



Journal of Symbolic Computation 42 (2007) 290-297

Journal of Symbolic Computation

www.elsevier.com/locate/jsc

An elementary proof of Sylvester's double sums for subresultants

Carlos D'Andrea^a, Hoon Hong^b, Teresa Krick^c, Agnes Szanto^{b,*}

^a Department d'Àlgebra i Geometria, Facultat de Matemàtiques, Universitat de Barcelona, Gran Via de les Corts Catalanes, 585; 08007, Spain

^b Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA
^c Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires,
Ciudad Universitaria, 1428 Buenos Aires, Argentina

Received 20 September 2005; accepted 14 September 2006 Available online 13 November 2006

Abstract

In 1853 Sylvester stated and proved an elegant formula that expresses the polynomial subresultants in terms of the roots of the input polynomials. Sylvester's formula was also recently proved by Lascoux and Pragacz using multi-Schur functions and divided differences. In this paper, we provide an elementary proof that uses only basic properties of matrix multiplication and Vandermonde determinants.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Subresultants; Double-sum formula; Vandermonde determinant

1. Introduction

Subresultants play a fundamental role in Computer Algebra and Computational Algebraic Geometry (for instance, see Collins (1967), Brown and Traub (1971), Collins (1975), Gonzalez-Vega et al. (1989), Renegar (1992), Apéry and Jouanolou (2005), Gonzalez-Vega (1996), Lombardi et al. (2000) and Hong (2001)). Sylvester (1853) stated and proved an elegant formula that expresses the polynomial subresultants of two polynomials in terms of their roots, the so-called *double-sum* formula. This identity was proved also by Lascoux and Pragacz (2003), by using the theory of multi-Schur functions and divided differences.

^{*} Corresponding author. Tel.: +1 919 622 7454; fax: +1 919 513 7336. *E-mail addresses:* cdandrea@ub.edu (C. D'Andrea), hong@math.ncsu.edu (H. Hong), krick@dm.uba.ar (T. Krick), aszanto@ncsu.edu (A. Szanto).

In this paper we provide a new and elementary proof that uses only the basic properties of matrix multiplication and Vandermonde determinants. As apparent in our proof, Sylvester's double-sum formula is only one simple step further from a particular case, the so-called *single-sum* formula. Such connection between the single and the double-sum formulae was originally thought to be unlikely, as remarked in page 691 of Lascoux and Pragacz (2003). There have been various proofs for the single-sum formula (Borchardt, 1860; Chardin, 1990; Apéry and Jouanolou, 2005; Hong, 1999; Diaz-Toca and Gonzalez-Vega, 2004).

The matrix multiplication technique, presented in this paper, has proven to be quite powerful in that it is easily generalizable to multivariate polynomials: similar techniques were successfully applied to obtain expressions for multivariate subresultants in roots in D'Andrea et al. (2006), and the generalization of Sylvester's single and double-sum formulae to the multivariate case is the subject of ongoing research.

2. Review of Sylvester's double sum for subresultants

Let $f = a_m x^m + \cdots + a_0$ and $g = b_n x^n + \cdots + b_0$, be two polynomials with coefficients in a commutative ring. The dth subresultant polynomial $\operatorname{Sres}_d(f,g)$ is defined for $0 \le d < \min\{m,n\}$, and if $m \ne n$ holds, also for $d = \min\{m,n\}$, as the following determinant:

$$Sres_{d}(f,g) := det \begin{bmatrix} a_{m} & \cdots & \cdots & a_{d+1-(n-d-1)} & x^{n-d-1}f(x) \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m} & \cdots & a_{d+1} & f(x) \end{bmatrix} \\ b_{n} & \cdots & \cdots & b_{d+1-(m-d-1)} & x^{m-d-1}g(x) \\ \vdots & \vdots & & \vdots \\ b_{n} & \cdots & b_{d+1} & g(x) \end{bmatrix}$$
 (1)

where $a_{\ell} = b_{\ell} = 0$ for $\ell < 0$.

By developing this determinant by the last column, it is clear that $\operatorname{Sres}_d(f,g)$ is a polynomial combination of f and g. It is also a classic fact that $\operatorname{Sres}_d(f,g)$ is a polynomial of degree bounded by d, since it coincides with the determinant of the matrix obtained by replacing the last column C_{m+n-2d} with

$$C'_{m+n-2d} := C_{m+n-2d} - x^{d+1}C_{m+n-2d-1} - \dots - x^{m+n-d-1}C_1.$$

Now, let $A=(\ldots,\alpha,\ldots)$ and $B=(\ldots,\beta,\ldots)$ be finite lists (ordered sets) of distinct indeterminates. Sylvester (1853) introduced for $0 \le p \le |A|, \ 0 \le q \le |B|$ the following *double-sum* expression in A and B:

$$\operatorname{Sylv}^{p,q}(A,B;x) := \sum_{\substack{A' \subset A, B' \subset B \\ |A'| = p, |B'| = q}} R(x,A') R(x,B') \frac{R(A',B') R(A \setminus A', B \setminus B')}{R(A',A \setminus A') R(B',B \setminus B')},$$

where

$$R(X,Y) := \prod_{x \in X, y \in Y} (x-y), \quad R(x,Y) := \prod_{y \in Y} (x-y).$$

Sylvester (1853) gave the following elegant formula that expresses the subresultants in terms of the double sum, that is, in terms of the roots of f and g.

Theorem 1 (Sylvester's Double-sum Formula). Let f, g be the monic polynomials

$$f = \prod_{\alpha \in A} (x - \alpha), \quad g = \prod_{\beta \in B} (x - \beta) \in \mathbb{Z}[\alpha \in A, \beta \in B][x],$$

where |A| = m and |B| = n. Let $p, q \ge 0$ be such that $d := p + q < \min\{m, n\}$ or $d \le \min\{m, n\}$ if $m \ne n$ holds. Then

$$\operatorname{Sres}_{d}(f,g) = \frac{(-1)^{p(m-d)}}{\binom{d}{p}} \operatorname{Sylv}^{p,q}(A,B;x).$$

When p = d and q = 0, the above expression immediately simplifies to the *single-sum* formula:

$$\operatorname{Sres}_{d}(f,g) = \sum_{\substack{A' \subset A \\ |A'| = d}} R(x,A') \frac{R(A \backslash A',B)}{R(A \backslash A',A')}.$$
 (2)

Complete proofs of Sylvester's double sum can be found in Sylvester (1853) and Lascoux and Pragacz (2003), while the single-sum formula has various proofs (Borchardt, 1860; Chardin, 1990; Apéry and Jouanolou, 2005; Hong, 1999; Diaz-Toca and Gonzalez-Vega, 2004). Here we present in Section 4 an alternative elementary proof for both results.

3. Notations

We recall that $0 \le d < \min\{m, n\}$ or $d \le \min\{m, n\}$ if $m \ne n$ holds. We let M_f and M_g denote the following matrices:

$$M_f := \begin{bmatrix} a_0 & \dots & a_m \\ & \ddots & & \ddots \\ & & a_0 & \dots & a_m \end{bmatrix}_{n-d} , \qquad M_g := \begin{bmatrix} b_0 & \dots & b_n \\ & \ddots & & \ddots \\ & & b_0 & \dots & b_n \end{bmatrix}_{m-d} .$$

We now define

Finally, we define for a polynomial p(t) and two lists, $\Gamma := (\gamma_1, \dots, \gamma_u)$ of scalars and $E := (e_1, \dots, e_v)$ of non-negative integers, the (not-necessarily square) matrix of size $v \times u$:

$$\langle p(t), \Gamma \rangle_E := \begin{bmatrix} \gamma_1^{e_1} p(\gamma_1) & \dots & \gamma_u^{e_1} p(\gamma_u) \\ \vdots & & \vdots \\ \gamma_1^{e_v} p(\gamma_1) & \dots & \gamma_u^{e_v} p(\gamma_u) \end{bmatrix}_v.$$

For instance, under this notation, if we take E := (0, ..., u - 1), we have the following equality for the Vandermonde determinant $\mathcal{V}(\Gamma)$ associated to Γ :

$$\mathcal{V}(\Gamma) := |(\gamma_i^{i-1})_{1 \le i, j \le u}| = |\langle 1, \Gamma \rangle_E|.$$

When E is of the form E = (0, ..., v - 1), we directly write $\langle p(t), \Gamma \rangle_v$. We mention the following useful equalities that hold since $m + n - d \ge \max(m, n)$:

$$M_f \cdot \langle 1, \Gamma \rangle_{m+n-d} = \langle f(t), \Gamma \rangle_{n-d}$$

$$M_g \cdot \langle 1, \Gamma \rangle_{m+n-d} = \langle g(t), \Gamma \rangle_{m-d}$$

$$M_{t-r} \cdot \langle 1, \Gamma \rangle_{m+n-d} = \langle t - x, \Gamma \rangle_{d}.$$

4. The proof

The proof is divided into a series of lemmas which are interesting on their own. For an easier understanding, we recommend not to pay attention to signs in a first approach.

Lemma 1. Under the previous assumptions and notations, we have

$$Sres_d(f, g) = (-1)^{d + (n-d)(m-d)} |S_d|.$$

Proof. We denote by C_i the *i*th column of the matrix S_d and we replace its first column C_1 by $C'_1 := C_1 + xC_2 + \cdots + x^{m+n-d-1}C_{m+n-d}$. This operation does not change the determinant of this matrix, and

$$C'_{1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ f(x) \\ \vdots \\ x^{n-d-1}f(x) \end{bmatrix} \xrightarrow{m-d} .$$

$$\begin{bmatrix} x^{m-d-1}g(x) \\ \vdots \\ x^{m-d-1}g(x) \end{bmatrix} \xrightarrow{m-d} .$$

We now perform a Laplace expansion of the determinant of the new matrix over the first d rows, and we observe that only one block survives, which corresponds to columns 2 to d+1 of M_{t-x} . Moreover, this block is lower triangular with diagonal entries 1. Thus

$$|S_d| = (-1)^d \det \begin{bmatrix} f(x) & a_{d+1} & \dots & a_m \\ \vdots & \vdots & & \ddots & \\ x^{n-d-1}f(x) & a_{d+1-(n-d-1)} & \dots & & \dots & a_m \\ g(x) & b_{d+1} & \dots & b_n \\ \vdots & \vdots & & \ddots & \\ x^{m-d-1}g(x) & b_{d+1-(m-d-1)} & \dots & & \dots & b_n \end{bmatrix}^{n-d}$$

$$= (-1)^{d+(n-d)(m-d)} \operatorname{Sres}_d(f, g),$$

since the matrix in the right-hand side above is the matrix of (1) viewed backward. \square

For simplicity, from now on, we assume f and g to be the monic polynomials $f = \prod_{\alpha \in A} (x - \alpha)$, $g = \prod_{\beta \in B} (x - \beta)$ where A and B are lists with |A| = m and |B| = n. (As pointed out by a referee, under this assumption one has in the language of multi-Schur functions: $|S_d| = S_{1d:(m-d)^{n-d}:0^{m-d}}(-x, -A, -B)$ (see Lascoux and Pragacz (2003)).)

The lemmas below generalize in an obvious manner to non-monic polynomials. The first one corresponds to Th. 3 in Hong (1999). We prove it here with a different technique that follows from Lemma 1.

Lemma 2 (Hong's Subresultant in Roots (Hong, 1999, Th. 3.1)). Under the previous notations, we have

$$\operatorname{Sres}_{d}(f,g) \, \mathcal{V}(A) = \det \begin{bmatrix} \frac{m}{\langle x-t,A\rangle_{d}} & d\\ \langle g(t),A\rangle_{m-d} & m-d \end{bmatrix}.$$

Proof. We note that $|S_d| \mathcal{V}(A)$ is the determinant of the following product of matrices:

since
$$\langle f(t), A \rangle_{n-d} = \left[\alpha_j^{i-1} f(\alpha_j) \right] = [0].$$

By permuting the rows of the second block with those of the third, we obtain

$$\operatorname{Sres}_{d}(f,g) \mathcal{V}(A) = (-1)^{d+(m-d)(n-d)} |S_{d}| \mathcal{V}(A)$$

$$= (-1)^{d} \det \frac{\langle t - x, A \rangle_{d}}{\langle g(t), A \rangle_{m-d}} |M'_{f}|$$

$$= \det \frac{\langle x - t, A \rangle_{d}}{\langle g(t), A \rangle_{m-d}},$$

since M'_f is a lower triangular matrix with diagonal entries $a_m = 1$. \square

(3)

Let us remark here that the Poisson product formula $\operatorname{Res}(f,g) = \prod_{\alpha \in A} g(\alpha)$ is a direct consequence of the previous lemma for the case d=0.

For $S \subseteq T$ finite lists, let $\operatorname{sg}(S,T) := (-1)^{\sigma}$ where σ is the number of transpositions needed to take T to $S \cup (T \setminus S)$. Here, " \cup " stands for list concatenation and " \setminus " means list subtraction.

Lemma 3. Let P and Q be two disjoint sublists of E := (0, ..., d-1) that satisfy $P \cup Q = E$, and let p := |P|, q := |Q|. Then

$$\operatorname{Sres}_{d}(f,g)\,\mathcal{V}(A)\mathcal{V}(B) = (-1)^{q+(m-d)n}\operatorname{sg}(P,E)\det\begin{bmatrix} \frac{m}{\langle x-t,A\rangle_{P}} & 0\\ 0 & \langle x-t,B\rangle_{Q}\\ \hline \langle 1,A\rangle_{m+n-d} & \langle 1,B\rangle_{m+n-d} \end{bmatrix} \stackrel{p}{q}_{m+n-d}.$$

Proof. Recalling that $\mathcal{V}(B) = |\langle 1, B \rangle_n|$, we have by Lemma 2:

$$\operatorname{Sres}_{d}(f,g) \mathcal{V}(A) \mathcal{V}(B) = \det \begin{bmatrix} m & n \\ \langle x - t, A \rangle_{d} & 0 \\ \langle g(t), A \rangle_{m-d} & 0 \\ \langle 1, A \rangle_{n} & \langle 1, B \rangle_{n} \end{bmatrix}_{n}^{d}$$

$$= (-1)^{(m-d)n} \det \begin{bmatrix} m & n \\ \langle x-t, A \rangle_d & 0 \\ \langle 1, A \rangle_n & \langle 1, B \rangle_n \\ \langle g(t), A \rangle_{m-d} & 0 \end{bmatrix}_{m-d}^{d}$$

$$= (-1)^{(m-d)n} \det \begin{pmatrix} & d & n & m-d & m & n \\ d & I_d & 0 & 0 \\ n & 0 & I_n & 0 \\ m-d & 0 & M_g & & & & & & & & & \\ \end{pmatrix} \frac{\langle x-t,A\rangle_d & 0}{\langle 1,A\rangle_{m+n-d} & \langle 1,B\rangle_{m+n-d}} d_{m+n-d},$$

since $M_g \cdot \langle 1, B \rangle_{m+n-d} = \langle g(t), B \rangle_{m-d} = [0]$. Now, since the first matrix is lower triangular with diagonal entries 1, we have

$$\operatorname{Sres}_{d}(f,g) \mathcal{V}(A) \mathcal{V}(B) = (-1)^{(m-d)n} \det \begin{bmatrix} \frac{m}{\langle x-t,A\rangle_{d}} & 0\\ \hline \langle 1,A\rangle_{m+n-d} & \langle 1,B\rangle_{m+n-d} \end{bmatrix}^{d}_{m+n-d} . \tag{4}$$

Finally, recalling that $\langle x - t, A \rangle_d = \left(\alpha_j^{i-1} x - \alpha_j^i\right)_{1 \le i \le d, 1 \le j \le m}$ and $\langle 1, A \rangle_{m+n-d} = \left(\alpha_j^{i-1}\right)_{1 \le i \le m+n-d, 1 \le j \le m}$, the obvious subtractions and permutations of rows yield

$$\operatorname{Sres}_d(f,g)\,\mathcal{V}(A)\mathcal{V}(B) = (-1)^{(m-d)n}\operatorname{sg}\left(P,E\right)\det\begin{bmatrix} \frac{m}{\langle x-t,A\rangle_P} & 0 \\ 0 & -\langle x-t,B\rangle_Q \\ \hline \langle 1,A\rangle_{m+n-d} & \langle 1,B\rangle_{m+n-d} \end{bmatrix} \stackrel{p}{q}_{m+n-d}.$$

The lemma follows by moving $(-1)^q$ out of the determinant. \square

We will also need the following observation:

Observation 1. Let $\Gamma := (\gamma_1, \dots, \gamma_d)$. Then

$$|\langle x - t, \Gamma \rangle_d| = R(x, \Gamma)|\langle 1, \Gamma \rangle_d|. \tag{5}$$

Proof. The claim follows from

$$\begin{pmatrix} x - \gamma_1 & \dots & x - \gamma_d \\ \vdots & & \vdots \\ \gamma_1^{d-1} x - \gamma_1^d & \dots & \gamma_d^{d-1} x - \gamma_d^d \end{pmatrix} = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ \gamma_1^{d-1} & \dots & \gamma_d^{d-1} \end{pmatrix} \begin{pmatrix} x - \gamma_1 & & & \\ & \ddots & & \\ & & x - \gamma_d \end{pmatrix}.$$

4.1. Proof of Theorem 1

For any P and Q disjoint sublists of E := (0, ..., d-1) that satisfy $P \cup Q = E$, with |P| = p and |Q| = q, a Laplace expansion over the first d rows in Identity (3) gives that $\operatorname{Sres}_d(f, g) \mathcal{V}(A) \mathcal{V}(B)$ equals

$$(-1)^{\sigma} \operatorname{sg}(P, E) \sum_{\substack{A' \subset A, B' \subset B \\ |A'| = p, |B'| = q}} \operatorname{sg}(A', A) \operatorname{sg}(B', B) \cdot |\langle x - t, A' \rangle_{P}| \cdot |\langle x - t, B' \rangle_{Q}|$$

$$\cdot \mathcal{V}(A \backslash A' \cup B \backslash B')$$

where $\sigma := q + (m-d)n + (m-p)q \equiv (m-d)(n-q) \pmod{2}$. Adding over all such choices of $P \subset E$ with |P| = p, we deduce that $\operatorname{Sres}_d(f, g) \mathcal{V}(A) \mathcal{V}(B)$ equals

$$\frac{1}{\binom{d}{p}} \sum_{P} (-1)^{\sigma} \operatorname{sg}(P, E) \sum_{A', B'} \operatorname{sg}(A', A) \operatorname{sg}(B', B) \cdot |\langle x - t, A' \rangle_{P}| \cdot |\langle x - t, B' \rangle_{Q}|$$

$$\cdot \mathcal{V} \left(A \backslash A' \cup B \backslash B' \right)$$

$$= \frac{(-1)^{\sigma}}{\binom{d}{p}} \sum_{A', B'} \operatorname{sg}(A', A) \operatorname{sg}(B', B) \mathcal{V} \left(A \backslash A' \cup B \backslash B' \right)$$

$$\cdot \left(\sum_{P} \operatorname{sg}(P, E) |\langle x - t, A' \rangle_{P}| \cdot |\langle x - t, B' \rangle_{Q}| \right).$$

We observe now that, by another Laplace expansion and Identity (5),

$$\sum_{P} \operatorname{sg}(P, E) |\langle x - t, A' \rangle_{P}| \cdot |\langle x - t, B' \rangle_{Q}| = |\langle x - t, A' \cup B' \rangle_{d}|$$

$$= R(x, A') R(x, B') |\langle 1, A' \cup B' \rangle_{d}|.$$

Recalling that $|\langle 1, A' \cup B' \rangle_d| = \mathcal{V}(A' \cup B')$, this gives

$$\begin{split} &\operatorname{Sres}_{d}(f,g) \\ &= \frac{(-1)^{\sigma}}{\binom{d}{p}} \sum_{A',B'} R(x,A') R(x,B') \frac{\operatorname{sg}(A',A) \operatorname{sg}(B',B) \mathcal{V} \left(A \backslash A' \cup B \backslash B'\right) \mathcal{V} \left(A' \cup B'\right)}{\mathcal{V}(A) \mathcal{V}(B)} \\ &= \frac{(-1)^{\sigma} (-1)^{\tau}}{\binom{d}{p}} \sum_{A',B'} R(x,A') R(x,B') \frac{R(A',B') R(A \backslash A',B \backslash B')}{R(A',A \backslash A') R(B',B \backslash B')}, \end{split}$$

where $\tau = (m-p)(n-q) + pq - (m-p)p - (n-q)q = (m-d)(n-d)$ since for any finite lists X, Y, one has $\mathcal{V}(X \cup Y) = \mathcal{V}(X)\mathcal{V}(Y)R(Y, X) = (-1)^{|X| \cdot |Y|}\mathcal{V}(X)\mathcal{V}(Y)R(X, Y)$. The claim follows now from the fact that $(m-d)(n-q) + (m-d)(n-d) \equiv (m-d)p \pmod{2}$.

As a final remark, we mention that if in the previous proof we start with a Laplace expansion over the first d rows in Identity (4) instead of Identity (3), we obtain in the same manner Sylvester's single sum formulation (2).

Acknowledgements

We are grateful to the anonymous referees for their careful reading of our preliminary manuscript and their very precise indications to improve our presentation.

HH's research was supported by NSF grant CCR-0097976, TK's by CONICET PIP 2461/01 and UBACYT X-112 grants and AS's by NSF grants CCR-0306406 and CCR-0347506.

References

Apéry, F., Jouanolou, J.P., 2005. Résultant et sous-résultants: le cas d'une variable, Cours DESS 1995-1996.

Borchardt, C.W., 1860. Über eine Interpolationsformel für eine Art Symmetrischer Functionen und über Deren Anwendung. Math. Abh. der Akademie der Wissenschaften zu Berlin, pp. 1–20.

Brown, W.S., Traub, J.F., 1971. On Euclid's algorithm and the theory of subresultants. Journal of the ACM 18 (4), 505–514.

Chardin, M., 1990. Thèse. Université Pierre et Marie Curie (Paris VI).

Collins, G.E., 1967. Subresultants and reduced polynomial remainder sequences. Journal of the ACM 14, 128-142.

Collins, G.E., 1975. Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition. In: Lecture Notes in Computer Science, vol. 33. Springer-Verlag, Berlin, pp. 134–183.

D'Andrea, C., Krick, T., Szanto, A., 2006. Multivariate subresultants in roots. Journal of Algebra 302 (1), 16–36.

Diaz-Toca, G.M., Gonzalez-Vega, L., 2004. Various new expressions for subresultants and their applications. Applicable Algebra Engineering Communication Computing 15 (3–4), 233–266.

Gonzalez-Vega, L., 1996. A combinatorial algorithm solving some quantifier elimination problems. In: Caviness, B., Johnson, J. (Eds.), Quantifier Elimination and Cylindrical Algebraic Decomposition. In: Texts and Monographs in Symbolic Computation, Springer-Verlag.

Gonzalez-Vega, L., Lombardi, H., Recio, T., Roy, M.-F., 1989. Sturm–Habicht sequences. In: Proceedings of the ACM-SIGSAM 1989 International Symposium on Symbolic and Algebraic Computation, July 1989, pp. 136–146.

Hong, H., 1999. Subresultants in roots. Technical Report. Department of Mathematics, North Carolina State University. Hong, H., 2001. Ore subresultant coefficients in solutions. Journal of Applicable Algebra in Engineering, Communication, and Computing 12 (5), 421–428.

Lascoux, A., Pragacz, P., 2003. Double Sylvester sums for euclidean division, multi-Schur functions. Journal of Symbolic Computation 35, 689–710.

Lombardi, H., Roy, M.-F., El Din, M.S., 2000. New structure theorem for subresultants. Journal of Symbolic Computation 29, 663–689.

Renegar, J., 1992. On the computational complexity and geometry of the first-order theory of the reals. Journal of Symbolic Computation 13 (3), 255–352.

Sylvester, J.J., 1853. On a theory of syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's function and that of the greatest algebraical common measure. Trans. Roy. Soc. London. Reprinted in: The Collected Mathematical Papers of James Joseph Sylvester, vol. 1, Chelsea Publ., New York, 1973, pp. 429–586.