On the resultant of degree-deficient polynomials

Alexander Lifshitz and Yuval Bistritz Department of Electrical Engineering Tel Aviv University Tel Aviv 69978 ISRAEL

Abstract: The resultant is an algebraic expression, computable in a finite number of arithmetic operations from the coefficients of two univariate polynomials, that vanishes if, and only if, the two polynomials have common zeros. The paper considers formal resultant for degree-deficient polynomials (polynomials whose actual degree is lower than their assumed degree). Some key properties of the resultant are extended to formal resultants including its expression by the finite zeros of the polynomials.

Key-Words: Resultant, Sylvester matrix, Bezout matrix, GCD of polynomials.

1 Introduction

The *resultant* is an algebraic concept that arises in seeking conditions for two univariate polynomials to have common zeros. It plays a role in determining co-primeness of also more than two (univariate, multivariate, matrix) polynomials, determining zero location of polynomials in the complex plane for testing and controlling stability of linear systems, and in more applications. The basic problem can be stated as follows. Let $\mathbb F$ be an arbitrary field, and let

$$a(z) = a_m z^m + a_{m-1} z^{m-1} + \dots + a_0$$

$$b(z) = b_n z^n + b_{n-1} z^{n-1} + \dots + b_0$$
(1)

be two general polynomials of degree m and n in $\mathbb{F}[z]$. Find necessary and sufficient conditions for a(z) and b(z) to have common zeros. The answer to this problem is that there exists a multivariate polynomial denoted by $\mathcal{R}(a,b)$ whose variables are the coefficients $a_i,\ i=0,\ldots,m$ and $b_j,\ j=0,\ldots,n$, such that a(z) and b(z) have a common zero if, and only if, $\mathcal{R}(a,b)=0$. This function is called the resultant of a(z) and b(z).

There are several ways to present the resultant of two polynomials. The most commonly used presentations relate $\mathcal{R}(a,b)$ to the determinant of certain resultant matrices formed from the polynomial coefficients. These expressions were introduced by Euler (1748) and Bézout (1764) and further considered in 19th century works by Jacobi, Sylvester and Cauchy (for an excellent introduction and background for the topic with also a careful report of historical perspective see [1]). The term resultant apparently comes from Bézout's celebrated paper on elimination theory [2]. He considered there ways to construct the $\mathcal{R}(a,b)$

as the determinant of matrices of size n+m following Euler as well as the determinant of an abridged matrix of size max(m,n) that is called today the Bézoutian matrix (as it was called already by Sylvester [3]). Euler constructed for a pair of polynomials a(z) and b(z) (1) a matrix of size (n+m),

$$\operatorname{Syl}(a,b) = \begin{pmatrix} a_{m}a_{m-1} \dots a_{0} \\ a_{m}a_{m-1} \dots a_{0} \\ \dots \\ a_{m}a_{m-1} \dots a_{0} \\ b_{n}b_{n-1} \dots b_{0} \\ b_{n}b_{n-1} \dots b_{0} \\ \dots \\ b_{n}b_{n-1} \dots b_{0} \end{pmatrix} \begin{cases} n \text{ rows} \\ m \text{ rows} \end{cases}$$

(all blank spaces must be filled with zeros), now called the Sylvester matrix . It can be shown that for two full degree non-zero $a(z),b(z)\in\mathbb{F}[z]$ det $\mathrm{Syl}(a,b)=0$ if, and only if, a(z) and b(z) have common zeros, see e.g. [5] or [1]. Thus det $\mathrm{Syl}(a,b)$ qualifies as an expression for $\mathcal{R}(a,b)$. Some mathematical texts take $\det\mathrm{Syl}(a,b)$ as the definition for the resultant of a(z) and b(z)

$$\mathcal{R}(a,b) = \det \text{Syl}(a,b) \tag{3}$$

There exists for a(z) and b(z) in (1) a field \mathbb{K} ($\mathbb{K} \supseteq \mathbb{F}$) such that a(z) and b(z) can be factored into linear terms

over \mathbb{K}

$$a(z) = a_m \prod_{i=1}^{m} (z - \alpha_i)$$

$$b(z) = b_n \prod_{i=1}^{n} (z - \beta_i)$$
(4)

$$b(z) = b_n \prod_{i=1}^{n} (z - \beta_i)$$

$$(5)$$

where $\{\alpha_i\}_{i=1}^m \in \mathbb{K}$ and $\{\beta_i\}_{i=1}^n \in \mathbb{K}$ are zeros of a(z)and b(z), respectively. This is known as the fundamental theorem of algebra. An alternative presentation for The resultant is its expression by the zeros of a(z) or b(z).

Theorem 1. Let a(z) and b(z) be two polynomials (1) of full degree $(a_m \neq 0 \text{ and } b_n \neq 0)$ in $\mathbb{F}[z]$ with factorization into their zeros as in (4) and (5). Then the resultant $\mathcal{R}(a,b)$, defined by (2)-(3), may be expressed as a function of the zeros of a(z) or b(z)

$$\mathcal{R}(a,b) = a_m^n \prod_{i=1}^m b(\alpha_i)$$

$$\mathcal{R}(a,b) = (-1)^{mn} b_n^m \prod_{i=1}^n a(\beta_i)$$
(6)

$$\mathcal{R}(a,b) = (-1)^{mn} b_n^m \prod_{i=1}^n a(\beta_i)$$
 (7)

For a proof for the above theorem see e.g. [4] or [5]. Combining (6) with (5), gives a third expression for the resultant that uses the zeros of both polynomials.

Addendum. The resultant $\mathcal{R}(a,b)$ of a(z) and b(z) with zeros (4) and (5) can be expressed by

$$\mathcal{R}(a,b) = a_m^n b_n^m \prod_{i=1}^m \prod_{j=1}^n (\alpha_i - \beta_j)$$
 (8)

Some texts use (8) as definition for the resultant $\mathcal{R}(a,b)$ for a(z) and b(z).

The expression of the resultant by the determinant of the Sylvester matrix provides a useful algebraic tool to determine whether two polynomials have common zeros because it is computable from the coefficients of the polynomials in a finite number of arithmetic operations. The expression of the resultant by the zeros of the polynomials does not seem to serve a similarly useful purpose. The expressions (6)-(8) make transparent the capacity of the resultant to detect common zeros. However, if the zeros of both a(z) and b(z) that participate in (8) are known, the common zeros can be determined by inspection. Similarly, when the zeros of one of the two polynomials are known, the expressions(6) and (7) do a straightforward detection of common zeros - evaluation of the other polynomial at the known zeros. On the other hand, these expressions make most transparent the ability of the resultant to detect common zeros. They become instrumental at various stages in developing and proving properties of algorithms associated with the resultant. Furthermore, if (2)-(3) is taken as the definition for $\mathcal{R}(a,b)$ (as many textbooks do) then showing

that it is equal to one of the expressions (6) (7) or (8) proves that (2)-(3) indeed functions as detector for common zeros for the two polynomials. We shall apply later the latter argument to establish the resultant for also degree-deficient polynomials.

The resultant can also be expressed as the determinant of the above mentioned Bezoutian matrix. The Bezoutian and the Sylvester matrix have each its own advantage for various algebraic tasks. The advantage of the Bezoutian matrix for some algebraic tasks stems from its having the minimal matrix size for posing the relevant algebraic conditions. The size reduction from the double sized Sylvester matrix is achieved by elimination of the structural zeros in the Sylvester matrix. But, as a consequence, the matrix entries of the Bezoutian are no longer the coefficients of the two polynomials. The fact that the coefficients of the polynomials appear in the Sylvester matrix directly renders it more useful for some other tasks including the derivations in this paper. Thus, we focus on the Sylvester matrix. Obviously, properties established for the determinant of the Sylvester matrix hold also for the determinant of a (properly set) corresponding Bezoutian matrix.

Resultants and Bezoutian play an important role in determining whether the zeros of a polynomial are in certain regions of the complex plane. For example stability of continuous-time and discrete-time linear systems require a polynomial to have zeros in the left-half of the complex plane and in a unit-circle centered at the origin, respectively. In such problems, the tested polynomial is paired with a second polynomial whose zeros are the reflection of the zeros of the polynomial with respect to the boundary of the desirable domain [6]. The zero location problems extends the stability conditions into asking the distribution of the polynomial zeros with respect to the respective distinguished boundary. In fact the main result in this paper stems from a need that arose in our study on the extension of the zero location with respect to the unit-circle method in [7] into an integer-arithmetic algorithm (the paper under preparation will employ the results established here). It is apparent from [7] that the method there constructs for the tested polynomial a sequence of polynomials of decreasing degrees that at times may have vanishing leading coefficients. The situation needed resultants for degree deficient polynomials and mobility between its expression by $\det \operatorname{Syl}(a,b)$ and by the expressions in Theorem 1. It is noteworthy that the problem in this paper is presented and solved regardless of our initial motivation and without restriction to the specific form of polynomial in our intended immediate application.

The paper considers $\mathcal{R}(a,b)$ for polynomials whose degrees are formally m and n given as in (1) but their highest non-vanishing coefficient is lower then their formal leading coefficient. Note that the expressions for the resultant in Theorem 1 is restricted to full degree polynomials. It needs $a_m \neq 0$ and the *n* zeros of b(z) or $b_n \neq 0$ and the *m*

zeros of a(z). In difference, the expression $\det Syl(a,b)$ seems to be more tolerant to degree-deficient polynomials. It will become apparent from our main result that $\det \operatorname{Syl}(a,b)$ still functions well as a common zeros detector as long as only one of the polynomials has zero leading coefficient. The complementary situation of $a_m = b_n = 0$ obviously implies $\det \text{Syl}(a, b) = 0$. It is possible to regard this situation as detection of common zeros at *infinity*, and then claim that " $\det \text{Syl}(a, b) = 0$ if, and only if, a(z) and b(z) have common zeros" holds equally for both full and degree deficient polynomials.

In this paper we extend Theorem 1 to degree-deficient polynomials. Namely, we obtain exact relations between $\det \operatorname{Syl}(a,b)$ and the finite zeros of a(z) or b(z) for any difference between the formal and the actual degree of the two polynomials. If (2)-(3) is taken as a formal definition for the resultant for also degree deficient polynomials then our derivation provides in addition a constructive proof that $\det \operatorname{Syl}(a,b)$ detects common zeros for also degree deficient polynomials.

2 **Degree-deficient polynomials**

Let \mathbb{F} be an arbitrary field, and let

$$a(z) = a_m z^m + a_{m-1} z^{m-1} + \dots + a_0$$

be a polynomial in $\mathbb{F}[z]$. We shall not exclude the possibility that $a_m = 0$, that is, that the degree of a(z) is actually lower than m. If the polynomial a(z) is written in the above form beginning with a possibly vanishing term $a_m z^m$, then m is called the formal degree of a(z)and a_m is the formal leading coefficient. If the formal leading coefficient is different from zero, then the polynomial is said to be of full degree. If, on the other hand, $a_m = a_{m-1} = \cdots = a_{m-\lambda_a+1} = 0 \text{ and } a_{m-\lambda_a} \neq 0,$ then a(z) is said to be degree-deficient and λ_a is the degree deficiency of a(z). Clearly, $\lambda_a = \operatorname{fdeg} a - \operatorname{deg} a$, where fdeg a is the formal degree and deg a is the nominal degree of a(z).

Theorem 1 provides relationship between the resultant and zeros of a(z) and b(z) with restriction to full degree polynomials. Assume that the resultant is defined by Eqs. (2)-(3) for also degree-deficient a(z) and b(z), in the following we state and prove extensions expressions to (6), (7) and (8) for polynomials with any difference between their formal and actual degrees.

Theorem 2. Let a(z) and b(z) be two polynomials in $\mathbb{F}[z]$ with degree deficiency of $\lambda_a \geq 0$ and $\lambda_b \geq 0$ respectively, and let \mathbb{K} be a field ($\mathbb{K} \supseteq \mathbb{F}$) such that a(z) and b(z) can be factored into linear terms over K

$$a(z) = a_{m-\lambda_a} \prod_{i=1}^{m-\lambda_a} (z - \alpha_i)$$

$$b(z) = b_{n-\lambda_b} \prod_{i=1}^{n-\lambda_b} (z - \beta_i)$$
(9)

$$b(z) = b_{n-\lambda_b} \prod_{i=1}^{n-\lambda_b} (z - \beta_i)$$
 (10)

where $\{\alpha_i\}_{i=1}^{m-\lambda_a} \in \mathbb{K}$ and $\{\beta_i\}_{i=1}^{n-\lambda_b} \in \mathbb{K}$ are zeros of a(z)and b(z) respectively.

Then the resultant defined by Eqs. (2)-(3) may be expressed as a function of zeros of a(z) and b(z) as

$$\mathcal{R}(a,b) = (-1)^{n\lambda_a} a_m^{\lambda_b} b_n^{\lambda_a} a_{m-\lambda_b}^{n-\lambda_b} \prod_{i=1}^{m-\lambda_a} b(\alpha_i)$$
(11)

$$\mathcal{R}(a,b) = (-1)^{n\lambda_a} a_m^{\lambda_b} b_n^{\lambda_a} a_{m-\lambda_a}^{n-\lambda_b} \prod_{i=1}^{m-\lambda_a} b(\alpha_i)$$

$$\mathcal{R}(a,b) = (-1)^{n\lambda_a + (m-\lambda_a)(n-\lambda_b)} a_m^{\lambda_b} b_n^{\lambda_a} b_{n-\lambda_a}^{m-\lambda_a} \prod_{i=1}^{n-\lambda_b} a(\beta_i)$$
(12)

Proof. Denote by $\hat{a}(z)$ and $\hat{b}(z)$ the polynomials a(z) and b(z) with respect to their nominal degrees. The resultant $\mathcal{R}\left(\hat{a},\hat{b}\right)$ is equal to the determinant of the corresponding Sylvester matrix of size $(m - \lambda_a + n - \lambda_b)$

$$\operatorname{Syl}(\hat{a}, \hat{b}) = \begin{pmatrix} a_{m-\lambda_{a}} a_{m-\lambda_{a}-1} \dots a_{0} \\ a_{m-\lambda_{a}} a_{m-\lambda_{a}-1} \dots a_{0} \\ \dots \\ a_{m-\lambda_{a}} a_{m-\lambda_{a}-1} \dots a_{0} \\ b_{n-\lambda_{b}} b_{n-\lambda_{b}-1} \dots b_{0} \\ b_{n-\lambda_{b}} b_{n-\lambda_{b}-1} \dots b_{0} \\ \dots \\ b_{n-\lambda_{b}} b_{n-\lambda_{b}-1} \dots b_{0} \end{pmatrix} m - \lambda_{a} \text{ rows}$$

$$(13)$$

By applying Theorem 1 to $\hat{a}(z)$ and $\hat{b}(z)$ we have

$$\mathcal{R}\left(\hat{a}, \hat{b}\right) = a_{m-\lambda_{a}}^{n-\lambda_{b}} \prod_{i=1}^{m-\lambda_{a}} b(\alpha_{i})$$

$$\mathcal{R}\left(\hat{a}, \hat{b}\right) = (-1)^{(m-\lambda_{a})(n-\lambda_{b})} b_{n-\lambda_{b}}^{m-\lambda_{a}} \prod_{i=1}^{n-\lambda_{b}} a(\beta_{i})$$

Thus, by comparing these equations to Eqs. (11)-(12), we must show that

$$\det \operatorname{Syl}(a,b) = (-1)^{n\lambda_a} a_m^{\lambda_b} b_n^{\lambda_a} \det \operatorname{Syl}(\hat{a},\hat{b}) \tag{14}$$

For simplicity of the following, we specify submatrices of Syl(a, b) by participating columns and rows as follows.

$$Syl(a, b) \equiv S_{a, b}(1 : m + n, 1 : m + n)$$

We consider four cases:

Case 1:
$$\lambda_a = 0, \lambda_b > 0$$

In this case we must show that $\det Syl(a, b) =$ $a_m^{\lambda_b} \det \operatorname{Syl}(\hat{a}, \hat{b}).$ Since each submatrix $S_{a,b}(k:m+n,k:m+n)$ for $1 \le k \le \lambda_b$ has only one non zero element in the first column $(S_{a,b}(k,k) = a_m)$, then by successive expansion of the determinant $\det \operatorname{Syl}(a, b)$ along the first column of each submatrix we obtain

$$\det S_{a,b}(1:m+n,1:m+n)$$
= $a_m \det S_{a,b}(2:m+n,2:m+n)$
...
= $a_m^{\lambda_b} \det S_{a,b}(\lambda_b+1:m+n,\lambda_b+1:m+n)$
= $a_m^{\lambda_b} \det \text{Syl}(\hat{a},\hat{b})$

Case 2:
$$\lambda_a > 0, \lambda_b = 0$$

In this case we must show that $\det Syl(a, b)$ $(-1)^{n\lambda_a}b_n^{\lambda_a}\det \mathrm{Syl}(\hat{a},\hat{b}).$ Since each submatrix $S_{a,b}([1:n, n+k:m+n], k:m+n) \text{ for } 1 \leq k \leq \lambda_b$ has only one non zero element in the first column $(S_{a,b}(n+k,k)=b_n)$, then by successive expansion of the determinant $\det Syl(a, b)$ along the first column of each submatrix we obtain

$$\det S_{a,b}(1:m+n,1:m+n)$$

$$= (-1)^{n+2}b_n \det S_{a,b}([1:n, n+2:m+n], 2:m+n)$$
...
$$= [(-1)^{n+2}b_n]^{\lambda_a} \det S_{a,b}([1:n, n+1+\lambda_a:m+n], \lambda_a+1:m+n)$$

$$= (-1)^{n\lambda_a}b_n^{\lambda_a} \det \text{Syl}(\hat{a}, \hat{b})$$

Case 3:
$$\lambda_a > 0, \lambda_b > 0$$

Since $a_m = b_n = 0$, the resultant must be zero.

$$\det \operatorname{Syl}(a,b) = (-1)^{n\lambda_a} 0^{\lambda_b} 0^{\lambda_a} \det \operatorname{Syl}(\hat{a},\hat{b}) = 0$$

Case 4:
$$\lambda_a = 0, \lambda_b = 0$$

Since $\hat{a}(z) = a(z)$ and $\hat{b}(z) = b(z)$, we must have $Syl(a, b) = Syl(\hat{a}, \hat{b})$

$$\det \operatorname{Syl}(a,b) = (-1)^{n\lambda_a} a_m^0 b_n^0 \det \operatorname{Syl}(\hat{a},\hat{b}) = \det \operatorname{Syl}(\hat{a},\hat{b})$$

The next property extends to formal resultants the property stated as addendum to Theorem 1.

Property 1. Let $a(z), b(z) \in \mathbb{F}[z]$ with fdeg a = m, fdeg b = n, then

$$\mathcal{R}(a,b) = (-1)^{n\lambda_a} a_m^{\lambda_b} b_n^{\lambda_a} a_{m-\lambda_a}^{n-\lambda_b} b_{n-\lambda_b}^{m-\lambda_a} \prod_{i=1}^{m-\lambda_a} \prod_{j=1}^{n-\lambda_b} (\alpha_i - \beta_j)$$

$$\tag{15}$$

Proof. By evaluating Eq. (10) at $z = \alpha_i$ we obtain

$$b(\alpha_i) = b_{n-\lambda_b} \prod_{i=1}^{n-\lambda_b} (\alpha_i - \beta_j)$$

Substituting $b(\alpha_i)$ into (11) gives

$$\mathcal{R}(a,b) = (-1)^{n\lambda_a} a_m^{\lambda_b} b_n^{\lambda_a} a_{m-\lambda_a}^{n-\lambda_b} \prod_{i=1}^{m-\lambda_a} b_{n-\lambda_b} \prod_{j=1}^{n-\lambda_b} (\alpha_i - \beta_j)$$
$$= (-1)^{n\lambda_a} a_m^{\lambda_b} b_n^{\lambda_a} a_{m-\lambda_a}^{n-\lambda_b} b_{n-\lambda_b}^{m-\lambda_a} \prod_{i=1}^{m-\lambda_a} \prod_{j=1}^{n-\lambda_b} (\alpha_i - \beta_j)$$

The expressions (9), (10) and (15) are the extension to degree-deficient polynomials of the expressions (6) (7) and (8), respectively. Also note that regarding (2)-(3) as the definition of the resultant for also degree-deficient polynomials, the next corollary has been proved.

Corollary. Let a(z) and b(z) be two polynomials in $\mathbb{F}[z]$ with degree deficiency of $\lambda_a \geq 0$ and $\lambda_b \geq 0$ respectively, then $\mathcal{R}(b, a) = 0$ if, and only if, a(z) and b(z) have at least one common finite zero or are both degree-deficient.

Here are two other useful properties of the formal resultant. They look like known corresponding properties of full degree polynomials resultant.

Property 2. Let $a(z), b(z) \in \mathbb{F}[z]$ with fdeg a = m, f deg b = n, then

$$\mathcal{R}(b,a) = (-1)^{mn} \mathcal{R}(a,b)$$

Proof. Evaluating (15) with a and b reversed gives

$$\mathcal{R}\left(b,a\right) = (-1)^{m\lambda_b} a_m^{\lambda_b} b_n^{\lambda_a} a_{m-\lambda_a}^{n-\lambda_b} b_{n-\lambda_b}^{m-\lambda_a} \prod_{i=1}^{m-\lambda_a} \left(\beta_j - \alpha_i\right)$$

$$= (-1)^{m\lambda_b + (n-\lambda_b)(m-\lambda_a)} a_m^{\lambda_b} b_n^{\lambda_a} a_{m-\lambda_a}^{n-\lambda_b} b_{n-\lambda_b}^{m-\lambda_a} \prod_{i=1}^{m-\lambda_a} \prod_{j=1}^{n-\lambda_b} (\alpha_i - \beta_j)$$

$$= (-1)^{n\lambda_a + mn} a_m^{\lambda_b} b_n^{\lambda_a} a_{m-\lambda_a}^{n-\lambda_b} b_{n-\lambda_b}^{m-\lambda_a} \prod_{i=1}^{m-\lambda_a} \prod_{j=1}^{n-\lambda_b} (\alpha_i - \beta_j)$$

$$= (-1)^{mn} \mathcal{R} (a, b)$$

Property 3. Let $K_1, K_2 \in \mathbb{F}$ and $a(z), b(z) \in \mathbb{F}[z]$ with f deg a = m, f deg b = n, then

$$\mathcal{R}\left(K_{1}a,K_{2}b\right)=K_{1}^{n}K_{2}^{m}\mathcal{R}\left(a,b\right)$$

Proof. Follows immediately by direct substitution into (2) and taking K_1 and K_2 out of determinant or by straightforward evaluation of (11).

3 Conclusion

The paper considered the determinant of the Sylvester matrix (the resultant) for degree-deficient polynomials. Exact relations between this definition for the formal resultant and the finite zeros of one of the polynomials or both were derived. These relations also led to the conclusion that the formal resultant vanishes if, and only if, the two polynomials have common zeros or are both degree-deficient.

References:

[1] J. von zur Gathen and J. Gerhard, *Modern Computer Algebra*, Cambridge University, 1999.

- [2] É . Bézout, "Recherchez sur le degré des équations résultantes de l'évanouissement des inconnues et sur le moyens qu'il convient d'employer pour trouver ces équations," Mem. Acad, Paris, 288-338, 1764.
- [3] J. Sylvester, "On a Theory of Syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm functions, and that of the greatest Algebraical Common Measure," *Phil. Trans. Roy. Soc. London*, Vol. 143, pp. 407-548, 1853.
- [4] M. Mignotte, *Mathematics for Computer Algebra*, Springer-Verlag, 1992
- [5] R. Zippel, *Effective Polynomial Computation*, Cornell University, 1994.
- [6] H. Lev-Ari, Y. Bistritz, and T. Kailath, "Generalized Bezoutians and families of efficient zero-location procedures," *IEEE Trans. Circuits Syst.* vol. 38, pp. 170-186, Feb. 1991.
- [7] Y. Bistritz "Zero location of polynomials with respect to the unit-circle unhampered by nonessential singularities," *IEEE Trans. on Circuits and Systems, I:Fundam. Theory Appl.* vol. 49, no. 3, pp. 305-314, March 2002.