Various New Methods for Computing Subresultant Polynomial Remainder Sequences (PRS's)

Alkiviadis G. Akritas

University of Thessaly, Volos, Greece, akritas@uth.gr

To the memory of Anna Johnson Pell[1] *and R. L. Gordon,*
for their inspiring Theorem of 1917![2]

Teaching subresultant prs's is an unpleasant experience because there is a misunderstanding about the role of Sylvester's two matrices and how they affect the signs of the sequences. Almost all articles and texts on the subject perform operations in $\mathbf{Z}[x]$ and use a form of pseudo-division that distorts the signs of the polynomial remainders; hence, sentences like "forget about the signs" appear quite often in the literature. In this talk we clarify the mystery about the signs and show how to compute the subresultant prs's in various ways — performing operations even in $\mathbf{Q}[x]$. Briefly stated, here is how.

Consider the polynomials $f, g \in \mathbf{Z}[x]$ of degrees n, m, respectively, with $n > m$. We call *Euclidean* prs the sequence of polynomial remainders obtained during the execution of the Euclidean algorithm for polynomial gcd. If the polynomials in the sequence are of degrees $n = m+1, m, m-1, m-2, \ldots, 0$, the sequence is called *complete*. Otherwise it is called *incomplete*.

The *complete* Euclidean prs of two polynomials can be computed either by doing polynomial divisions over the integers/rationals or by evaluating determinants of submatrices of `sylvester1` — J.J. Sylvester's matrix of 1840 [1], [12]. In the latter case, the coefficients of each polynomial remainder are the above mentioned determinants (or subresultants) and we are talking about the *subresultant* prs [6], [7], [8], [10].

Caveat 1: As demonstrated by $f = x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5$ and $g = 3x^6 + 5x^4 - 4x^2 - 9x + 21$, the signs of the polynomials in an *incomplete* Euclidean prs may differ from those of the corresponding subresultant prs [3].

Analogous to Euclidean prs's are the Sturm sequences, which are obtained by *modifying* Euclid's algorithm; that is, at each step we take the negative of the remainder obtained.

Like their cousins, *complete* Sturm sequences can be computed either by doing polynomial divisions over the integers/rationals or by evaluating determinants of

[1] See the link `http://en.wikipedia.org/wiki/Anna_Johnson_Pell_Wheeler` for her biography.

[2] Discovered by Panagiotis S. Vigklas.

submatrices of sylvester2 — J.J. Sylvester's matrix of 1853 [13]. In the latter case, the coefficients of each polynomial remainder are the *modified* subresultants and we have the *modified subresultant* prs [5].

Caveat 2: As demonstrated by $f = x^5 - 3x - 1$ and $g = 5x^4 - 3$, the signs of the polynomials in an *incomplete* Sturm sequence may differ from those of the corresponding modified subresultant prs [5].

Recall that det(sylvester1) defines the resultant of two polynomials and that in general det(sylvester1) $\neq$ det(sylvester2). A detailed discussion on these two matrices can be found elsewhere [4].

In 1900, E.B. Van Vleck, [14], computed *complete* Sturm sequences by triangularizing sylvester2. Akritas extended Van Vleck's triangularization method for *incomplete* subresultant prs's, [2], but in this case it was impossible to compute the correct sign of the polynomials in the sequence. The solution [4] came with the discovery, by Vigklas, of the Pell-Gordon theorem of 1917 [11].

In short, the Pell-Gordon theorem was a response to Van Vleck's work and is precisely the tool needed to compute the correct sign of the polynomials in an *incomplete* Sturm sequence computed with the triangularization method [5]. The only difference from what we are used today is the fact that Pell and Gordon do their computations in $\mathbf{Q}[x]$. Their theorem is stated below, whereas a detailed example can be found elsewhere [5].

Theorem (Pell-Gordon, 1917): Let

$$A = a_0x^n + a_1x^{n-1} + \cdots + a_n$$

and

$$B = b_0x^n + b_1x^{n-1} + \cdots + b_n$$

be two polynomials of the n-th degree. Modify the process of finding the highest common factor of A and B by taking at each stage the negative of the remainder. Let the i-th modified remainder be

$$R^{(i)} = r_0^{(i)}x^{m_i} + r_1^{(i)}x^{m_i-1} + \cdots + r_{m_i}^{(i)}$$

where $(m_i + 1)$ is the degree of the preceeding remainder, and where the first $(p_i - 1)$ coefficients of $R^{(i)}$ are zero, and the p_i-th coefficient $\rho_i = r_{p_i-1}^{(i)}$ is different from zero. Then for $k = 0, 1, \ldots, m_i$ the coefficients $r_k^{(i)}$ are given by[3]

$$r_k^{(i)} = \frac{(-1)^{u_{i-1}}(-1)^{u_{i-2}}\cdots(-1)^{u_1}(-1)^{v_{i-1}}}{\rho_{i-1}^{p_{i-1}+1}\rho_{i-2}^{p_{i-2}+p_{i-1}}\cdots\rho_1^{p_1+p_2}\rho_0^{p_1}} \cdot \det(i,k), \tag{1}$$

[3] It is understood in (1) that $\rho_0 = b_0$, $p_0 = 0$, and that $a_i = b_i = 0$ for $i > n$.

where $\quad u_{i-1} = 1+2+\cdots+p_{i-1}, \quad v_{i-1} = p_1+p_2+\cdots+p_{i-1} \quad$ and

$$\det(i,k) = \begin{vmatrix} a_0 & a_1 & a_2 & \cdots & \cdot & \cdot & \cdots & a_{2v_{i-1}} & a_{2v_{i-1}+1+k} \\ b_0 & b_1 & b_2 & \cdots & \cdot & \cdot & \cdots & b_{2v_{i-1}} & b_{2v_{i-1}+1+k} \\ 0 & a_0 & a_1 & \cdots & \cdot & \cdot & \cdots & a_{2v_{i-1}-1} & a_{2v_{i-1}+k} \\ 0 & b_0 & b_1 & \cdots & \cdot & \cdot & \cdots & b_{2v_{i-1}-1} & b_{2v_{i-1}+k} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & a_0 & a_1 & \cdots & a_{v_{i-1}} & a_{v_{i-1}+1+k} \\ 0 & 0 & 0 & \cdots & b_0 & b_1 & \cdots & b_{v_{i-1}} & b_{v_{i-1}+1+k} \end{vmatrix}.$$

Proof: The proof by induction of this theorem depends on two Lemmas that can be found in the original paper of Pell and Gordon.

As indicated elsewhere [5], we use a modification of formula (1) to compute the coefficients of the Sturm sequence. In that case $p_0 = \deg(A) - \deg(B) = 1$, since B is the derivative of A and, hence, the modified formula is shown below with the changes appearing in bold:

$$r_k^{(i)} = \frac{(-1)^{u_{i-1}}(-1)^{u_{i-2}}\cdots(-1)^{u_1}(-1)^{u_0}(-1)^{v_{i-1}}}{\rho_{i-1}^{\mathbf{p_{i-1}+p_i-degDiffer}}\rho_{i-2}^{p_{i-2}+p_{i-1}}\cdots\rho_1^{p_1+p_2}\rho_0^{\mathbf{p_0}+p_1}} \cdot \frac{\det(i,k)}{\rho_{\mathbf{-1}}}, \tag{2}$$

where $\rho_{-1} = a_0$, the leading coefficient of A and `degDiffer` is the difference between the expected degree m_i and the actual degree of the remainder.

It should be noted that in our (general) case $p_0 = \deg(A) - \deg(B)$ and that the division $\frac{\det(i,k)}{\rho_{-1}}$ is possible if the leading coefficient of A is the only element in the first column of `sylvester2`. Moreover, if the leading coefficient of A is negative we work with the polynomial negated and at the end we reverse the signs of all polys in the sequence.

Using formula (2) above we were able to compute [5]:

- complete and incomplete Sturm sequences in $\mathbf{Z}[x]$ by doing divisions in $\mathbf{Q}[x]$;
- complete and incomplete *modified* subresultant prs's by evaluating the *sign* of the determinant of an appropriate submatrix of `sylvester2` — one sign computation for each polynomial.

We also wondered whether the Pell-Gordon theorem can help us compute subresultant prs's and we came up with the following rule.

The Sign/Value Rule for subresultant prs's:

To compute the exact sign of a polynomial and (possibly) adjust its value in a complete or incomplete subresultant prs we evaluate the determinant of an appropriate submatrix of `sylvester1` — one determinant computation for each polynomial.

Three new methods were developed using the above rule [3]. They have been implemented in `Sympy` and can be downloaded from `http://inf-server.inf.uth.gr/~akritas/publications/subresultants.py`.

- In the first method, `subresultants_prem2(f, g, x)`, we incorporate the new pseudo remainder function,[4] `prem2(f, g, x)`, which uses the *absolute value* of the leading coefficient of the divisor; that is, `prem2` is based on the identity $|lc(g)|^{deg(f)-deg(g)+1} \cdot f = q \cdot g + r$. This way, we preserve the "correct" sign sequence of the Euclidean prs, as discussed elsewhere [5].
- The second method, `subresultants_PG(f, g, x)`, does divisions over the rationals and uses the Pell-Gordon theorem to convert the coefficients of the polynomial remainders to integers. Here we have an implicit interplay between the two Sylvester matrices, `sylvester1` and `sylvester2`.
- Finally, in the third method, `subresultants_triang(f, g, x)`, we see — for the first time in the literature — an explicit interplay between the two Sylvester matrices, `sylvester1` and `sylvester2`. While we triangularize the latter to obtain polynomial-candidates for the subresultant prs, we evaluate determinants of submatrices of the former in order to make the candidates actual members of the prs by adjusting, if needed, their coefficients accordingly — both in value and sign!

Note that, for all three methods, the cost of computing a single subresultant per remainder is negligible if a probabilistic algorithm is available for computing large determinants — as is the case in the free computer algebra system `Xcas`. Moreover, as mentioned in [3], this cost can be further decreased if in the Sign/Value Rule — instead of `sylvester1` — we use submatrices of other, equivalent, matrices with *smaller* dimensions [9].

References

[1] A.G. Akritas, *A Simple Proof of the Validity of the Reduced prs Algorithm*, Computing, **38**, pp. 369-372 (1987).

[4] To compute pseudo-remainders `Sympy` has the built-in function `prem(f, g, x)`, which uses the leading coefficient of the divisor along with its sign, as described in [7], [8]; that is, `prem` is based on the identity $lc(g)^{deg(f)-deg(g)+1} \cdot f = q \cdot g + r$.

[2] A.G. Akritas, *A New Method for Computing Polynomial Greatest Common Divisors and Polynomial Remainder Sequences*, Numerische Mathematik, **52**, pp. 119-127 (1988).

[3] A.G. Akritas, *Three New Methods for Computing Subresultant Polynomial Remainder Sequences (PRS's)*, Serdica Journal of Computing, to appear.

[4] A.G. Akritas, G.I. Malaschonok and P.S. Vigklas, *On a Theorem by Van Vleck Regarding Sturm Sequences*, Serdica Journal of Computing, **7**(4), pp. 101-134 (2013).

[5] A.G. Akritas, G.I. Malaschonok and P.S. Vigklas, *Sturm Sequences and Modified Subresultant Polynomial Remainder Sequences*, Serdica Journal of Computing, to appear.

[6] W.S. Brown and J.F. Traub, *On Euclids Algorithm and the Theory of subresultants*, Journal of the ACM, **18**, pp. 505-514 (1971).

[7] W.S. Brown, *The subresultant PRS Algorithm*, ACM Transactions on Mathematical Software, **4**(3), pp. 237-249 (1978).

[8] G.E. Collins, *Subresultants and Reduced Polynomial Remainder Sequences*, Journal of the ACM, **14**, pp. 128-142 (1967).

[9] G.M. Diaz-Toca, and L. Gonzalez-Vega, *Various New Expressions for Subresultants and Their Applications*, Applicable Algebra in Engineering, Communication and Computing, **15**, pp. 233-266 (2004).

[10] W. Habicht, *Eine Verallgemeinerung des Sturmschen Wurzelzählverfahrens*, Commentarii Mathematici Helvetici, **21**, pp. 99-116 (1948).

[11] A.J. Pell and R.L. Gordon, *The Modified Remainders Obtained in Finding the Highest Common Factor of Two Polynomials*, Annals of Mathematics, Second Series, **18**(4), pp. 188-193 (Jun., 1917).

[12] J.J. Sylvester, *A method of determining by mere inspection the derivatives from two equations of any degree*, Philosophical Magazine, **16**, pp. 132-135 (1840).

[13] J.J. Sylvester, *On a theory of the syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraical common measure*, Philoshophical Transactions, **143**, pp. 407-548 (1853)

[14] E.B. Van Vleck, *On the Determination of a Series of Sturm's Functions by the Calculation of a Single Determinant*, Annals of Mathematics, Second Series, **1**(1/4), pp. 1-13 (1899 - 1900).