Computer tools in education, 2016
№ 2: 13–35
UDC: 512.622
http://ipo.spb.ru/journal

ANNA JOHNSON
AND HER SEMINAL THEOREM OF 1917

Alkiviadis Akritas

Abstract

In this article we present the life of Anna Johnson, a woman exceptionally gifted in Mathematics, along with what we consider her greatest contribution: to wit, the theorem of 1917 on modified Euclidean polynomial remainder sequences (prs's), which laid the foundations of the theory of subresultant prs's.

To demonstrate the various mathematical concepts presented in this article we use the `python` based computer algebra system `sympy (version 1.0)`, which is freely available.

Keywords: *Euclidean polynomial remainder sequence (prs), modified Euclidean prs, subresultant prs, modified subresultant prs, Van Vleck's method, Sturm's prs.*

1. ANNA'S BIOGRAPHY

Anna Johnson (May 5, 1883 – March 26, 1966) was born to Swedish immigrant parents in Hawarden, Iowa, USA. As detailed below, in her adult life she is known as Anna Johnson Pell and also as Anna Johnson Pell Wheeler. As time advances in the narrative, we use all three names.

In the fall of 1899 Anna Johnson entered the University of South Dakota, where her older sister Esther was also a student. She and her sister rented a room from the mathematics professor, Alexander Pell, a Russian emigrant (and former Russian double agent) who had received his Ph.D. in mathematics in 1897 from Johns Hopkins University, after having had to flee Russia.[1]

Anna Johnson pursued her graduate studies in mathematics at the University of Iowa, earning a master's degree in 1904 with a thesis on *The extension of the Galois theory to linear differential equations*. A second master's degree followed a year later from Radcliffe College.[2] After an

[1] Alexander Pell was Sergei Degaev (1857–1921) for about the first thirty years of his life. It is a strange story for this man led two totally distinct lives:

- one as Sergei Degaev, a Russian revolutionary and member of Narodnaya Volya, a group which is best remembered for the assassination of Tsar Alexander II in March 1881, and
- the other as Alexander Pell a highly respected American mathematics professor.

Details can be found in http://www-history.mcs.st-andrews.ac.uk/Biographies/Pell_Alexander.html.

[2] One of the "Seven Sisters Colleges", a loose association of seven liberal arts colleges in the Northeastern United States that are historically women's colleges. All were founded between 1837 and 1889. Four are in Massachusetts (MA), two are in New York (NY), and one is in Pennsylvania (PA). The seven colleges are: Barnard College (NY), Bryn Mawr College (PA), Mount Holyoke College (MA), Radcliffe College (MA), Smith College (MA), Vassar College (NY), and Wellesley College (MA). Radcliffe (which merged with Harvard College) and Vassar (which is now coeducational) are no longer women's colleges. The name "Seven Sisters" was given because of their parallel to the Ivy League men's colleges in 1927.

additional year of study at Radcliffe, she won an *Alice Freeman Palmer* Fellowship from Wellesley College[2] to study at Göttingen University[3] during the academic year 1906–1907.

In Göttingen Anna Johnson attended lectures by David Hilbert, Felix Klein, and Hermann Minkowski, who were some of the great mathematicians of the early 20th century.

Alexander Pell, whose wife had died in 1904, continued to correspond with Anna during her years of graduate studies. After the end of her fellowship in July 1907, the two were married in Göttingen despite their 25 years difference in age.

They returned to the University of South Dakota where Anna Johnson Pell taught two courses in the mathematics department during the fall semester. She then returned to Göttingen in the spring of 1908 with the intention of completing her doctoral thesis. However, due to conflicts with Hilbert, she returned to the United States without her degree. As she later put it in a letter to her friend Mary Coes (1861–1913), Dean at Radcliffe (1909–1913):

> *In Göttingen I had some trouble with Professor Hilbert and came back to America without a degree* (1910, [8]).

In January 1909 Anna Johnson Pell entered the graduate program at the University of Chicago where she studied with another distinguished mathematician, E.H. Moore, chair of the department. Her husband at this time was teaching at the Armout Institute of Chicago. After a year of classes at Chicago, she received her Ph.D. in 1909 with a thesis on *Biorthogonal Systems of Functions* that she had originally written (independently of Hilbert) during her time at Göttingen.

With Moore's help Anna Johnson Pell tried to find a teaching position at universities near Chicago. However, as she wrote to Mary Coes:

> *I had hoped for a position in one of the good univ. like Wisc., Ill. etc., but there is such an objection to women that they prefer a man even if he is inferior both in training and research. It seems that professor Moore has also given up hope for he has inquired at some of the Eastern Girls' Colleges and Bryn Mawr[2] is apparently the only one with a vacancy in mathematics* (1910, [8]).

A further setback occurred in the spring of 1911 when her husband suffered a stroke. Anna Johnson Pell stepped in to teach his mathematics classes at the Armout Institute. The following fall she accepted a teaching position at Mt. Holyoke College[2] where she taught for seven years before moving to Bryn Mawr College in 1918 as an associate professor. Alexander Pell never returned to teaching except for one semester at Northwestern, although he did continue working on his research. He died in Bryn Mawr in 1921.

One of the attractions of Bryn Mawr was the chance to work with graduate students in mathematics. Anna Johnson Pell supervised eight Ph.D. students during her career. In 1924 she became head of the mathematics department and in 1925 she was promoted to professor. That same year she also married Arthur Wheeler, a colleague in classics at Bryn Mawr, who moved to Princeton as a professor of Latin.

Anna Johnson Pell Wheeler moved to Princeton also, commuting to Bryn Mawr to teach on a part-time basis. This also allowed her to participate in the mathematical activities at Princeton. When her husband died in 1932, however, she returned to Bryn Mawr as a full-time faculty member and department chair until her retirement in 1948. During this period she played an important role in bringing Emmy Noether[4] to Bryn Mawr (1933). However, Emmy Noether died in 1935.

[3] At that time its Mathematics Department was the best in the world.

[4] Amalie Emmy Noether (March 23, 1882 – April 14, 1935) was a German-Jewish mathematician who had worked in Göttingen.

Anna Johnson Pell Wheeler received numerous honors during her life. In 1927 she became the first woman to give the Colloquium Lectures at the American Mathematical Society (AMS) meetings. Her topic was *Theory of quadratic forms in infinitely many variables and applications*. She was an active member of both the American Mathematical Society and the Mathematical Association of America. She served on the Board of Trustees (1923-1924) and the Council (1924-1926) of the AMS. She was an editor of the *Annals of Mathematics* for 18 years. She received honorary doctorate degrees from the New Jersey College for Women (1932) and Mount Holyoke College (1937). The 1940 Women's Centennial Congress named her as one of the 100 American women to have succeeded in careers not open to women a century before.

Anna Johnson Pell Wheeler was also an influential and dedicated teacher. After her retirement she continued to attend mathematics meetings and correspond with former students. She suffered a stroke in early 1966 and died a few weeks later. Following her wishes, she was buried beside her first husband Alexander Pell in Lower Merion Baptist Church Cemetery in Bryn Mawr.

Additional information about this extraordinary woman can be found in various sites in the internet[5] as well as in the article by Greenstein [8].

2. THE SEMINAL THEOREM OF 1917 ON MODIFIED EUCLIDEAN POLYNOMIAL REMAINDER SEQUENCES

In 1917 Anna Johnson Pell published — together with Ruth L. Gordon — in the Annals of Mathematics a theorem, which laid the foundations of the theory of subresultant and modified subresultant polynomial remainder sequences (prs's) [9].

As indicated by the order of the names, Anna Johnson Pell was the person mostly responsible for the theorem. Therefore, we will mostly refer to this theorem as belonging to Anna; however, at the end we will present it as Pell-Gordon.

This theorem along with the algorithm implied by it were dormant for about a century. It was Akritas, Malaschonok and Vigklas the ones who discovered it, realized its importance and used it for the first time in the literature [2–5].

Below we describe the history of the problem that was solved by this theorem and then discuss its ramifications.

In section 2.1 we introduce Euclidean and modified Euclidean polynomial remainder sequences computed over $\mathbb{Q}[x]$, the rationals and over $\mathbb{Z}[x]$, the integers.

Subsequently, in section 2.2, we define two matrices introduced by Sylvester in 1840 and 1853 and relate them to the polynomial remainder sequences introduced in section 2.1.

A special section, 2.3, is devoted to incomplete polynomial remainder sequences — see Definition 3 in section 2.1 — for which Anna Johnson Pell's theorem was intended.

Finally, in section 2.4 we present the Pell-Gordon theorem of 1917 and its ramifications.

For our discussion we will need the `python` based computer algebra system `sympy 1.0` and in particular its module `subresultants_qq_zz.py`.[6]

We begin our discussion with a review of the Euclidean algorithm for integers. This is by far the most widely known algorithm for the computation of the greatest common divisor (gcd) of

[5] See for example https://en.wikipedia.org/wiki/Anna_Johnson_Pell_Wheeler and http://www-history.mcs.st-andrews.ac.uk/Biographies/Wheeler.html

[6] If one has an earlier version of `sympy`, the module can be downloaded from https://github.com/sympy/sympy/blob/master/sympy/polys/subresultants_qq_zz.py. Obviously, the module can be `load`-ed or `attach`-ed in a session of `sage`, the other freely available `python` based computer algebra system.

two integers $a, b \neq 0$. It is based on the equation

$$\gcd(a, b) = \gcd(b, r),$$

where $r = a - q \cdot b$, the remainder obtained on dividing[7] a by b. The algorithm is already implemented in `sympy` but we implement it again as `our_gcd(a, b)` in order to highlight its main operation, namely the computation of the remainders.

So we open a `python` interface in `TeXmacs`, import from `sympy` the necessary functions and define `our_gcd(a, b)`.

```
   Python 2.7.11 |Anaconda 2.1.0 (x86_64)|(default, Dec  6 2015, 18:57:58)
   [GCC 4.2.1 (Apple Inc. build 5577)]
   Python plugin for TeXmacs.
   Please see the documentation in Help -> Plugins -> Python
>  from sympy import Abs, var, diff, sturm, det
>  from sympy.polys.subresultants_qq_zz import *     # import all functions
>  x = var('x')
>  def our_gcd(a, b):
       """
       Calculate the Greatest Common Divisor of a and b, where b≠0.
       """
       while b:
           a, b = b, a%b
       return Abs(a)
>  our_gcd(-15, 3)
3
```

Notice that inside the `while`-loop of `our_gcd(a, b)` the operation a%b repeatedly computes the integer remainder until b becomes zero. With this observation in mind we are ready to proceed with our discussion.

2.1. Euclidean and Modified Euclidean Polynomial Remainder Sequences

The Euclidean algorithm for computing the greatest common divisor (gcd) of two polynomials works exactly like its counterpart for integers — the only difference being that the polynomial remainders can be computed

- either in $\mathbb{Q}[x]$, that is over the rationals, in which case the polynomials obtained are *uniquely* defined and of little interest,
- or in $\mathbb{Z}[x]$, that is over the integers, in which case the polynomials obtained are *not* uniquely defined. However, of great interest are the polynomial remainders whose integer coefficients can also be computed as the determinants of submatrices of a given matrix. We will examine only this case in detail.

In the sequel we examine these two cases separately.

Definition 1 *The sequence of polynomials obtained by applying the Euclidean algorithm on two polynomials with rational coefficients is called a* ***Euclidean polynomial remainder sequence*** *or* ***Euclidean prs*** *for short.*[8]

[7] By Euclidean division.

[8] It is understood that the two original polynomials belong to the sequence.

However, of great interest in our discussion is another kind of polynomial remainder sequences, which are obtained by **modifying the Euclidean algorithm**. The modification consists in negating — at each iteration of the `while`-loop in the Euclidean algorithm — the polynomial remainder and using the negated polynomial in the next iteration.

The modified Euclidean algorithm is of great importance because when applied to f, g where $g = f'$, the derivative of f, we obtain **Sturm's sequence** of f.[9] The Sturm sequence of a polynomial f can be used to isolate its real roots.

Anna Johnson Pell discovered her seminal theorem while attempting to compute Sturm's sequences in a very peculiar way described in section 2.4.1.

The terminology in the following definition was inspired by the title of Anna Johnson Pell's article [9].

Definition 2 *The sequence of polynomials obtained by applying the modified Euclidean algorithm on two polynomials with rational coefficients is called a **modified Euclidean polynomial remainder sequence** or **modified Euclidean prs** for short.*

Note 1 There is no universally accepted terminology regarding the two prs's defined above. Since Anna's theorem was until recently dormant, people are not even aware of the need to differentiate the two prs's mentioned above.

The vast majority of people concentrate on Euclidean prs's and "mess up" the signs of the Euclidean algorithm — as was observed in http://planetmath.org/sturmstheorem. The rest form a tiny minority, who concentrate on modified Euclidean prs's and call them "signed" prs's to indicate that their signs are important; they erroneously think that Euclidean prs's are "non signed" sequences, implying that their signs are not important and may be arbitrarily changed.

The following definition will be helpful in our discussion.

Definition 3 *A polynomial remainder sequence is called **complete** if the degrees of any two consecutive polynomials differ by one. Otherwise, it is called **incomplete**.*

Great mathematicians of the 19th and early 20th century, like Sylvester and Van Vleck, could handle complete prs's but the incomplete ones were beyond their grasp [11, p. 419], [13, p. 4]. It was on the latter sequences that Anna Johnson Pell left her mark [9].

2.1.1. Polynomial Remainder Sequences in $\mathbb{Q}[x]$

To compute prs's over the rationals we use the `sympy` function `rem(f, g, x)`. In the following example we first compute the Euclidean prs and then the modified Euclidean prs.

Example 1 Consider the polynomials $f = x^3 + 5x^2 - 7x + 7$ and $g = 5x^2 - 6x + 8$. To compute the Euclidean prs in $\mathbb{Q}[x]$, we utilize the function `euclid_q(f, g, x)` and obtain the sequence of polynomials shown below

```
>  f = x**3 + 5*x**2 - 7*x + 7
>  g = 5*x**2 - 6*x + 8
>  euclid_q(f, g, x)
[x**3 + 5*x**2 - 7*x + 7, 5*x**2 - 6*x + 8, -29*x/25 - 73/25, 46075/841]
```

[9] See also section 2.4.1.

Notice that the first remainder $r_1 = -\frac{29}{25}x - \frac{73}{25}$ — which is the third polynomial in the above sequence — is obtained by dividing f by g. Likewise, the second remainder $r_2 = \frac{46075}{841}$ — which is the constant (last) polynomial in the above sequence — is obtained by dividing g by r_1. Indeed, we have:

```
>  r1 = rem(f, g, x); print(r1)
-29*x/25 - 73/25
>  r2 = rem(g, r1, x); print(r2)
46075/841
```

The constant polynomial $r_2 = \frac{46075}{841}$ is a multiple of 1, and it is the latter that is taken as the gcd of the f, g. Indeed, we have[10]

```
>  gcd(f, g, x)
1
```

For the same polynomials f, g defined above, the modified Euclidean prs in $\mathbb{Q}[x]$ is obtained by employing the function `sturm_q(f, g, x)`.

```
>  sturm_q(f, g, x)
[x**3 + 5*x**2 - 7*x + 7, 5*x**2 - 6*x + 8, 29*x/25 + 73/25, -46075/841]
```

The first remainder $R_1 = \frac{29}{25}x + \frac{73}{25}$ — which is the third polynomial in the above sequence — is the negation of the remainder obtained by dividing f by g. Likewise, the second remainder $R_2 = -\frac{46075}{841}$ — which is the constant (last) polynomial in the above sequence — is the negation of the remainder obtained by dividing g by R_1. Indeed, we have:

```
>  R1 = -rem(f, g, x); print(R1)
29*x/25 + 73/25
>  R2 = -rem(g, R1, x); print(R2)
-46075/841
```

The prs's $[f, g, r_1, r_2]$ and $[f, g, R_1, R_2]$ in $\mathbb{Q}[x]$ computed above are both complete.

Note that if $g = fp = \frac{d}{dx}(f)$, that is, g is the derivative of f, then the modified Euclidean prs is identical to the Sturm sequence of f, which in sympy is computed by the function `sturm(f, x)`. Hence, the name of the function to compute the modified Euclidean prs. Indeed, if

```
>  fp = diff(f, x, 1); print(fp)
3*x**2 + 10*x - 7
```

then `sturm_q(f, fp, x)` = `sturm(f, x)`.

```
>  sturm_q(f, fp, x)
[x**3 + 5*x**2 - 7*x + 7, 3*x**2 + 10*x - 7, 92*x/9 - 98/9, -14931/2116]
>  sturm(f, x)
[x**3 + 5*x**2 - 7*x + 7, 3*x**2 + 10*x - 7, 92*x/9 - 98/9, -14931/2116]
```

[10] In sympy the function gcd is applied to integers as well as to polynomials.

2.1.2. Polynomial Remainder Sequences in $\mathbb{Z}[x]$

As indicated by equations (1) and (2) below, to compute prs's over the integers we have to *first* premultiply the dividend times the *absolute value* of the leading coefficient of the divisor raised to the power δ and *then* use the `sympy` function `rem(f, g, x)`. This is done by the function `rem_z(f, g, x)`, which is defined by

$$|\mathrm{LC}(g,x)|^{\delta} \cdot f = q \cdot g + h, \tag{1}$$

where h is the remainder, $\mathrm{LC}(g,x)$ is the leading coefficient of the divisor g, and

$$\delta = \mathrm{degree}(f,x) - \mathrm{degree}(g,x) + 1. \tag{2}$$

It is clear that premultiplying the dividend times $|\mathrm{LC}(g,x)|^{\delta}$ in each iteration of the `while`-loop leads to ever increasing integer coefficients for the polynomial remainders. To reduce the size of the coefficients we could divide them — after each iteration — by their greatest common divisor, and this would result in the smallest possible integer coefficients. Even though this sounds attractive, it is not what we will do because the coefficients obtained this way are *not* determinants of any submatrices of the matrices discussed in section 2.2.

To avoid the extra computation of the gcd of the coefficients in each iteration we will instead divide the coefficients by the *absolute value* of the so called **coefficients-reduction factor** β_i. In other words, we form the prs:

$$\begin{array}{rcl} r_{-1} & = & f, \\ r_0 & = & g, \\ r_1 & = & \dfrac{\texttt{rem\_z}(r_{-1}, r_0, x)}{|\beta_1|}, \\ & \vdots & \\ r_i & = & \dfrac{\texttt{rem\_z}(r_{i-2}, r_{i-1}, x)}{|\beta_i|}, \quad etc, \end{array} \tag{3}$$

where r_i is exactly divided by the coefficients-reduction factor β_i given by

$$\begin{array}{rl} \psi_1 = -1, \; \beta_1 = (-1)^{\delta_1}, & i = 1, \\ \psi_i = \dfrac{(-\mathrm{LC}(r_{i-2},x))^{\delta_{i-1}-1}}{\psi_{i-1}^{\delta_{i-1}-2}}, & i > 1, \\ \beta_i = -\mathrm{LC}(r_{i-2},x) \cdot \psi_i^{\delta_i - 1}, & i > 1, \end{array} \tag{4}$$

and

$$\delta_i = \mathrm{degree}(r_{i-2},x) - \mathrm{degree}(r_{i-1},x) + 1, \qquad i \geqslant 1.$$

The coefficients-reduction factors β_i (4) were discovered — in another context that we will not go into; details can be found elsewhere [6] — about 50 years *after* Anna Johnson Pell published her seminal theorem. We mention them because they speed up the implementation of certain functions — such as `euclid_amv(f, g, x)` — in the module `subresultants_qq_zz.py`.[11]

[11] Functions in `subresultants_qq_zz.py` of the form `function_name_amv(f, g, x)` perform *all* their computations in $\mathbb{Z}[x]$ and make direct or indirect use of the coefficients-reduction factors β_i (4).

The coefficients of the polynomial remainders computed this way *are* determinants of appropriately chosen submatrices of the matrices discussed in the section 2.2.

The process described by equations (3) and (4) above applies mainly to incomplete prs's. For complete prs's we have $\delta_i = 2$ for all i, in which case $\beta_1 = 1$ and $\beta_i = \mathrm{LC}(r_{i-2}, x)^2$ for $i > 1$ — a fact that Sylvester had proved back in 1853 [11]. In other words, Sylvester had discovered the following theorem, whose proof can be found elsewhere [1]:

Theorem 1 (Sylvester, 1853) *Let $r_{-1}, r_0, r_1, \ldots, r_i, \ldots, r_h$ be a complete polynomial remainder sequence, $r_i \in \mathbb{Z}[x]$ for $i = -1, 0, 1, \ldots, h$. Then for $1 < i \leq h-2$ we have $\mathrm{LC}(r_{i-2}, x)^2 \mid r_i$; that is, the square of the leading coefficient of r_{i-2} exactly divides r_i.*

This theorem is hardly mentioned in the literature. Sylvester proved it using determinants, which means he was aware of the relation that exists between prs's and the matrices that he himself had discovered. However, in his article he concedes that he cannot tackle incomplete prs's:

> *the same explicit method might be applied to show, that if the first divisor were e degrees instead of being only one degree lower than the first dividend, α^{e+1} would be contained in every term of the second residue; the difficulty, however, of the proof by this method augments with the value of e* [11, p. 419].

Below we continue with our original example of a complete prs.

Example 1 **(*continued*)** We first compute the Euclidean prs and then the modified Euclidean prs of $f = x^3+5x^2-7x+7$ and $g = 5x^2-6x+8$ over the integers, i.e. in $\mathbb{Z}[x]$, utilizing the function `rem_z(f, g, x)`.

To compute the Euclidean prs of f, g in $\mathbb{Z}[x]$ we use the function `euclid_amv(f, g, x)` and obtain the sequence of polynomials shown below

```
> euclid_amv(f, g, x)
[x**3 + 5*x**2 - 7*x + 7, 5*x**2 - 6*x + 8, -29*x - 73, 1843]
```

Notice that the first remainder $r_1 = -29x - 73$ — which is the third polynomial in the above sequence — is obtained by dividing f by g and then dividing the result by $|\beta_1| = 1$. Likewise, the second remainder $r_2 = 1843$ — which is the constant (last) polynomial in the above sequence — is obtained by dividing g by r_1 and then dividing the result by $|\beta_2| = 25$. Indeed, we have:

```
> r1 = rem_z(f, g, x) / 1; print(r1)
-29*x - 73
> r2 = rem_z(g, r1, x) / 25; print(r2)
1843
```

The constant polynomial $r_2 = 1843$ is a multiple of 1, which is taken as the gcd of f, g.

To compute the modified Euclidean prs of f, g in $\mathbb{Z}[x]$ we employ the function `sturm_amv(f, g, x)` and obtain the sequence of polynomials shown below

```
> sturm_amv(f, g, x)
[x**3 + 5*x**2 - 7*x + 7, 5*x**2 - 6*x + 8, 29*x + 73, -1843]
```

the first remainder $R_1 = 29x + 73$ — which is the third polynomial in the above sequence — is the negation of the remainder obtained by dividing f by g and then dividing the result by $|\beta_1| = 1$. Likewise, the second remainder $R_2 = -1843$ — which is the constant (last) polynomial in the above sequence — is the negation of the remainder obtained by dividing g by R_1 and then dividing the result by $|\beta_2| = 25$. Indeed, we have:

```
> R1 = -rem_z(f, g, x) / 1; print(R1)
29*x + 73

> R2 = -rem_z(g, R1, x) / 25; print(R2)
-1843
```

The prs's $[f, g, r_1, r_2]$ and $[f, g, R_1, R_2]$ in $\mathbb{Z}[x]$ computed above are both complete.

Note that if $g = fp = \frac{d}{dx}(f)$, that is, g is the derivative of f, then the modified Euclidean prs is identical to the Sturm sequence of f, computed in $\mathbb{Z}[x]$. The only way to do this in `sympy` is with the help of the function `sturm_amv(f, g, x)`.

2.2. The two Sylvester Matrices and their Relation to Polynomial Remainder Sequences

Consider the polynomials $f, g \in \mathbb{Z}[x]$ of degrees n, m, respectively, with $n \geq m$. In this section we will present `sylvester1` and `sylvester2`, two matrices discovered by Sylvester in 1840 and 1853, respectively. The elements of both matrices are the coefficients of the polynomials f, g.

Of the two matrices, `sylvester1` is widely known and used (not only because of its smaller dimensions), whereas `sylvester2` is barely known and hardly used. Nonetheless, `sylvester2` contains more information, due to its bigger dimensions, and it is this matrix that was used by Anna Johnson Pell [9].

Sylvester was well aware that `sylvester1` was related to Euclidean prs's and that `sylvester2` was related to modified Euclidean prs's (or at least to Sturm sequences).[12] Below we examine these two cases separately.

2.2.1. Euclidean prs's and Sylvester's Matrix of 1840

For the polynomials $f, g \in \mathbb{Z}[x]$ of degrees n, m, respectively, with $n \geq m$, Sylvester's matrix `sylvester1` of 1840 [10] has dimensions $(n + m) \times (n + m)$ and consists of two groups of rows: the first one with m rows and the second one with n. Concatenation of the two groups yields the matrix `sylvester1`.

In the first row of the first group (of m rows) are the coefficients of f with $m - 1$ trailing zeros. The second row in this group differs from the first one in that its elements have been rotated to the right by one. A total of $m - 1$ rotations are needed to construct the first group of rows.

In the first row of the second group (of n rows) are the coefficients of g with $n - 1$ trailing zeros. The second row in this group differs from the first one in that its elements have been rotated to the right by one. A total of $n - 1$ rotations are needed to construct the second group of rows.

Sylvester used `sylvester1` to compute in $\mathbb{Z}[x]$ the ***resultant*** of the polynomials f, g along with the coefficients of the polynomial remainders obtained by applying Euclid's algorithm on

[12] This was not exactly Sylvester's discovery as we can see in Van Vleck's article [13], but we follow established practice and give Sylvester all the credit.

f, g [10]. Recall that if the resultant of f, g is 0, then these polynomials have a common factor; otherwise, $\gcd(f, g) = 1$.

The resultant of f, g is obtained by evaluating the determinant of `sylvester1`, whereas the coefficients of the polynomial remainders are obtained as determinants of submatrices, or ***subresultants***, of `sylvester1`. The coefficients obtained this way are the *smallest possible* without introducing rationals and without computing (integer) greatest common divisors.

Example 1 (*continued*) To create matrix `sylvester1`, Sylvester's matrix of 1840, for the polynomials $f = x^3 + 5x^2 - 7x + 7$ and $g = 5x^2 - 6x + 8$ we employ the function `sylvester(f, g, x, 1)`. Note the fourth argument "1".

```
>  s1 = sylvester(f, g, x, 1); pprint(s1)
    [1  5   -7  7   0]
    [                ]
    [0  1   5   -7  7]
    [                ]
    [5  -6  8   0   0]
    [                ]
    [0  5   -6  8   0]
    [                ]
    [0  0   5   -6  8]
```

The resultant of f, g is the determinant of s_1 and is equal to $r_2 = 1843$, which was computed by the polynomial division `r2 = rem_z(g, r1, x) / 25` in $\mathbb{Z}[x]$, in section 2.1.2.

```
>  det(s1)
1843
```

The coefficients of $r_1 = -29x - 73$, which was computed in section 2.1.2 by the polynomial division `r1 = rem_z(f, g, x) / 1` in $\mathbb{Z}[x]$ can be also computed from s_1 as follows:

Since r_1 is of degree 1, we delete 1 (the last) row from each group of rows in the `sylvester1` matrix and we are left with the 3 × 5 matrix:

$$\begin{pmatrix} 1 & 5 & -7 & 7 & 0 \\ 5 & -6 & 8 & 0 & 0 \\ 0 & 5 & -6 & 8 & 0 \end{pmatrix}.$$

Then,

$$-29 = \begin{vmatrix} 1 & 5 & -7 \\ 5 & -6 & 8 \\ 0 & 5 & -6 \end{vmatrix},$$

and, after we swap the 3rd and 4th columns of the 3 × 5 matrix

$$-73 = \begin{vmatrix} 1 & 5 & 7 \\ 5 & -6 & 0 \\ 0 & 5 & 8 \end{vmatrix}.$$

In `sympy` the code is[13]

[13] Enumeration in `sympy` begins with 0.

```
>  s1.row_del(4); s1.row_del(1); print(det(s1[:, 0:3]))
```

```
-29
```

```
>  s1.col_swap(2, 3); print(det(s1[:, 0:3]))
```

```
-73
```

Definition 4 *A polynomial remainder sequence, where the coefficients of the remainder polynomials are computed as subresultants of* `sylvester1` *is called* ***subresultant prs***.

Note 2 Instead of using subresultants of `sylvester1` we can obtain the same result using subresultants of Bezout's matrix, which was introduced by Sylvester in 1853 [11], is equivalent to `sylvester2` and has smaller dimensions, namely $n \times n$.

The Bezout matrix for the polynomials f, g of our example is

```
>  b = bezout(f, g, x, 'prs'); pprint(b)
```

```
[5   -6   8 ]
[           ]
[-6  13   5 ]
[           ]
[8   5   -14]
```

To obtain the subresultant prs of f, g we use the function `subresultants_bezout(f, g, x)`, where the signs of the determinants are appropriately adjusted [7]. As Van Vleck informs us, when working with Bezout's matrices,

> *the $n - i + 1$ coefficients of the i-th remainder are the minors obtained from the first i rows by associating those constituents which are found in the first $i - 1$ columns with those of each succeeding column in turn*[14] [13, p. 2].

For complete prs's the output of `subresultants_bezout(f, g, x)` is identical to that of `euclid_amv(f, g, x)`. Indeed, we have

```
>  subresultants_bezout(f, g, x)
```

```
[x**3 + 5*x**2 - 7*x + 7, 5*x**2 - 6*x + 8, -29*x - 73, 1843]
```

```
>  euclid_amv(f, g, x)
```

```
[x**3 + 5*x**2 - 7*x + 7, 5*x**2 - 6*x + 8, -29*x - 73, 1843]
```

2.2.2. Modified Euclidean prs's and Sylvester's Matrix of 1853

For the polynomials $f, g \in \mathbb{Z}[x]$ of degrees n, m, respectively, with $n \geq m$, Sylvester's matrix `sylvester2` of 1853 [11] has dimensions $2n \times 2n$ and consists of n *pairs* of rows.

In the first row of the first pair are the coefficients of f whereas in the second row of the first pair are the coefficients of g; $n - m$ zeros have been prepended to g to also make it of degree n. Both rows in the first pair have $2n - (n + 1)$ trailing zeros and both rows of the last pair have $2n - (n + 1)$ leading zeros. The second pair of rows differs from the first one in that the elements

[14] Van Vleck's enumeration of the remainders differs from our own.

of both rows have been rotated to the right by one. A total of $2n-(n+1)$ rotations are needed to construct `sylvester2`.

Sylvester used `sylvester2` to compute in $\mathbb{Z}[x]$ the ***modified resultant***[15] of the polynomials f, g along with the coefficients of the polynomial remainders obtained by applying the *modified* Euclidean algorithm on f, g [11].

The modified resultant of f, g is obtained by evaluating the determinant of `sylvester2`, whereas the coefficients of the polynomial remainders are obtained as determinants of submatrices, or ***modified subresultants***, of `sylvester2`. The coefficients obtained this way are the *smallest possible* without introducing rationals and without computing (integer) greatest common divisors.

Example 1 (*continued*) To create matrix `sylvester2`, Sylvester's matrix of 1853, for the polynomials $f = x^3 + 5x^2 - 7x + 7$ and $g = 5x^2 - 6x + 8$ we employ the function `sylvester(f, g, x, 2)`. Note the fourth argument “2”.

```
>  s2 = sylvester(f, g, x, 2); pprint(s2)
```

```
[1  5  -7  7   0   0]
[                   ]
[0  5  -6  8   0   0]
[                   ]
[0  1  5   -7  7   0]
[                   ]
[0  0  5   -6  8   0]
[                   ]
[0  0  1   5   -7  7]
[                   ]
[0  0  0   5   -6  8]
```

The modified resultant of f, g is the determinant of s_2 and is equal to $R_2 = -1843$, which was computed by the polynomial division `R2 = -rem_z(g, R1, x) / 25` in $\mathbb{Z}[x]$, in section 2.1.2.

```
>  det(s2)
```

```
-1843
```

The coefficients of $R_1 = 29x + 73$, which was computed in section 2.1.2 by the polynomial division `R1 = - rem_z(f, g, x) / 1` in $\mathbb{Z}[x]$ can be also computed from s_2 as follows:

Since R_1 is of degree 1, we delete 1 (the last) pair from the `sylvester2` matrix and we are left with the 4×6 matrix:

$$\begin{pmatrix} 1 & 5 & -7 & 7 & 0 & 0 \\ 0 & 5 & -6 & 8 & 0 & 0 \\ 0 & 1 & 5 & -7 & 7 & 0 \\ 0 & 0 & 5 & -6 & 8 & 0 \end{pmatrix}.$$

[15] It may differ from the resultant of f, g in sign and/or by an integer factor.

Then,

$$29 = \begin{vmatrix} 1 & 5 & -7 & 7 \\ 0 & 5 & -6 & 8 \\ 0 & 1 & 5 & -7 \\ 0 & 0 & 5 & -6 \end{vmatrix},$$

and, after we swap the 4th and 5th columns of the 4 × 6 matrix,

$$73 = \begin{vmatrix} 1 & 5 & -7 & 0 \\ 0 & 5 & -6 & 0 \\ 0 & 1 & 5 & 7 \\ 0 & 0 & 5 & 8 \end{vmatrix}.$$

In sympy the code is[13]

```
>  s2.row_del(4); s2.row_del(4); print(det(s2[:, 0:4]))
29
```

```
>  s2.col_swap(3, 4); print(det(s2[:, 0:4]))
73
```

Definition 5 *A polynomial remainder sequence, where the coefficients of the remainder polynomials are computed as subresultants of* `sylvester2` *is called* ***modified subresultant prs****.*[16]

Note 3 Instead of using subresultants of `sylvester2` we can obtain the same result using subresultants of Bezout's matrix, which is equivalent to `sylvester2` and has smaller dimensions, namely $n \times n$.

The function `modified_subresultants_bezout(f, g, x)` computes the modified subresultant prs of f, g.

For complete prs's the output of `modified_subresultants_bezout(f, g, x)` is identical to that of `sturm_amv(f, g, x)`. Indeed, we have

```
>  modified_subresultants_bezout(f, g, x)
[x**3 + 5*x**2 - 7*x + 7, 5*x**2 - 6*x + 8, 29*x + 73, -1843]
```

```
>  sturm_amv(f, g, x)
[x**3 + 5*x**2 - 7*x + 7, 5*x**2 - 6*x + 8, 29*x + 73, -1843]
```

2.3. Incomplete Polynomial Remainder Sequences

So far, the polynomials f, g that we examined, had complete prs's, which were thoroughly investigated in the 19th century. As we saw, the polynomials in the Euclidean prs of f, g, computed in $\mathbb{Z}[x]$, are identicall to those in the subresultant prs of f, g and, likewise, the polynomials in the modified Euclidean prs of f, g, computed in $\mathbb{Z}[x]$, are identicall to those in the modified subresultant prs of f, g. That is, in this case, we *can* compute one sequence from the other.

[16] Modified subresultants may differ from subresultants in sign and/or by the factor of $\mathrm{LC}(f,x)^{\delta_1}$, where δ_1 was defined elsewhere (4). If $\mathrm{LC}(f,x) = 1$, this factor obviousy does not appear; otherwise, the factor $\mathrm{LC}(f,x)^{\delta_1}$ may be divided out of the modified subresultants for smaller coefficients.

However, things become very complicated when f, g have incomplete prs's. The polynomials in the Euclidean prs of f, g, computed in $\mathbb{Z}[x]$, may differ in sign from those in the subresultant prs of f, g and, likewise, the polynomials in the modified Euclidean prs of f, g, computed in $\mathbb{Z}[x]$, may differ in sign from those in the modified subresultant prs of f, g. Therefore, in this case we *cannot* compute one sequence from the other.

For our discussion we need the following definition:

Definition 6 *The* ***sign sequence*** *of a polynomial remainder sequence (prs) is the sequence of signs of the leading coefficients of its polynomials.*

The sign sequence of a random prs `sq` is computed with the function `sign_seq(sq, x)`. The difficulties with incomplete prs's are demonstrated by the following example.

Example 2 Consider the polynomials $f = x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5$ and $g = 3x^6 + 5x^4 - 4x^2 - 9x + 21$, whose prs is "very" incomplete since its degree sequence is $[8, 6, 4, 2, 1, 0]$.

```
>  f = x**8 + x** 6 - 3*x** 4 - 3*x** 3 + 8*x **2 + 2*x - 5;
    g = 3*x** 6 + 5*x** 4 - 4*x** 2 - 9*x + 21
```

Notice that the polynomials in `eprs`, the Euclidean prs of f, g computed in $\mathbb{Z}[x]$, differ in sign from those in `sprs`, the subresultant prs. However, the absolute values of the coefficients are the same in both sequences. Therefore, it is not at all obvious how to compute the correct signs of one sequence from those of the other.

```
>  eprs = euclid_amv(f, g, x); print(eprs)
```

```
[x**8 + x**6 - 3*x**4 - 3*x**3 + 8*x**2 + 2*x - 5, 3*x**6 + 5*x**4 - 4*x**2
- 9*x + 21, -15*x**4 + 3*x**2 - 9, -65*x**2 - 125*x + 245, 9326*x - 12300,
-260708]
```

```
>  sign_seq(eprs, x)
```

```
[1, 1, -1, -1, 1, -1]
```

```
>  sprs = subresultants_bezout(f, g, x); print(sprs)
```

```
[x**8 + x**6 - 3*x**4 - 3*x**3 + 8*x**2 + 2*x - 5, 3*x**6 + 5*x**4 - 4*x**2
- 9*x + 21, 15*x**4 - 3*x**2 + 9, 65*x**2 + 125*x - 245, 9326*x - 12300,
260708]
```

```
>  sign_seq(sprs, x)
```

```
[1, 1, 1, 1, 1, 1]
```

Likewise, the polynomials in `meprs`, the modified Euclidean prs of f, g computed in $\mathbb{Z}[x]$, differ in sign from those in `msprs`, the modified subresultant prs. However, the absolute values of the coefficients are the same in both sequences. Again, it is not at all obvious how to compute the correct signs of one sequence from those of the other.

```
>  meprs = sturm_amv(f, g, x); print(meprs)
```

```
[x**8 + x**6 - 3*x**4 - 3*x**3 + 8*x**2 + 2*x - 5, 3*x**6 + 5*x**4 - 4*x**2
- 9*x + 21, 15*x**4 - 3*x**2 + 9, 65*x**2 + 125*x - 245, 9326*x - 12300,
-260708]
```

```
>  sign_seq(meprs, x)
[1, 1, 1, 1, 1, -1]
```

```
>  msprs = modified_subresultants_bezout(f, g, x); print(msprs)
[x**8 + x**6 - 3*x**4 - 3*x**3 + 8*x**2 + 2*x - 5, 3*x**6 + 5*x**4 - 4*x**2
- 9*x + 21, -15*x**4 + 3*x**2 - 9, 65*x**2 + 125*x - 245, -9326*x + 12300,
260708]
```

```
>  sign_seq(msprs, x)
[1, 1, -1, 1, -1, 1]
```

Anna Johnson Pell was faced with the much more difficult and daunting problem of having to compute the modified Euclidean prs of f, g, ***computed in*** $\mathbb{Q}[x]$, from the modified subresultant prs and vice-versa. In this case the sequences to be related are `meprs_q` and `msprs` shown below:

Example 2 ***(continued)*** For the same polynomials $f = x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5$ and $g = 3x^6 + 5x^4 - 4x^2 - 9x + 21$, whose prs is "very" incomplete, we have:

```
>  meprs_q = sturm_q(f, g, x); print(meprs_q)
[x**8 + x**6 - 3*x**4 - 3*x**3 + 8*x**2 + 2*x - 5, 3*x**6 + 5*x**4 - 4*x**2
- 9*x + 21, 5*x**4/9 - x**2/9 + 1/3, 117*x**2/25 + 9*x - 441/25, 233150*x/19773
- 102500/6591, -1288744821/543589225]
```

```
>  sign_seq(meprs_q, x)
[1, 1, 1, 1, 1, -1]
```

```
>  msprs = modified_subresultants_bezout(f, g, x); print(msprs)
[x**8 + x**6 - 3*x**4 - 3*x**3 + 8*x**2 + 2*x - 5, 3*x**6 + 5*x**4 - 4*x**2
- 9*x + 21, -15*x**4 + 3*x**2 - 9, 65*x**2 + 125*x - 245, -9326*x + 12300,
260708]
```

```
>  sign_seq(msprs, x)
[1, 1, -1, 1, -1, 1]
```

Notice that `sign_seq(meprs_q, x)=sign_seq(meprs, x)`, where the sequence on the right hand of the equation was computed in the first part of the example.

The impatient reader can jump to section 2.4.2 and find out how the two sequences `meprs_q` and `msprs` were coupled together. In section 2.4.1 we set the stage for the problem as encountered by Anna Johnson Pell and her co-author Ruth L. Gordon.

2.4. The Pell-Gordon Theorem of 1917 and its Ramifications

The article of 1917 [9] was written by Anna Johnson Pell and a certain Ruth L. Gordon, for whom no information whatsoever can be found. However, the order of the names in the article indicate that Anna was the driving force behind the theorem, which is referred to as "Pell-Gordon."

The Pell-Gordon theorem of 1917 gave a solution to the problem of incomplete prs's that Van Vleck — more on him in section 2.4.1 — was not able to tackle; details in [2]. By the way, this was the same problem that Sylvester had alluded to back in 1853.

In section 2.4.1 we present Van Vleck's method for computing Sturm sequences and thereby set the stage for the Pell-Gordon theorem, which is presented in section 2.4.2.

Finally, in section 2.4.3 we present the various ramifications of this remarkable theorem.

2.4.1. Van Vleck's Method for Computing Sturm Sequences

Up until 1827 Euclidean prs's were the only sequences of interest to mathematicians. However, in 1827 Sturm discovered his theorem for isolating by bisection the real roots of a polynomial f with rational coefficients. Sturm's theorem relies on the so called Sturm's sequence, which — as we have already mentioned in section 2.1 — is a special case of the modified Euclidean prs.

Therefore, after 1827 modified Euclidean prs's — and in particular Sturm's sequences — were of greater interest to mathematicians.

Van Vleck, at his time a renowned mathematician at the University of Wisconsin-Madison,[17] USA, published an article in 1899-1900 [13] in which he developed a method for computing the Sturm sequence of a polynomial f by triangularizing `sylvester2`, Sylvester's matrix of 1853.

Van Vleck's triangularization method is amply described elsewhere [2] and we will not dwell on it here. Instead we simply present without proof two of his theorems that form the backbone of his method.

The first theorem was first presented by Sylvester as a way to exactly compute the coefficients of the polynomial remainders in *complete* Sturm sequences as modified subresultants of `sylvester2`. This was reiterated by Van Vleck about 50 years later:

Theorem 2 (Sylvester 1853, Van Vleck 1899-1900) *Consider the polynomials $f = c_n x^n + \ldots + c_0$ and $g = d_m x^m + \ldots + d_0$, in $\mathbb{Z}[x]$, with $c_n \neq 0, d_m \neq 0, n \geqslant m$. Then the successive polynomials that are formed from the first $2j$ rows, $j = 2, \ldots, n$, of Sylvester's matrix (`sylvester2`) for f, g, constitute a Sturm sequence.*

Notice that the theorem makes *no* reference to the Sturm sequence being complete, but clearly this is what both Sylvester and Van Vleck had in mind. Recall that Sylvester himself was aware of the difficulties involved with incomplete sequences, but he did not attempt to tackle the problem. Van Vleck, on the other hand, solved the problem by assuming all Sturm sequences complete:

> *Beginning with this polynomial and remainder, the degree of each succeeding polynomial, respectively remainder is, in general, one less than that of the preceding* [13, p. 4].

Van Vleck realized that one does not have to compute modified subresultants of Sylvester's matrix `sylvester2` in order to find the coefficients of the polynomial remainders in the Sturm sequence. It suffices to simply triangularize `sylvester2` using integer preserving transformations, in which case the modified subresultants (the coefficients) can be read off the triangularized matrix. We have the following [13, p. 8]:

[17] The building of the Mathematics Department in Madison is named after Van Vleck.

Theorem 3 (Van Vleck 1899-1900) *Let f and $g = f'$ be two polynomials of degree n and $n-1$, respectively, and let S_2 be their Sylvester matrix* `sylvester2`*. If, using integer preserving transformations, we bring S_2 into its upper triangular form, $T(S_2)$, then the even rows of $T(S_2)$ furnish the coefficients of the successive polynomial remainders of the Sturm sequence. The coefficients taken from a given row are multiplied times $(-1)^k$, where* k *is the number of negative elements on the principle diagonal above the row under consideration.*

Van Vleck takes advantage of the special form of `sylvester2` and computes $T(S_2)$ by updating only two rows at a time; to update these two rows he triangularizes a matrix of only three rows, a fact that makes his procedure extremely efficient. To keep the coefficients small he removes at each step the greatest common divisor (content) of the elements in both updated rows, and uses those reduced coefficients in the next three-row matrix.

Van Vleck's computation is justified by the fact that in `sylvester2` the elements (entries) of any two consecutive rows are the same as those of the two preceding rows.

Therefore, if in any row the values of the elements are changed by adding a multiple of the preceding row, exactly the same change can be made in the elements of each alternate row thereafter, without altering the value of any modified subresultant that appears as a coefficient in one of the polynomials of the Sturm sequence.

In the following example we show how to compute the Sturm sequence of a polynomial f from the corresponding Euclidean prs of f, f'. This relation dates back to an observation by Sylvester [12] that was recently proved by Akritas and Malaschonok — hence the label $\mathscr{SAM}$ in Figure 1 of section 2.4.3.

Example 3 In this example we take the polynomial $f = x^3 + 5x^2 - 7x + 7$ and its derivative and use the function `subresultants_vv(f, g, x)`[18] to compute their subresultant prs which — since their prs is complete — is the same as their Euclidean prs.

```
>  f = x**3 - 5*x**2 - 7*x + 7; g = diff(f, x, 1)

>  subresultants_vv(f, g, x, 1)
     [1   -5   -7     7     0      0  ]
     [                                ]
     [0   3    -10   -7     0      0  ]
     [                                ]
     [0   0     3    -10   -7      0  ]
     [                                ]
     [0   0     0    -92   28      0  ]
     [                                ]
     [0   0     0     0    -92    28  ]
     [                                ]
     [0   0     0     0     0    -9184]
```

According to Theorem 3 the coefficients of the first remainder are in the fourth row and the constant term is in the sixth row. Indeed, applying the function `euclid_amv(f, g, x)` we see that by triangularizing the matrix `sylvester2` of f, g we have obtained the Euclidean prs of f, g in $\mathbb{Z}[x]$.

[18] A modification of Van Vleck's original algorithm, which makes use of the Pell-Gordon theorem in section 2.4.2.

```
>  euclid_amv(f, g, x)
[x**3 - 5*x**2 - 7*x + 7, 3*x**2 - 10*x - 7, -92*x + 28, -9184]
```

Sturm's sequence in $\mathbb{Z}[x]$ differs only in signs from the Euclidean prs computed in $\mathbb{Z}[x]$. The signs in both sequences follow the pattern [4, Theorem 3]

$$+,+,-,-,+,+,-,-,\dots$$

and, therefore, the Sturm sequence of f is

$$[x^3-5x^2-7x+7,\ 3x^2-10x-7,\ 92x-28,\ 9184].$$

2.4.2. A Seminal Theorem

Anna Johnson Pell was well aware of the deficiencies of the Van Vleck approach and in their article of 1917 Pell and Gordon explicitly mention that his method needs to be modified in case of incomplete prs's.

Theorem 4 below, [9], helps us compute the coefficients of the modified Euclidean prs of f,g — be it complete or incomplete — from the corresponding modified subresultants prs of f,g,[19] and vice versa.[20] The theorem is stated below

Theorem 4 (Pell-Gordon, 1917) *Let*

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

and

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

be two polynomials of the n-th degree. Modify the process of finding the highest common factor of f and g 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 $\varrho_i=r_{p_i-1}^{(i)}$ is different from zero. Then for $k=0,1,\dots,m_i$ the coefficients $r_k^{(i)}$ are given by[21]

$$r_k^{(i)}=\frac{(-1)^{u_{i-1}}(-1)^{u_{i-2}}\cdots(-1)^{u_1}(-1)^{v_{i-1}}}{\varrho_{i-1}^{p_{i-1}+1}\varrho_{i-2}^{p_{i-2}+p_{i-1}}\cdots\varrho_1^{p_1+p_2}\varrho_0^{p_1}}\cdot\operatorname{Det}(i,k), \tag{5}$$

where

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

and

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

[19] *Without* polynomial divisions, that is, with the help of determinants of submatrices of the `sylvester2` matrix of f,g.

[20] *Without* determinant evaluations, that is, with the help of the coefficients of the modified Euclidean prs of f,g.

[21] It is understood in (5) that $\varrho_0=b_0$, $p_0=0$, and that $a_i=b_i=0$ for $i>n$.

Proof The proof is by structural induction[22] on the the polynomials in the prs, but it is too lengthy to be presented here. For details see [9]. □

This way, by “modifying” Van Vleck’s theorem, Pell and Gordon solved the problem of computing the coefficients of the modified Euclidean prs of f, g via the corresponding modified subresultants prs of f, g, and vice versa. Their paper went unnoticed for about 100 years, until it was discovered by us in the journal archives.

Note that the first fraction in formula (5) depends only on i and is independent of k. Denote by $PG^{(i)}$ that fraction and call it the $PG^{(i)}$***-factor*** [4]; that is, we have

$$PG^{(i)} = \frac{(-1)^{u_{i-1}}(-1)^{u_{i-2}}\cdots(-1)^{u_1}(-1)^{u_0}(-1)^{v_{i-1}}}{\varrho_{i-1}^{p_{i-1}+1}\varrho_{i-2}^{p_{i-2}+p_{i-1}}\cdots\varrho_1^{p_1+p_2}\varrho_0^{p_1}}, \tag{6}$$

in which case, the coefficients of the polynomials in the modified Euclidean prs are exactly

$$r_k^{(i)} = PG^{(i)} \times \mathrm{Det}\,(i,k). \tag{7}$$

Example 4 Consider again the polynomials $f = x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5$ and $g = 3x^6 + 5x^4 - 4x^2 - 9x + 21$, last encountered in the second part of Example 2. Recall that `msprs`, the modified subresultant prs of f, g, is

$$\begin{aligned} &x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5,\ \ 3x^6 + 5x^4 - 4x^2 - 9x + 21, \\ &\qquad -15x^4 + 3x^2 - 9,\ \ 65x^2 + 125x - 245,\ \ -9326x + 12300,\ \ 260708, \end{aligned} \tag{8}$$

where the coefficients of the last 4 polynomials in the second line of equation (8) are all determinants (modified subresultants $\mathrm{Det}\,(i,k)$) of appropriate submatrices of `sylvester2`.

To compute the coefficients of the polynomials in `meprs_q`, the modified Euclidean prs in $\mathbb{Q}[x]$, we have to compute the $PG^{(i)}$*-factor*, $i = 1, 2, 3, 4$, for each remainder. Using (6) we find

$$PG^{(i)} = \left\{-\frac{1}{27}, \frac{9}{125}, -\frac{25}{19773}, -\frac{19773}{2174356900}\right\}, \quad i = 1, 2, 3, 4, \tag{9}$$

and from (7), we obtain

$$\begin{aligned} &x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5,\ \ 3x^6 + 5x^4 - 4x^2 - 9x + 21, \\ &\qquad 5x^4/9 - x^2/9 + 1/3,\ \ 117x^2/25 + 9x - 441/25, \\ &\qquad\quad 233150x/19773 - 102500/6591,\ \ -1288744821/543589225. \end{aligned} \tag{10}$$

Theorem 4 is the main pillar of the theory of subresultants. Using Sylvester’s observation of 1853 [12] — see Example 3 of section 2.4.1 and the arrow labeled $\mathscr{SAM}$ in Figure 1 of section 2.4.3 — it establishes a one-to-one correspondence between modified subresultant prs’s, on one hand, and modified Euclidean and Euclidean prs’s on the other. This one-to-one correspondence unequivocally refutes the claim — see Note 1 in section 2.1 — that Euclidean prs’s are “non signed” sequences and that the signs of their polynomials can be changed arbitrarily.

Using Theorem 4 we have developed four functions in the module `subresultants_qq_zz.py` with which we can compute the Euclidean and modified Euclidean prs as well as the subresultant and modified subresultant prs of two polynomials f, g; these functions are `euclid_pg(f, g, x)`, `subresultants_pg(f, g, x)`, `sturm_pg(f, g, x)`,

[22] It seems to be the first such proof in the literature.

and `modified_subresultants_pg(f, g, x)`. However, they all perform their operations in $\mathbb{Q}[x]$, as a result of which these functions are slower than the equivalent ones of the form `function_name_amv(f, g, x)`, which perform all their operations in $\mathbb{Z}[x]$.

Example 4 (*continued*) Consider again the polynomials $f = x^8+x^6-3x^4-3x^3+8x^2+2x-5$ and $g = 3x^6+5x^4-4x^2-9x+21$, which form an incomplete prs. Using Theorem 4, their Euclidean and subresultant prs's are:

```
>  eprs = euclid_pg(f, g, x); print(eprs)

[x**8 + x**6 - 3*x**4 - 3*x**3 + 8*x**2 + 2*x - 5, 3*x**6 + 5*x**4 - 4*x**2
- 9*x + 21, -15*x**4 + 3*x**2 - 9, -65*x**2 - 125*x + 245, 9326*x - 12300,
-260708]

>  sign_seq(eprs, x)

[1, 1, -1, -1, 1, -1]

>  sprs = subresultants_pg(f, g, x); print(sprs)

[x**8 + x**6 - 3*x**4 - 3*x**3 + 8*x**2 + 2*x - 5, 3*x**6 + 5*x**4 - 4*x**2
- 9*x + 21, 15*x**4 - 3*x**2 + 9, 65*x**2 + 125*x - 245, 9326*x - 12300,
260708]

>  sign_seq(sprs, x)

[1, 1, 1, 1, 1, 1]
```

As we see the sign sequences of the two prs's differ. Likewise their modified Euclidean and modified subresultant prs's are

```
>  meprs = sturm_pg(f, g, x); print(meprs)

[x**8 + x**6 - 3*x**4 - 3*x**3 + 8*x**2 + 2*x - 5, 3*x**6 + 5*x**4 - 4*x**2
- 9*x + 21, 15*x**4 - 3*x**2 + 9, 65*x**2 + 125*x - 245, 9326*x - 12300,
-260708]

>  sign_seq(meprs, x)

[1, 1, 1, 1, 1, -1]

>  msprs = modified_subresultants_pg(f, g, x); print(msprs)

[x**8 + x**6 - 3*x**4 - 3*x**3 + 8*x**2 + 2*x - 5, 3*x**6 + 5*x**4 - 4*x**2
- 9*x + 21, -15*x**4 + 3*x**2 - 9, 65*x**2 + 125*x - 245, -9326*x + 12300,
260708]

>  sign_seq(msprs, x)

[1, 1, -1, 1, -1, 1]
```

Again, the sign sequences of the two prs's differ.

2.4.3. Ramifications of the Pell-Gordon Theorem

Theorem 4 led us to the discovery of the following theorem [5], which establishes a one-to-one correspondence between subresultant prs's, on one hand, and Euclidean and modified Euclidean prs's on the other. This one-to-one correspondence unequivocally refutes, a second time, the claim — see Note 1 in section 2.1 — that Euclidean prs's are "non signed" sequences and that the signs of their polynomials can be changed arbitrarily. The complete picture is given by Figure 1.

Theorem 5 (Akritas-Malaschonok-Vigklas, 2015) *Let*

$$\begin{aligned} f &= a_0x^n + a_1x^{n-1} + \cdots + a_n, \\ g &= b_0x^n + b_1x^{n-1} + \cdots + b_n \end{aligned}$$

be two polynomials of degree n *and* $n - p_0$*, respectively, with* $b_0 = b_1 = \ldots = b_{p_0-1} = 0$, $b_{p_0} \neq 0$, $p_0 \geq 0$*. Moreover, for* $i = 1, 2, \ldots$*, let*

$$\begin{aligned} R^{(i)} &= r_0^{(i)}x^{m_i} + r_1^{(i)}x^{m_i-1} + \cdots + r_{m_i}^{(i)}, \\ R^{E(i)} &= r_0^{E(i)}x^{m_i} + r_1^{E(i)}x^{m_i-1} + \cdots + r_{m_i}^{E(i)}, \end{aligned} \tag{11}$$

be the i-th modified Euclidean and Euclidean remainders, respectively, of f, g*, with* $R^{(i)}$ *and* $R^{E(i)}$ *both of degree* $m_i - p_i + 1$*, where* $(m_i + 1)$ *is the degree of the preceding remainder and*

$$r_0^{(i)} = r_0^{E(i)} = \ldots = r_{p_i-2}^{(i)} = r_{p_i-2}^{E(i)} = 0,\ \varrho_i = r_{p_i-1}^{(i)} \neq 0,\ \sigma_i = r_{p_i-1}^{E(i)} \neq 0.$$

Then for $k = 0, 1, \ldots, m_i$ *the coefficients* $r_k^{(i)}$ *and* $r_k^{E(i)}$ *in (11) are given by*

$$r_k^{(i)} = \frac{(-1)^{\varphi_i}}{\varrho_{i-1}^{p_{i-1}+1}\varrho_{i-2}^{p_{i-2}+p_{i-1}}\cdots\varrho_0^{p_0+p_1}} \times \frac{\mathrm{Det}_{i,k}(f,g)}{a_0^{p_0}}, \tag{12}$$

$$r_k^{E(i)} = \frac{(-1)^{\psi_i}}{\sigma_{i-1}^{p_{i-1}+1}\sigma_{i-2}^{p_{i-2}+p_{i-1}}\cdots\sigma_0^{p_0+p_1}} \times \frac{\mathrm{Det}_{i,k}(f,g)}{a_0^{p_0}}, \tag{13}$$

where $\varrho_0 = \sigma_0 = b_{p_0}$*,*

$$\varphi_i = \lfloor (s_{i-1}+1)/2 \rfloor,$$

$$s_{i-1} = \textit{the number of odd integers in the list } \{p_0, p_1, \ldots, p_{i-1}\},$$

$$\psi_i = i + \varphi_i + p_1 + p_3 + p_5 + .. + p_{2\lfloor i/2 \rfloor - 1},\ \textit{with } p_{-1} = 0,$$

$$\mathrm{Det}_{i,k}(f,g) = \begin{vmatrix} a_0 & a_1 & \cdots & a_{p_0} & \cdots & a_{v_{i-1}} & \cdots & a_{2v_{i-1}} & a_{2v_{i-1}+k+1} \\ 0 & a_0 & \cdots & a_{p_0-1} & \cdots & a_{v_{i-1}-1} & \cdots & a_{2v_{i-1}-1} & a_{2v_{i-1}+k} \\ \vdots & & \ddots & & \cdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & a_0 & \cdots & a_{v_{i-1}-p_0} & \cdots & a_{2v_{i-1}-p_0} & a_{2v_{i-1}+k+1-p_0} \\ \vdots & & \ddots & & \ddots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & a_0 & \cdots & a_{v_{i-1}} & a_{v_{i-1}+k+1} \\ b_0 & b_1 & \cdots & b_{p_0} & \cdots & b_{v_{i-1}} & \cdots & b_{2v_{i-1}} & b_{2v_{i-1}+k+1} \\ 0 & b_0 & \cdots & b_{p_0-1} & \cdots & b_{v_{i-1}-1} & \cdots & b_{2v_{i-1}-1} & b_{2v_{i-1}+k} \\ \vdots & & \ddots & & \cdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & b_0 & \cdots & a_{v_{i-1}-p_0} & \cdots & a_{2v_{i-1}-p_0} & a_{2v_{i-1}+k+1-p_0} \\ \vdots & & \ddots & & \ddots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & b_0 & \cdots & b_{v_{i-1}} & b_{v_{i-1}+k+1} \end{vmatrix},$$

and

$$v_{i-1} = p_0 + p_1 + \cdots + p_{i-1}.$$

Proof By structural induction on the polynomials in the prs. Compared to Theorem 4, our proof has been considerably simplified. □

Using Theorem 5 we have developed four functions in the module `subresultants_qq_zz.py` with which we can compute the Euclidean and modified Euclidean prs as well as the subresultant and modified subresultant prs of two polynomials f, g; these functions are `euclid_amv(f, g, x)`, `subresultants_amv(f, g, x)`, `sturm_amv(f, g, x)`, and `modified_subresultants_amv(f, g, x)` and have already been used in the examples of this article.

Figure 1 below, summarizes everything we talked about.

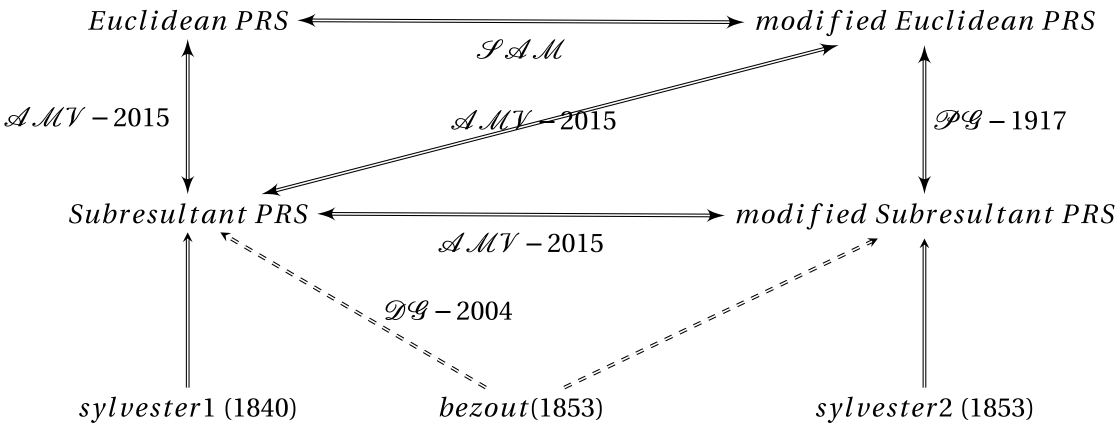

Figure 1. The double ended arrows indicate one-to-one correspondences that exist between the coefficients of the polynomials in the respective nodes. The labels indicate those who first established the correspondences and when. The dashed arrow labeled $\mathscr{DG} - 2004$ is due to Diaz–Toca and Gonzalez–Vega [7]

References

1. Akritas, A.G.: “A simple proof of the validity of the reduced prs algorithm.” *Computing*, **38** (1987), 369-372.
2. Akritas, A.G., G.I. Malaschonok, P.S. Vigklas: “On a Theorem by Van Vleck Regarding Sturm Sequences.” *Serdica Journal of Computing*, 7(4) (2013), 101–134.
3. Akritas, A.G., G.I. Malaschonok, P.S. Vigklas: “Sturm Sequences and Modified Subresultant Polynomial Remainder Sequences.” *Serdica Journal of Computing*, **8**(1) (2014), 29–46.
4. Akritas, A.G., G.I. Malaschonok, P.S. Vigklas: “On the Remainders Obtained in Finding the Greatest Common Divisor of Two Polynomials.” *Serdica Journal of Computing*, **9**(2) (2015), 123–138.
5. Akritas, A.G., G.I. Malaschonok, P.S. Vigklas: “A Basic Result on the Theory of Subresultants.” Preprint submitted for publication.
6. Cohen, J.E.:*Computer Algebra and Symbolic Computation – Mathematical Methods*. A.K. Peters, Massachusetts, (2003).
7. Diaz–Toca G. M., L. Gonzalez–Vega: “Various New Expressions for Subresultants and Their Applications”. *Applicable Algebra in Engineering, Communication and Computing*, **15**, (2004), 233–266.
8. Greenstein, L.S., P.J. Campbell: “Anna Johnson Pell Wheeler: Her Life and Work.” *Historia Mathematica* **9** (1982), 37–53.

9. Pell, A.J., R.L. Gordon: “The Modified Remainders Obtained in Finding the Highest Common Factor of Two Polynomials.” *Annals of Mathematics*, Second Series **18**(4) (1917), 188–193.
10. Sylvester J.J.: “A method of determining by mere inspection the derivatives from two equations of any degree”. *Philosophical Magazine*, **16** (1840), 132–135.
11. Sylvester, J.J.: “On the Theory of 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.” *Philosophical Transactions*, **143**, (1853), 407–548.
12. Sylvester J. J.: “On a remarkable modification of Sturm’s theorem.” *Philosophical Magazine and Journal of Science*, **V**, Fourth Series, (January–June, 1853), 446–456. http://books.google.gr/books?hl=el&id=3Ov22-gFMnEC&q=sylvester#v=onepage&q&f=false
13. Van Vleck, E.B.: “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), (1899–1900), 1–13.

АННА ДЖОНСОН И ЕЕ ГЕНИАЛЬНАЯ ТЕОРЕМА 1917 ГОДА

Алкивиадис Акритас

Аннотация

В этой статье мы представляем жизнь Анны Джонсон, женщины исключительно одаренной в области математики, наряду с тем, что мы считаем ее наибольшим вкладом: а именно, теоремой 1917 на модифицированных евклидовых последовательностях полиномиальных остатков (ППО), которая заложила основы теории субрезультантных ППО.

Для того, чтобы продемонстрировать различные математические понятия, представленные в этой статье, мы используем систему компьютерной алгебры SumPy (версия 1.0), которая основана на Python и находится в свободном доступе.

Ключевые слова: *евклидова последовательность полиномиальных остатков (ППО), модифицированная евклидова ППО, субрезультантная ППО, модифицированная субрезультантная ППО, метод Ван Флека, ППО Штурма.*

Alkiviadis Akritas,
University of Thessaly
Department of Electrical and
Computer Engineering
Volos, Greece,
akritas@uth.gr