d)$ , which is  $\frac{d^2}{2}$ . In the proof of Lemma 10, we have already seen that the volume of the elementary

simplex of the integer lattice affinely generated by  $\mathcal{C}_2$  is  $\frac{d}{2}$ . Therefore, the total degree of the right-hand side is  $\frac{d^2}{2}/\frac{d}{2} = d$ . Thus  $\delta = 1$  and

$$\text{Res}_{e_1, d}(h_1, y_2 g_1 - y_1 g_2) = \lambda \times \text{Res}_{c_1, c_2, c_3}(h_1, y_1 - g_1, y_2 - g_2).$$

In order to determine  $\lambda$ , specialize  $h_1$  to  $x_1^{e_1}$ ,  $g_1$  to  $-x_2^d$ ,  $g_2$  to 0,  $y_1$  to 0 and  $y_2$  to 1 in the above equality. Since we have normalized  $\text{Res}_{c_1, c_2, c_3}$  as described in Theorem 1, we have

$$(-1)^{e_1} = \lambda \times 1$$

and thus we have shown the lemma.  $\square$

*Proof of the main theorem, Theorem 1*

By combining Lemma 8 and Lemma 13 we have shown Theorem 1.  $\square$

## 4 Computational efficiency of main results

In this section we describe how one can apply Theorem 1 to efficiently compute resultants of partially composed polynomials.

*Step 1: Computation of  $\mathbf{f}_1$*

We ask the reader to examine the resultant in (3). Note that the bi-homogeneity of this resultant implies that the polynomial  $f_1$  is homogeneous in the variables  $y_1$  and  $y_2$ . Furthermore the total degree of  $f_1$  is  $e_1$ . Thus, in order to compute  $f_1$  it is sufficient to compute the polynomial  $p(y_1) = \text{Res}_{e_1, d}(h_1(y_1, 1), g_1 - y_1 g_2)$ . This polynomial  $p$  can be computed via interpolation letting  $y_1$  range over the values  $0, 1, \dots, e_1$ .

*Step 2: Computation of  $\mathbf{Res}_{c_1, c_2}(\mathbf{f}_1, \mathbf{f}_2)$*

Note that  $f_1$  and  $f_2$  are *bivariate homogeneous* polynomials. Therefore the resultant  $\text{Res}_{c_1, c_2}(f_1, f_2)$  can be computed as the univariable (Sylvester) resultant  $\text{Res}_{c_1, c_2}(f_1(y_1, 1), f_2(y_1, 1))$ .

### Running Time experiments

Now, we discuss some practical running time experiments carried out under Maple 9 on a PC with a 2.2 GHz processor and 3 GB main memory. For this subsection, we assume that all the polynomials  $h_1, f_2, g_1, g_2$  have **integer coefficients modulo a fixed 32 bit prime number**. The author has measured how the running times of the method described in Step 1 and Step 2 above compare to the running times of computing resultants of partially composed polynomials without utilizing the composition structure of  $f_2 \circ (g_1, g_2)$ . For the rest of this subsection, in order to be able to easily compare both methods, we refer to the first method with “UseStruc” (use the structure via Step 1 and Step 2) and to the second one with “NoStruc” (do not use the structure, expand the composed polynomial and compute the resultant).

The measurements have been taken for random dense  $g_1$ 's and  $g_2$ 's of equal degrees ranging from 10 to 30 and for random dense  $h_1$ 's and  $f_2$ 's of degrees independently ranging from 10 to 30 as well. This choice of inputs results in a large amount of computations and running times measured. That is, the degrees  $(c_2, d, e_1)$  of the inputs range over the set  $\{10, \dots, 30\}^3$  and for each triple in the latter set we get running time measurements. In order to make the presentation of the timings more compact, we compute averages of the running times in a systematic way described as follows. For fixed degree  $e_1$  of  $h_1$ , we partition the set  $\{10, \dots, 30\}^2 \times \{e_1\}$  into small sets of four triples. That is, these partitioning sets are  $P_{l,e_1} = \{10 + 2l, 10 + (2l + 1)\}^2 \times \{e_1\} = \{(10 + 2l, 10 + 2l, e_1), (10 + 2l, 10 + (2l + 1), e_1), (10 + (2l + 1), 10 + 2l, e_1), (10 + (2l + 1), 10 + (2l + 1), e_1)\}$ . For each triple in  $P_{l,e_1}$ , we generate random polynomials of corresponding degrees and measure the running times of methods UseStruc and NoStruc. Then we compute the averages  $\text{time}_{l,e_1}^{\text{UseStruc}}$  and  $\text{time}_{l,e_1}^{\text{NoStruc}}$ , of these measured times for the four triples in  $P_{l,e_1}$ . One can observe that these averages vary not very much as  $e_1$  ranges from 10 to 30. Thus we compute the averages  $\text{time}_l^{\text{UseStruc}}$  and  $\text{time}_l^{\text{NoStruc}}$ , for  $e_1$  ranging from 10 to 30, further simplifying the presentation of the running times but still remaining faithful to the experimental measurements. Finally, these values are listed in Table 1.

The author believes that intuitively it is not surprising that the averages  $\text{time}_{l,e_1}^{\text{UseStruc}}$  and  $\text{time}_{l,e_1}^{\text{NoStruc}}$  vary little for varying  $e_1$ . That is,  $e_1$ , the degree of the unstructured  $h_1$ , is relatively small in comparison to the degree of the composed polynomial  $f_2 \circ (g_1, g_2)$ . Therefore, changes of  $e_1$  have little impact on the running time. Furthermore, note that in this case utilizing the composition structure is also very efficient computationally. If  $e_1$  becomes larger then the efficiency of Step 1 and Step 2 decreases. This behavior is expected because, intuitively, for large  $e_1$ , in comparison to the degree of the composed polynomial  $f_2 \circ (g_1, g_2)$ , one expects to achieve only little or even no gain in efficiency through utilizing the composition structure of  $f_2 \circ (g_1, g_2)$ .

$l$	$\text{time}_l^{\text{NoStruc}}$ in sec.	$\text{time}_l^{\text{UseStruc}}$ in sec.
Application of Theorem 1		
0	0.763	.025
1	1.320	.027
2	3.059	.027
3	4.902	.028
4	7.675	.030
5	12.414	.031
6	18.843	.031
7	31.393	.033
8	58.322	.035
9	99.768	.036

Fig. 1. Running times for increasing degrees of  $f_2, g_1, g_2$ . Averages for  $(c_2, d, e_1)$  in  $\{10 + 2l, 10 + 2l + 1\}^2 \times \{10, 11, \dots, 30\}$ .

In Table 1 one can see that the speedup of Method UseStruc (Theorem 1 applied in Step 1 and Step 2) is quite dramatic as  $l$ , i.e. the degrees of  $f_2, g_1$  and  $g_2$ , increases.

## 5 Conclusion

This paper has studied resultants of partially composed polynomials. We have found that these resultants are certain iterated resultants of the component polynomials. Furthermore, we saw in experiments that, in many cases, these iterated resultants can be computed with dramatically increased efficiency.

Future research might address multi-variable generalizations of the results of this paper.

## References

- [1] P. H. Berman and M. F. Singer. Calculating the Galois group of  $L_1(L_2(y)) = 0$ ,  $L_1, L_2$  completely reducible operators. *J. Pure Appl. Algebra*, 139(1-3):3–23, 1999. Effective methods in algebraic geometry (Saint-Malo, 1998).

- [2] L. Busé, M. Elkadi, and B. Mourrain. Generalized resultants over unirational algebraic varieties. *J. Symbolic Computation*, 29(4-5):515–526, 2000.
- [3] J. Canny and I. Emiris. A subdivision-based algorithm for the sparse resultant. *J. ACM*, 47(3):417–451, May 2000.
- [4] J. Canny, E. Kaltofen, and Lakshman Y. Solving systems of nonlinear polynomial equations faster. In *Proc. ACM-SIGSAM 1989 Internat. Symp. Symbolic Algebraic Comput. ISSAC’89*, pages 121–128. ACM, 1989.
- [5] E. Cattani, A. Dickenstein, and B. Sturmfels. Residues and resultants. *J. Math. Sci. Univ. Tokyo*, 5(1):119–148, 1998.
- [6] A. Cayley. On the theory of elimination. *Cambridge and Dublin Math. J.*, 3:116–120, 1848.
- [7] M. Chardin. *Contributions à l’algèbre commutative effective et à la théorie de l’élimination*. PhD thesis, Université Paris VI, 1990.
- [8] 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.
- [9] D. Cox, J. Little, and D. O’Shea. *Using Algebraic Geometry*. Springer Verlag, New York, Berlin, Heidelberg, 1998.
- [10] C. D’Andrea and A. Dickenstein. Explicit formulas for the multivariate resultant. *Pure Appl. Algebra*, 164(1-2):59–86, 2001.
- [11] A.-L. Dixon. The eliminant of three quantics in two independent variables. *Proc. London Math. Soc.*, 7:49–69, 473–492, November 1908.
- [12] I. Z. Emiris and V. Pan. The structure of sparse resultant matrices. In *Proc. Int. Symp. on Symbolic and Algebraic Computation (ISSAC)*. ACM Press, 1997.
- [13] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, Resultants and Multidimensional Determinants*. Birkhäuser, Boston, 1994.
- [14] L. González-Vega. Une théorie des sous-résultants pour les polynômes en plusieurs variables. *C. R. Acad. Sci. Paris Sér. I Math.*, 313/13:905–908, 1991.
- [15] J. Gutierrez and R. Rubio San Miguel. Reduced Gröbner bases under composition. *J. Symbolic Computation*, 26(4):433–444, 1998.
- [16] H. Hong. Subresultants under composition. *J. Symbolic Computation*, 23(4):355–365, 1997.
- [17] H. Hong. Groebner basis under composition I. *J. Symbolic Computation*, 25(5):643–663, 1998.
- [18] H. Hong and M. Minimair. Sparse resultant of composed polynomials I. *J. Symbolic Computation*, 33:447–465, 2002.
- [19] J. P. Jouanolou. Le formalisme du résultant. *Adv. Math.*, 90(2):117–263, 1991.
- [20] D. Kapur and T. Saxena. Sparsity considerations in Dixon resultants. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing (Philadelphia, PA, 1996)*, pages 184–191, New York, 1996.

- ACM.
- [21] D. Kapur and T. Saxena. Extraneous factors in the Dixon resultant formulation. In *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation (Kihei, HI)*, pages 141–148, New York, 1997. ACM.
  - [22] R. Lewis and P. Stiller. Solving the recognition problem for six lines using the Dixon resultant. *Math. Comput. Simulation*, 49(3):205–219, 1999.
  - [23] F. S. Macaulay. *The algebraic theory of modular systems*. Cambridge Mathematical Library, 1916.
  - [24] D. Manocha and J. Canny. Multipolynomial resultant algorithms. *J. Symbolic Computation*, 15(2):99–122, 1993.
  - [25] J. H. McKay and S. S. Wang. A chain rule for the resultant of two polynomials. *Arch. Math.*, 53(4):347–351, 1989.
  - [26] J. H. McKay and S. S. Wang. A chain rule for the resultant of two homogeneous polynomials. *Arch. Math.*, 56(4):352–361, 1991.
  - [27] M. Minimair. *Sparse Resultants of Composed Polynomials*. PhD thesis, North Carolina State University, Raleigh, NC, USA, 2001.
  - [28] M. Minimair. Sparse resultant of composed polynomials II. *J. Symbolic Computation*, 33:467–478, 2002.
  - [29] M. Minimair. Dense resultant of composed polynomials. *J. Symbolic Computation*, 36(6):825–834, December 2003.
  - [30] 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.
  - [31] M. Minimair. Sparse resultant under vanishing coefficients. *J. Algebraic Combinatorics*, 18(1):53–73, July 2003.
  - [32] M. Minimair and M. B. Barnett. Solving polynomial equations for chemical problems using Gröbner bases. *Molecular Physics*, 102(23-24):2521–2535, December 2004.
  - [33] G. Nakos and R. M. Williams. Elimination with the Dixon resultant. *Mathematica for Education and Research*, 6/3:11–21, 1997.
  - [34] P. Nordbeck. SAGBI bases under composition. *J. Symbolic Computation*, 33(1):67–76, 2002.
  - [35] P. Pedersen and B. Sturmfels. Product formulas for resultants and Chow forms. *Mathematische Zeitschrift*, 214:377–396, 1993.
  - [36] J. M. Rojas. Solving degenerate sparse polynomial systems faster. *J. Symbolic Computation*, 28(1 and 2):155–186, July/August 1999. Special Issue Polynomial Elimination – Algorithms and Applications.