On Macaulay's Form of the Resultant

K. Kalorkoti
School of Informatics, University of Edinburgh
(kk@inf.ed.ac.uk)

Abstract

Macaulay's form of the resultant, as the ratio of two determinants, has been used to good effect in Computer Algebra applications. The first part of this paper gives a shorter self contained version of Macaulay's proof in modern notation. One problem with Macaulay's form is that under some conditions the denominator can vanish. This problem was addressed by Canny in 1990. Here we present an alternative solution. We also present a method for computing the u-resultant that does not suffer from exceptional cases and study the special case when the given forms have no common zeros at infinity.

§1. Introduction. Let $x_1, x_2, \ldots,$ be indeterminates over an algebraically closed field k. Given $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}^n$ we set $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ and $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ (we include 0 in $\mathbb{N}$). For a given $n > 0$, fix n non-zero degrees $d_1, d_2, \ldots, d_n$ and distinct indeterminates $u_{i,\alpha}$ for $1 \le i \le n$ and $\alpha \in \mathbb{N}^n$ with $|\alpha| = d_i$. Let $F_i = \sum_{|\alpha|=d_i} u_{i,\alpha} x^\alpha$ be n generic homogeneous forms. As is well known there is a polynomial, called the *resultant*, $\mathrm{Res}(F_1, F_2, \ldots, F_n)$ in the $u_{i,\alpha}$ with integer coefficients such that:

1. For all specializations of the $u_{i,\alpha}$ to values from k, the resulting homogeneous forms have a non-trivial zero if and only if the resultant vanishes.
2. The resultant is irreducible over k.
3. Set $D = d_1 d_2 \cdots d_n$. For each i, the resultant is homogeneous of degree D/d_i in the $u_{i,\alpha}$ and of total degree $\sum_{i=1}^n D/d_i$.
4. $\mathrm{Res}(x_1^{d_1}, x_2^{d_2}, \ldots, x_n^{d_n}) = 1$.

See van der Waerden [**14**], Jouanolou [**5**] or Gelfand, Kapranov and Zelevinsky [**3**] for material on resultants. For $s \ge 2$ we will use $F_i^{(s)}$ to denote the generic form in $x_1, x_2, \ldots, x_{s-1}$ obtained from F_i by the substitution $x_i \mapsto 0$ for all $i \ge s$.

§2. Macaulay's Construction. Given a degree t set

$$
\begin{aligned}
S(n,t,0) &= \{x^\alpha \mid \alpha \in \mathbb{N}^n \text{ with } |\alpha| = t \text{ and } x_1^{d_1} \mid x^\alpha\},\\
S(n,t,1) &= \{x^\alpha \mid \alpha \in \mathbb{N}^n \text{ with } |\alpha| = t \text{ and } x_2^{d_2} \mid x^\alpha \text{ and } x_1^{d_1} \nmid x^\alpha\},\\
&\ \ \vdots\\
S(n,t,n-1) &= \{x^\alpha \mid \alpha \in \mathbb{N}^n \text{ with } |\alpha| = t \text{ and } x_n^{d_n} \mid x^\alpha \text{ and } x_1^{d_1} \nmid x^\alpha \text{ and } \ldots \text{ and } x_{n-1}^{d_{n-1}} \nmid x^\alpha\}.\\
S(n,t,n) &= \{x^\alpha \mid \alpha \in \mathbb{N}^n \text{ with } |\alpha| = t \text{ and } x_1^{d_1} \nmid x^\alpha \text{ and } \ldots \text{ and } x_n^{d_n} \nmid x^\alpha\}.
\end{aligned}
$$

Note that these sets are disjoint and $S(n,t,i)$ is empty for $0 \le i \le n-1$ if and only if $t < d_{i+1}$ while $S(n,t,n)$ is empty if and only if $t \ge d$ where

$$
d = 1 + \sum_{i=1}^n (d_i - 1),
$$

(d has this meaning throughout the paper). We use $x^\beta S(n,t,i)$ to denote $\{x^\beta x^\alpha \mid x^\alpha \in S(n,t,i)\}$. We record a couple of simple observations.

Lemma 2.1 *For $0 \le i \le n-2$ the power products from $S(n,t-1,i)$ are in 1-1 correspondence with the power products of form $x^\alpha x_n^{-1}$ where $x^\alpha \in S(n,t,i)$ and $x_n \mid x^\alpha$. Moreover the power products from $S(n,t-1,n-1)$ are in 1-1 correspondence with the power products of form $x^\alpha x_n^{-1}$ where $x_n^\alpha \in S(n,t,n-1)$ and $x_n^{d_n+1} \mid x^\alpha$.*

PROOF. For $0 \le i \le n-2$ it is clear that every power product of the given form is in $S(n,t-1,i)$. Conversely if $x^\beta \in S(n,t-1,i)$ then $x^\beta x_n \in S(n,t,i)$.

For the second part, if $x^\beta \in S(n,t-1,n-1)$ then $x_n^{d_n} \mid x^\beta$ so that $x^\beta x_n \in S(n,t,n-1)$ and $x_n^{d_n+1} \mid x^\beta$. The converse is clear. □

LEMMA 2.2 *For $0 \le i \le t$ the power products from $S(n-1,t-i,0), S(n-1,t-i,1), \ldots, S(n-1,t-i,n-2)$ are in 1-1 correspondence with the power products of form $x^\beta x_n^i$ such that $\beta \in \mathbb{N}^{n-1}$, $|\beta| = t-i$ and $x^\beta x_n^i \notin S(n,t,n-1) \cup S(n,t,n)$.*

PROOF. If a power product is of the given form then $|\beta| = t-i$ and $x_j^{d_j} \mid x^\beta$ for some $1 \le j \le n-1$. It follows that x^β is in one of the given sets. Conversely if $x^\beta \in S(n-1,t-i,j)$ for some $0 \le j \le n-2$ then $x_{j+1}^{d_{j+1}} \mid x^\beta$ and so $x^\beta x_n^i \notin S(n,t,n-1) \cup S(n,t,n)$. □

Consider the following forms of degree t:

$$\begin{aligned}
&(x^\alpha/x_1^{d_1})F_1, \quad x^\alpha \in S(n,t,0),\\
&(x^\alpha/x_2^{d_2})F_2, \quad x^\alpha \in S(n,t,1),\\
&\qquad\qquad\vdots\\
&(x^\alpha/x_n^{d_n})F_n, \quad x^\alpha \in S(n,t,n-1).
\end{aligned}$$

These define a matrix consisting of the generic coefficients (the $u_{i,\alpha}$) whose rows are indexed by the elements of $S(n,t,0), S(n,t,1), \ldots, S(n,t,n-1)$ and columns by the power products of degree t. If we delete the columns (if any) that are indexed by the elements of $S(n,t,n)$ then we obtain a (possibly empty) square matrix which we denote by $M(F_1,F_2,\ldots,F_n;t)$. We will use $D(F_1,F_2,\ldots,F_n;t)$ to denote the determinant of this matrix (if the matrix is empty then we define its determinant to be 1). Let v_i be the coefficient of $x_i^{d_i}$ in F_i for $1 \le i \le n$. Then each diagonal entry of $M(F_1,F_2,\ldots,F_n;t)$ is one of $v_1, v_2, \ldots, v_n$. To see this suppose that $x^\alpha \in S(n,t,i-1)$ so that the entries of the row indexed by x^α are the coefficients of $(x^\alpha/x_i^{d_i})F_i$. Set $F_i = v_i x_i^{d_i} + \widehat{F}_i$ so that $(x^\alpha/x_i^{d_i})F_i = v_i x^\alpha + (x^\alpha/x_i^{d_i})\widehat{F}_i$ and so the entry indexed by (x^α, x^α) is v_i as claimed. It follows that $D(F_1,F_2,\ldots,F_n;t) \neq 0$ since $M(x_1^{d_1}, x_2^{d_2}, \ldots, x_n^{d_n};t)$ is the identity matrix.

Of course there is some ambiguity in our notation since the order of the rows and columns of $M(F_1,F_2,\ldots,F_n;t)$ has not been fixed. This is not a real problem provided we understand equalities involving determinants to be up to sign.

Suppose that $t \ge d$ and specialize the generic coefficients of the forms to values from k. If the resulting forms have a non-trivial common zero then it is clear that the rows of the matrix $M(F_1,F_2,\ldots,F_n;t)$ are linearly dependent and so $D(F_1,F_2,\ldots,F_n;t) = 0$. In other words, $D(F_1,F_2,\ldots,F_n;t)$ vanishes whenever $\operatorname{Res}(F_1,F_2,\ldots,F_n)$ vanishes. The irreducibility of the resultant and the algebraic closure of k now imply that

$$D(F_1,F_2,\ldots,F_n;t) = \operatorname{Res}(F_1,F_2,\ldots,F_n)\Delta(F_1,F_2,\ldots,F_n;t) \tag{1}$$

for some non-zero polynomial $\Delta(F_1,F_2,\ldots,F_n;t)$. Note that $D(F_1,F_2,\ldots,F_n;t)$ has degree at most (in fact equal to) D/d_n in the coefficients of F_n because $M(F_1,F_2,\ldots,F_n;t)$ has this many rows consisting of these coefficients. Since $\operatorname{Res}(F_1,F_2,\ldots,F_n)$ has degree exactly D/d_n in the same coefficients it follows, as Macaulay observed, that $\Delta(F_1,F_2,\ldots,F_n;t)$ is independent of the coefficients of F_n. In fact Macaulay proved that the extraneous factor $\Delta(F_1,F_2,\ldots,F_n;t)$ is given by a minor of $M(F_1,F_2,\ldots,F_n;t)$. To obtain this minor we delete all rows and columns that are indexed by any power product that is divisible by exactly one $x_i^{d_i}$ for $1 \le i \le n$. The aim of this section is to provide a modern (and shorter) version of the proof of this result which is self contained. (Proofs can also be found in the book by Gröbner [**4**] and the paper by Jouanolou [**6**].)

An examination of Macaulay's proof shows that for all $t \ge 1$ the matrix $M(F_1,F_2,\ldots,F_n;t)$ has a very useful structure provided we (partially) order the power products as follows.

	T_1	T_2	T_3
T_1	A		
T_2	0	B	
T_3	C_1	C_2	C_3

Figure 1: The structure of $M(F_1, F_2, \ldots, F_n; t)$ with power products partially ordered.

		T_1	T_2			
		1	x_n	x_n^2	x_n^3	$\ldots$
T_1	1	A_0				$\ldots$
	x_n	0	A_1			$\ldots$
	x_n^2	0	0	A_2		$\ldots$
T_2	x_n^3	0	0	0	A_3	$\ldots$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Figure 2: The structure of the part of $M(F_1, F_2, \ldots, F_n; t)$ indexed by T_1, T_2.

T_1 : First take those power products not divisible by x_n.

T_2 : Then take those power products that are divisible by x_n but are not in $S(n, t, n-1)$; order these by the highest power of x_n that divides them.

T_3 : Finally take the power products from $S(n, t, n-1)$.

The structure of $M(F_1, F_2, \ldots, F_n; t)$ when indexed by T_1, T_2, T_3 is shown in Figure 1 (the unmarked blocks do not play a significant role below). Note that A is $M\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}; t\big)$. Moreover if we let T_3' be all the power products in T_3 that are divisible by $x_n^{d_n+1}$ then, by Lemma 2.1, the sub-matrix indexed by T_2, T_3' is $M(F_1, F_2, \ldots, F_n; t-1)$. The sub-matrix indexed by T_1, T_2 has a finer structure that is determined by the largest power of x_n that divides each indexing power product. This is shown in Figure 2 in which A_0 is A. It follows from Lemma 2.2 that that

$$A_i = M\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}; t-i\big), \tag{2}$$

for all i (see also Gröbner [**4**, p.65]). The matrices C_1, C_2 and C_3 in Figure 1 consist of coefficients of F_n and the diagonal entries of C_3 consist of the coefficient of $x_n^{d_n}$ in F_n.

Lemma 2.3 *For all $t \geq 1$,*

1. $D(F_1, F_2, \ldots, F_{n-1}, x_n^{d_n}; t) = D\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}; t\big) D(F_1, F_2, \ldots, F_{n-1}, x_n^{d_n}; t-1)$.
2. $D(F_1, F_2, \ldots, F_{n-1}, x_n^{d_n}; t) = \prod_{i=0}^{t-1} D\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}; t-i\big)$.

proof. If we set $F_n = x_n^{d_n}$ then C_1, C_2 are both zero while C_3 is the identity matrix. This, together with the observations preceding (2), proves the first part. The second part follows by induction on t or directly from (2). □

The identities of the Lemma are contained in the proof of the Theorem in §5 of Macaulay's original paper. Gröbner [**4**] introduced a short cut in Macaulay's proof using the fact that

$$\operatorname{Res}(F_1, F_2, \ldots, F_{n-1}, x_n^e) = \operatorname{Res}\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}\big)^e,$$

which follows from

$$\begin{aligned}\operatorname{Res}(F_1, F_2, \ldots, F_n F_n') &= \operatorname{Res}(F_1, F_2, \ldots, F_n) \operatorname{Res}(F_1, F_2, \ldots, F_n'),\\ \operatorname{Res}(F_1, F_2, \ldots, F_{n-1}, x_n) &= \operatorname{Res}\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}\big).\end{aligned}$$

These follow from the basic properties of resultants given in §1 (although van der Waerden [**14**] proves the first of these and uses it to establish the third property of §1). Bearing in mind that the extraneous factor Δ is independent of the coefficients of F_n we have:

$$D(F_1, F_2, \ldots, F_{n-1}, x_n^{d_n}; t) = \operatorname{Res}\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}\big)^{d_n} \Delta(F_1, F_2, \ldots, F_n; t). \tag{3}$$

Theorem 2.1 *For all $t \geq d$ we have*

$$\Delta(F_1, F_2, \ldots, F_n; t) = \prod_{i=0}^{d_n-1} \Delta\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}; t-i\big) \prod_{i=d_n}^{t-1} D\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}; t-i\big)$$

proof. We have

$$D\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}; t-i\big) = \operatorname{Res}\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}\big) \Delta\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}; t-i\big)$$

for $0 \leq i \leq d_n - 1$. The result now follows from (3) and Lemma 2.3. □

It follows that $\Delta(F_1, F_2, \ldots, F_n; t)$ depends only on the coefficients of $F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}$; Macaulay also proved this by an argument based on weights.

Theorem 2.2 *For all $t \geq d$ the extraneous factor $\Delta(F_1, F_2, \ldots, F_n; t)$ is equal (up to sign) to the minor of $M(F_1, F_2, \ldots, F_n; t)$ obtained by deleting all rows and columns that are indexed by any power product that is divisible by exactly one $x_i^{d_i}$ for $1 \leq i \leq n$.*

proof. We use induction on n. The result is trivially true for $n = 1$. For $n > 1$ note that the entries indexed by T_3 are divisible by $x_n^{d_n}$ but not by $x_j^{d_j}$ for $1 \leq j \leq n-1$. With reference to Figure 2, note that a product of minors of the matrices $A_0, A_1, \ldots$ is a minor of $M(F_1, F_2, \ldots, F_n; t)$. Now the second product in Theorem 2.1 is the product of the determinants of the matrices $A_{d_n}, A_{d_{n+1}}, \ldots$ which consist of entries indexed by a power product from $x_n^r S(n, t, j)$ for some $r \geq d_n$ and $0 \leq j \leq n-2$. Thus every such power product is divisible by $x_n^{d_n}$ as well as $x_{j+1}^{d_{j+1}}$. By induction each factor $\Delta\big(F_1^{(n)}, F_2^{(n)}, \ldots, F_{n-1}^{(n)}; t-i\big)$ from the first product in Theorem 2.1 is the minor of A_i that is indexed by power products of degree $t - i$ that are divisible by at least two distinct powers $x_r^{d_r}$, $x_s^{d_s}$ for some $1 \leq r, s \leq n-1$. The induction is completed by noting that the correspondence between indices of A_i viewed as a stand alone matrix and as part of $M(F_1, F_2, \ldots, F_n; t)$ is given by $x^\alpha \leftrightarrow x^\alpha x_n^i$ (bearing in mind that $0 \leq i \leq d_n - 1$). □

§3. Vanishing of the Extraneous Factor. Macaulay's expression for the resultant has a serious drawback in applications; it is possible that upon specialization of the generic coefficients (usually to rationals) the extraneous factor vanishes even when the resulting polynomials do not have a non-trivial root. For example take $F_1 = x_2$, $F_2 = x_3$ and $F_3 = x_1^2$. Canny [**1**] provided a way round this using a method that computes the characteristic polynomial of $M(F_1, F_2, \ldots, F_n; d)$ (recall that $d = 1 + \sum_{i=1}^{n}(d_i - 1)$). This increases the cost somewhat since we now replace a computation involving only numbers with one that involves an unknown; however see Manocha and Canny [**11**]. It is therefore of interest to be able to detect situations when the extraneous factor does vanish more efficiently than just computing it directly.

When deciding whether the extraneous factor polynomial vanishes at a given specialization we do not need repeated factors so let us write $G \sim H$ to mean that the polynomials G, H have the same square free part. Furthermore we simplify notation by writing $\Delta(d_1, d_2, \ldots, d_r, t)$ in place of $\Delta(F_1^{(r+1)}, F_2^{(r+1)}, \ldots, F_r^{(r+1)}, t)$ (and similarly for $D(\cdot)$ as well as $\mathrm{Res}(\cdot)$).

Repeated application of Theorem 2.1 yields the following.

Lemma 3.1 *For all $t \geq d$ we have*

$$\begin{aligned}\Delta(d_1, d_2, \ldots, d_n; t) \sim\, & D(d_1, d_2, \ldots, d_{n-1}; t - d_n) \cdots D(d_1, d_2, \ldots, d_{n-1}, 1)\\ & D(d_1, d_2, \ldots, d_{n-2}; t - d_{n-1}) \cdots D(d_1, d_2, \ldots, d_{n-2}, 1)\\ & \vdots\\ & D(d_1, t - d_2) \cdots D(d_1, 1).\end{aligned}$$

We can carry out further simplifications by using the following result.

Lemma 3.2 *Suppose that $t < d_n$ or $t > d$. Then*

$$D(d_1, d_2, \ldots, d_n; t) = D(d_1, d_2, \ldots, d_{n-1}; t) D(d_1, d_2, \ldots, d_n; t - 1).$$

Moreover for $t < d_n$ we have

$$D(d_1, d_2, \ldots, d_n; t) \sim D(d_1, d_2, \ldots, d_{n-1}; t) D(d_1, d_2, \ldots, d_{n-1}; t - 1) \cdots D(d_1, d_2, \ldots, d_{n-1}; 1),$$

while for $t > d$ we have

$$D(d_1, d_2, \ldots, d_n; t) \sim D(d_1; t_1) D(d_1, d_2; t_2) \cdots D(d_1, d_2, \ldots, d_n; t_n)$$

where $t_i = 1 + \sum_{j=1}^{i}(d_i - 1)$.

proof. If $t < d_n$ then the matrices C_1, C_2, C_3 of Figure 1 are not present. The matrix A accounts for the first determinant of the product and B accounts for the second.

If $t > d$ then every power product in $S(n, t, n-1)$ is divisible by $x_n^{d_n+1}$ so that in the forms $(x^\alpha / x_n^{d_n}) F_n$ for $x^\alpha \in S(n, t, n-1)$ every power product is divisible by x_n. It follows that the matrix C_1 is zero.

The two expansions follow easily. □

The case $t > d$ was observed by Macaulay in §8 of [**8**] where he also gives explicit powers for the expansion. As examples we have

$$\begin{aligned}\Delta(1, 2, 3; 4) &\sim D(1; 1),\\ \Delta(1, 2, 3, 4; 7) &\sim D(1; 1) D(1, 2; 2) D(1, 2, 3; 3).\end{aligned}$$

We could also use (1) in simplifications but in general this introduces resultants which are harder to evaluate than determinants (in the case of the second example we would simply replace $D(1, 2; 2)$ with $R(1, 2)$ but these are equal).

Finally we observe that if $t < d_r$ for some $1 \leq r \leq n - 1$ then $D(d_1, d_2, \ldots, d_n; t)$ is (up to sign) equal to $D(d_1, \ldots, d_{r-1}, d_{r+1}, \ldots, d_n, d_r; t)$ where we have now changed the ordering of the indeterminates from $x_1, x_2, \ldots, x_n$ to $x_1, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n, x_r$. Of course this case does not arise if we order the forms by non-decreasing degree.

§4. Testing the Resultant. In this section we describe another way of testing the resultant (for equality to 0) that avoids the problem of the extraneous factor vanishing. One possible approach might be to change the order of the polynomials since the extraneous factor depends on the order they are used. Unfortunately this is not guaranteed to work as is shown by the simple example at the start of §3. Of course the construction of Macaulay's matrix sets up a correspondence between each form F_i and an indeterminate x_i (in our case). This correspondence is arbitrary and could be varied in an attempt to avoid the vanishing of the extraneous factor; however it would be very inefficient to try out all possibilities.

Consider the forms

$$\begin{aligned} &x^\alpha F_1, \quad |\alpha| = t - d_1, \\ &x^\alpha F_2, \quad |\alpha| = t - d_2, \\ &\qquad \vdots \\ &x^\alpha F_n, \quad |\alpha| = t - d_n. \end{aligned}$$

We define a matrix $L(F_1, F_2, \ldots, F_n; t)$ in a manner similar to §2; there is one row for each form $x^\alpha F_i$ whose entries are the coefficients of F_i disposed as before. If we assume that $t \geq d$ and choose any square sub-matrix of $L(F_1, F_2, \ldots, F_n; t)$ its determinant is either 0 or is a non-zero polynomial divisible by $\operatorname{Res}(F_1, F_2, \ldots, F_n)$ (Macaulay [**9**] defines the resultant to be the gcd of all such determinants with $t = d$; see also §6a of [**8**].) We note here that the the construction of the matrix L makes sense even when the number of forms is different from the number of indeterminates and this will be used in §5.

In this section we will work with $A = k[x_1, x_2, \ldots, x_n]$ and forms $G_1, G_2, \ldots, G_n \in A$ of degree $d_1, d_2, \ldots, d_n$ respectively. We set $I = (G_1, G_2, \ldots, G_n)$ and use A_s, I_s to denote the forms in A and I of degree s (including 0 so that we have a k-vector space).

Let λ be an indeterminate over $\mathbb{Z}$ and define the formal power series

$$H(A/I, \lambda) = \sum_{s=0}^{\infty} \dim_k(A_s/I_s)\lambda^s.$$

Suppose that $G_1, G_2, \ldots, G_r$, where $r \leq n$, is a regular sequence (called 'prime sequence' by Zariski and Samuel [**15**]) and set $J = (G_1, G_2, \ldots, G_r)$ (this has dimension $n - r$). Macaulay [**9**] shows that

$$H(A/J, \lambda) = (1 - \lambda^{d_1})(1 - \lambda^{d_2}) \cdots (1 - \lambda^{d_r})(1 - \lambda)^{-n}.$$

A modern treatment is given by Stanley [**13**] (Macaulay [**10**] states that the result was known before [**9**]). The key point is that for each i the linear map $A_{s-d_i}/(G_1, G_2, \ldots, G_{i-1})_{s-d_i} \to A_s/(G_1, G_2, \ldots, G_{i-1})_s$ given by $G \mapsto GG_i$ is 1-1. Note that for $r = n$ we have

$$\begin{aligned} H(A/J, \lambda) &= (1 + \lambda + \cdots + \lambda^{d_1-1})(1 + \lambda + \cdots + \lambda^{d_2-1}) \cdots (1 + \lambda + \cdots + \lambda^{d_n-1}) \\ &= c_0 + c_1\lambda + \cdots + c_{d-1}\lambda^{d-1}, \end{aligned}$$

where $c_{d-1} = 1$ and in fact the sequence $c_0, c_1, \ldots, c_{d-1}$ is equal to its reversed version (as Macaulay [**9**] observed). Moreover

$$c_0 + c_1 + \cdots + c_{d-1} = d_1 d_2 \cdots d_n, \tag{1}$$

as can be seen by setting $\lambda = 1$ (this is a version of Bézout's Theorem).

LEMMA 4.1 *Suppose that $G_1, G_2, \ldots, G_n$ have no common zero other than the trivial one. Then $\dim_k(A_s/I_s) = 0$ for all $s \geq d$, i.e., I_s has maximum possible dimension over k.*

PROOF. From the observations above, it suffices to show that $G_1, G_2, \ldots, G_n$ is a regular sequence. Let $M = (x_1, x_2, \ldots, x_n)$ then the local ring A_M is Cohen-Macaulay and A_M/IA_M has the same dimension as I which is 0 (since I has just one zero in affine space). The claim now follows from Theorem 2 in Appendix 6 of [**15**]. □

LEMMA 4.2 *$G_1, G_2, \ldots, G_n$ have a non-trivial common zero if and only if $L(G_1, G_2, \ldots, G_n; d)$ has rank strictly less than $\binom{n+d-1}{n-1}$.*

PROOF. If $G_1, G_2, \ldots, G_n$ have no common zero other than the trivial one the preceding lemma shows that $\dim_k I_d = \binom{n+d-1}{n-1}$ which is the same as the rank of $L(G_1, G_2, \ldots, G_n; d)$.

Conversely if the rank of $L(G_1, G_2, \ldots, G_n; d)$ is $\binom{n+d-1}{n-1}$ then it has a non-singular square sub-matrix that contains all the columns. This shows that the power products of degree d can be written as linear combinations of the $x^\alpha F_i$ corresponding to the rows of the square sub-matrix, i.e., all the power products are in I. In particular $x_i^d \in I$ for $1 \leq i \leq n$ and so $G_1, G_2, \ldots, G_n$ have no common zero other than the trivial one. □

An interesting consequence of this lemma is that we can find a square matrix in the coefficients of $G_1, G_2, \ldots, G_n$ whose determinant vanishes if and only if the resultant vanishes provided we enlarge k with new indeterminates. Let the rows of $L(G_1, G_2, \ldots, G_n; d)$ be $\mathbf{l}_1, \mathbf{l}_2, \ldots, \mathbf{l}_r$ and denote the number of columns by c. Introduce new indeterminates v_{ij} for $1 \leq i \leq c$ and $c+1 \leq j \leq r$. Then

$$\begin{pmatrix} \mathbf{l}_1 + v_{1,c+1}\mathbf{l}_{c+1} + \cdots + v_{1,r}\mathbf{l}_r \\ \mathbf{l}_2 + v_{2,c+1}\mathbf{l}_{c+1} + \cdots + v_{2,r}\mathbf{l}_r \\ \vdots \\ \mathbf{l}_c + v_{c,c+1}\mathbf{l}_{c+1} + \cdots + v_{c,r}\mathbf{l}_r \end{pmatrix}$$

has the claimed property. Of course it is too costly to compute the determinant of this matrix. Following Schwartz [**12**] we have the following simple probabilistic algorithm. Let S be a subset of k of size at least c and choose values for the v_{ij} from S uniformly at random. If the determinant of the matrix is non-zero then the resultant is non-zero, otherwise the resultant is zero with probability at least $1 - c/|S|$.

§5. The u-resultant. In this section we deal with n homogeneous forms in $n+1$ indeterminates that have finitely many zeros in projective space. We denote the forms by $H_1, H_2, \ldots, H_n$ and let the indeterminates be $x_0, x_1, \ldots, x_n$ where, as usual, the points at infinity are those with $x_0 = 0$. We denote the degrees of the forms by $d_1, d_2, \ldots, d_n$ and continue to use $d = 1 + \sum_{i=1}^{n}(d_i - 1)$. The u-resultant of the forms is simply $\mathrm{Res}(H_0, H_1, \ldots, H_n)$ in which $H_0 = u_0x_0 + u_1x_1 + \cdots + u_nx_n$ where $u_0, u_1, \ldots, u_n$ are new indeterminates over k. This is discussed by Macaulay [**9**] and van der Waerden [**14**]; see also Cox, Little and O'Shea [**2**]. It can be shown that

$$\mathrm{Res}(H_0, H_1, \ldots, H_n) = \textstyle\prod_p H_0(p)^{m(p)},$$

where the product ranges over all common zeros p of $H_1, H_2, \ldots, H_n$ and $m(p)$ denotes the multiplicity of p. Macaulay showed that $\Delta(H_1, H_2, \ldots, H_n, H_0; d)$ is independent of the u_i so that whenever this is non-zero we can obtain the u-resultant (up to a constant multiple) as $D(H_1, H_2, \ldots, H_n, H_0; d)$. We proceed to describe a method that ensures that the extraneous constant factor does not vanish.

Recall that an order $<$ on power products is said to be *admissible* provided that 1 is the smallest power product and whenever $x^\alpha < x^\beta$ then $x^\alpha x^\gamma < x^\beta x^\gamma$. We will use graded (or degree based) orders, i.e., orders with the extra property that whenever $x^\alpha < x^\beta$ then $|\alpha| \leq |\beta|$. From now on we assume that such an order has been fixed. For a non-zero polynomial f we use $\mathrm{lpp}(f)$ to denote the largest power product that occurs in f with a non-zero coefficient. We extend this notation to sets in the obvious way, i.e., $\mathrm{lpp}(S) = \{\mathrm{lpp}(f) \mid f \in S - \{0\}\}$.

An algorithm for computing the u-resultant of $H_1, H_2, \ldots, H_n$ is as follows.

1. Construct the matrix $L(H_1, H_2, \ldots, H_n; d)$. Use Gaussian elimination and keep the non-zero rows, these correspond to forms $B_1, \ldots, B_r$.

2. Construct $L(H_1, H_2, \ldots, H_n; d-1)$ with the columns indexed by power products sorted in decreasing order. Use Gaussian elimination on this and let P be the set of power products corresponding to the first non-zero entry in each non-zero row. Let $x^{\alpha_1}, ..., x^{\alpha_s}$ be all the power products of degree $d-1$ that are not in P. (If $d = 1$ then we just return 1 as the sequence of power products.)

3. Let $U(H_0, H_1, \ldots, H_n)$ be the coefficient matrix of the forms $B_1, \ldots, B_r, x^{\alpha_1}H_0, ..., x^{\alpha_s}H_0$.

LEMMA 5.1 *If $H_1, \ldots, H_n$ have finitely many common zeros then the matrix $U(H_0, H_1, \ldots, H_n)$ is square and*

$$\det U(H_0, H_1, \ldots, H_n) = a \operatorname{Res}(H_0, H_1, \ldots, H_n)$$

for some non-zero $a \in k$. Otherwise the matrix is either not square or the determinant is zero.

PROOF. Suppose that $H_1, \ldots, H_n$ have finitely many common zeros. Let K be the algebraic closure of $k(u_0, u_1, \ldots, u_n)$. Since $\operatorname{Res}(H_0, H_1, \ldots, H_n) \neq 0$ it follows that $H_0, H_1, \ldots, H_n$ have no common zero in $\mathbb{P}^n(K)$ other than the trivial one. It follows from the proof of Lemma 4.1 that $H_1, \ldots, H_n, H_0$ is a regular sequence in $K[x_0, x_1, \ldots, x_n]$.

Clearly the forms $B_1, \ldots, B_r$ constitute a K-vector basis for $(H_1, H_2, \ldots, H_n)_d$. We claim that $x^{\alpha_1}H_0, ..., x^{\alpha_s}H_0$ extend this basis to one for $(H_0, H_1, \ldots, H_n)_d$. For $x^{\alpha_1}, ..., x^{\alpha_s}$ are a basis for $K[x_0, x_1, \ldots, x_n]_{d-1}/(H_1, H_2, \ldots, H_n)_{d-1}$ since the set P constructed in step 2 of the algorithm is precisely $\operatorname{lpp}(H_1, H_2, \ldots, H_n)_{d-1}$. Since $H_1, \ldots, H_n, H_0$ is a regular sequence the claim follows. Lemma 4.1 now shows that $U(H_0, H_1, \ldots, H_n)$ is square and it follows from the remarks at the beginning of §4 that $\det U(H_0, H_1, \ldots, H_n) = a \operatorname{Res}(H_0, H_1, \ldots, H_n)$ for some $a \in K$. In fact $a \in k[u_0, u_1, \ldots, u_n]$ since the entries of the matrix $U(H_0, H_1, \ldots, H_n)$ are the coefficients of the forms and they are from k apart from those of H_0 which are $u_0, u_1, \ldots, u_n$. We show that the degree of $\det U(H_0, H_1, \ldots, H_n)$ as a polynomial in $u_0, u_1, \ldots, u_n$ is at most $d_1 d_2 \cdots d_n$. The lemma will then follow (for the case of finitely many common zeros) since the degree of $\operatorname{Res}(H_0, H_1, \ldots, H_n)$ is exactly $d_1 d_2 \cdots d_n$ (by the third property given in §1). Now the only rows of $U(H_0, H_1, \ldots, H_n)$ that involve $u_0, u_1, \ldots, u_n$ are those corresponding to $x^{\alpha_1}H_0, ..., x^{\alpha_s}H_0$ and from §4 we know that s is the coefficient of λ^{d-1} in $\prod_{i=1}^n (1-\lambda)^{d_i}/(1-\lambda)^{n+1}$, i.e., $s = c_0 + c_1 + \cdots + c_{d-1}$. It follows from (1) that $s = d_1 d_2 \cdots d_n$ as required.

Suppose now that $H_1, \ldots, H_n$ have infinitely many common zeros and the matrix is square. The u-resultant is then identically zero since for each $p \in \mathbf{V}(H_1, \ldots, H_n)$ we have that $l_p(u)$ divides $\operatorname{Res}(H_0, H_1, \ldots, H_n)$. The lemma follows since $\det U(H_0, H_1, \ldots, H_n)$ is a multiple of $\operatorname{Res}(H_0, H_1, \ldots, H_n)$. □

The proof of the lemma justifies the insertion of the following test between the first and second steps of the algorithm:

1.5 If $r \neq \binom{n+d-1}{n-1} - d_1 d_2 \cdots d_n$ then halt (the forms have infinitely many common zeros).

We note also that in step 2 we could use Gröbner bases to compute P and hence $x^{\alpha_1}, ..., x^{\alpha_s}$. However the runtime for such a computation is difficult to predict. In practical terms it would pay to run the two approaches in parallel.

We now consider the special case when $H_1, H_2, \ldots, H_n$ have no common zeros at infinity. We will use $h_1, h_2, \ldots, h_n$ to denote the dehomogenizations of the forms, i.e., h_i is H_i with $x_0 \mapsto 1$. We will also use $H_1^*, H_2^*, \ldots, H_n^*$ to denote the forms obtained from $H_1, H_2, \ldots, H_n$ by the substitution $x_0 \mapsto 0$.

Note that x_0 does not divide any of the forms $H_1, H_2, \ldots, H_n$. For if, w.l.o.g, H_n is divisible by x_0 then $H_n^* = 0$ and since the forms $H_1^*, H_2^*, \ldots, H_{n-1}^* \in k[x_1, x_2, \ldots, x_n]$ must have a non-trivial common zero it follows that $H_1, H_2, \ldots, H_n$ have a non-trivial common zero at infinity, contrary to assumption. We set

$$I = (h_1, h_2, \ldots, h_n),$$
$$I^* = (H_1^*, H_2^*, \ldots, H_n^*),$$

ideals of $k[x_1, x_2, \ldots, x_n]$ and

$$J = (H_1, H_2, \ldots, H_n),$$
$$K = (x_0, H_1, H_2, \ldots, H_n)$$

ideals of $k[x_0, x_1, \ldots, x_n]$. From now on we assume that the graded order on power products is such that if $x_0^r x^\alpha < x_0^s x^\beta$ then $x^\alpha < x^\beta$ where $r + |\alpha| = s + |\beta|$ and x_0 does not divide either of x^α or x^β. An example of such an order is obtained by sorting lexicographically within each degree with x_0 as the smallest indeterminate.

LEMMA 5.2 $\mathrm{lpp}(I) = \mathrm{lpp}(J) \cap k[x_1, x_2, \ldots, x_n] = \mathrm{lpp}(K) \cap k[x_1, x_2, \ldots, x_n]$.

PROOF. Suppose that $x^\alpha \in \mathrm{lpp}(I)$ so that there are $g, g_1, g_2, \ldots, g_n \in k[x_1, x_2, \ldots, x_n]$ such that

$$x^\alpha + g = g_1 h_1 + g_2 h_2 + \cdots + g_n h_n$$

where all the power products of g are less than x^α. Homogenizing we have

$$x_0^e x^\alpha + x_0^d G = x_0^{e_1} G_1 H_1 + x_0^{e_2} G_2 H_2 + \cdots + x_0^{e_n} G_n H_n,$$

where $e \leq d$ since we are using a graded order. If $e = 0$ then it follows immediately that $x^\alpha \in \mathrm{lpp}(J) \cap k[x_1, x_2, \ldots, x_n]$, so assume that $e > 0$. It follows that $x_0^e(x^\alpha + x_0^{d-e} G)$ is zero in $k[x_0, x_1, \ldots, x_n]/(H_1, H_2, \ldots, H_n)$. However the proof of Lemma 4.1 shows that $x_0, H_1, H_2, \ldots, H_n$ is a regular sequence and hence so is $x_0^e, H_1, H_2, \ldots, H_n$. It now follows that $x^\alpha + x_0^{d-e} G \in (H_1, H_2, \ldots, H_n)$ and so $x^\alpha \in \mathrm{lpp}(J) \cap k[x_1, x_2, \ldots, x_n]$.

If $x^\alpha \in \mathrm{lpp}(K) \cap k[x_1, x_2, \ldots, x_n]$ there are homogeneous polynomials $G, G_0, G_1, \ldots, G_n$ such that

$$x^\alpha + G = G_0 x_0 + G_1 H_1 + \cdots + G_n H_n$$

where each power product of G is less than x^α. It follows that $\deg G_0 = |\alpha| - 1$ and so each power product of G_0 is also less than x^α. Dehomogenizing we obtain

$$x^\alpha + g - g_0 = g_1 h_1 + g_2 h_2 + \cdots + g_n h_n$$

and so $x^\alpha \in \mathrm{lpp}(I)$.

The lemma follows since $J \subseteq K$. □

LEMMA 5.3 $\mathrm{lpp}(I^*) = \mathrm{lpp}(I)$.

PROOF. By Lemma 5.2 it suffices to prove that $\mathrm{lpp}(I^*) = \mathrm{lpp}(K) \cap k[x_1, x_2, \ldots, x_n]$. Suppose that $x^\alpha \in \mathrm{lpp}(K) \cap k[x_1, x_2, \ldots, x_n]$ so that there are polynomials $G, G_0, G_1, \ldots, G_n$ such that

$$x^\alpha + G = G_0 x_0 + G_1 H_1 + \cdots + G_n H_n$$

where each power product of G is less than x^α. Substituting $x_0 \mapsto 0$ shows that $x^\alpha \in \mathrm{lpp}(I^*)$.

The converse follows from the fact that $K = (x_0, H_1^*, H_2^*, \ldots, H_n^*)$. □

LEMMA 5.4 *Let $x^{\beta_1}, \ldots, x^{\beta_s}$ be all the power products of degree less than or equal to a degree e that are not in $\mathrm{lpp}(I^*)$ and $x^{\alpha_1}, \ldots, x^{\alpha_s}$ their homogenizations to degree e (i.e., $x^{\alpha_i} = x_0^{e-|\beta_i|} x^{\beta_i}$. Then $x^{\alpha_1}, \ldots, x^{\alpha_s}$ is a k-basis for $k[x_0, x_1, \ldots, x_n]_e / J_e$.*

PROOF. Suppose that $x_0^r x^\alpha \in \mathrm{lpp}(J)$ where $r + |\alpha| = e$ and $x^\alpha \in k[x_1, x_2, \ldots, x_n]$. The substitution $x_0 \mapsto 1$ shows that x^α is in $\mathrm{lpp}(I)$ and hence in $\mathrm{lpp}(I^*)$ by Lemma 5.3. Conversely if $x^\alpha \in \mathrm{lpp}(I^*)$ where $|\alpha| \leq e$ then $x_0^{e-|\alpha|} x^\alpha \in \mathrm{lpp}(J)_e$.

It follows that the power products of degree e that do not belong to $\mathrm{lpp}(J)_e$ are as described in the lemma and hence form a k-basis for $k[x_0, x_1, \ldots, x_n]_e / J_e$. □

The final lemma justifies the following modified version of step 2 of the algorithm for computing the u-resultant of $H_1, H_2, \ldots, H_n$:

2^* For $\min(d_1, d_2, \ldots, d_n) \leq i \leq d-1$, construct $L(H_1^*, H_2^*, \ldots, H_n^*; i)$ with the columns indexed by power products sorted in decreasing order. Use Gaussian elimination and let P_i be the set of power products corresponding to the first non-zero entry in each non-zero row. Let $x^{\beta_1}, \ldots, x^{\beta_s}$ be all the power products of degree at most $d-1$ that are not in $\cup_i P_i$. Let $x^{\alpha_1}, \ldots, x^{\alpha_s}$ be the homogenizations of the preceding power products to degree $d-1$. (If $d = 1$ then we just return 1 as the sequence of power products.)

In practice this might not produce a gain in computational cost as compared to the original step 2, however the Gröbner bases approach would usually benefit from replacing $H_1, H_2, \ldots, H_n$ with $H_1^*, H_2^*, \ldots, H_n^*$.

References

1. J. Canny, 'Generalized characteristic polynomials', *J. Symbolic Comput.* **9** (1990) 241–250.

2. D. Cox, J. Little, D. O'Shea, *Using Algebraic Geometry*, Springer, New York, (1998).

3. I. Gelfand, M. Kapranov and A. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston (1994).

4. W. Gröbner, *Moderne Algerbraische Geometrie*, Springer, Vienna (1949).

5. J-P. Jouanolou, 'La formalisme du résultant', Advances in Math. **90** (1991) 117–263.

6. J-P. Jouanolou, 'Formes d'inertie et résultant: un formulaire', Adv. in Math., **126** (1997) 119–250.

7. D. Lazard, 'Resolution des systèmes d'équations algébriques', *Theoretical Computer Science*, **15** (1981) 77-110.

8. F.S. Macaulay, 'On some formulas in elimination', *Proc. LMS* **3** (1902) 3–27.

9. F.S. Macaulay, *The Algebraic Theory of Modular Systems*, Cambridge U. Press (1916). Reprinted 1994.

10. F.S. Macaulay, 'Some properties of enumeration in the theory of modular systems', *Proc. LMS* **26** (1927) 531–555.

11. D. Manocha and J.F. Canny, 'MultiPolynomial Resultant Algorithms', *J. Symbolic Comput.* **15** (1993) 99–122.

12. J.T. Schwartz, 'Probabilistic algorithms for verification of polynomial identities', *J. ACM* **27** (1980), 701–717.

13. R.P. Stanley, 'Hilbert Functions of Graded Algebras', *Advances in Mathematics* **28** (1978) 57–83.

14. B.L. van der Waerden, *Modern Algebra, Vol. II*, F. Ungar Publishing Co., New York, (1950).

15. O. Zariski and P. Samuel, *Commutative Algebra*, Vol. II, Springer, New York, (1960).