Euclid's Algorithm over Rings – The Theory of Subresultants

Euclid's algorithm, as described in the previous lecture, works well when the coefficients of the input polynomials A, B, belong to a field F. There are many instances, however, when the coefficients do not come from a field but from a ring. The most standard examples being the ring of integers $\mathbb{Z}$ or the ring of multivariate polynomials $\mathbb{R}[x_1, \ldots, x_n]$. How do we implement Euclid's algorithm in these cases? Even more fundamentally, is the gcd of two polynomials well defined for such rings? Suppose the answer to the latter question is yes, then one way to compute the gcd is to do Euclid's algorithm in the quotient field w.r.t. the ring of coefficients, i.e., work in $\mathbb{Q}[x]$ when the input are integer polynomials, and $\mathbb{R}(x_1, \ldots, x_n)$ the field of rational functions when the input is a multivariate polynomial. There seems to be no problem. Let us try this approach for the following polynomials:

$$\begin{aligned} A_0 &:= x^8 + x^6 - 3\,x^4 - 3\,x^3 + 8\,x^2 + 2\,x - 5 \\ A_1 &:= 2\,x^6 + 5\,x^4 - 4\,x^2 - 9\,x + 21. \end{aligned}$$

We get the following remainder sequence

$$\begin{aligned} A_2 &= 11/4\,x^4 + 3/2\,x^3 - 11/2\,x^2 - \frac{19}{4}\,x + \frac{43}{4} \\ A_3 &= -\frac{5272}{1331}\,x^3 + \frac{664}{121}\,x^2 + \frac{15756}{1331}\,x - \frac{21972}{1331} \\ A_4 &= \frac{35003969}{3474248}\,x^2 - \frac{296813}{868562}\,x - \frac{39537355}{3474248} \\ A_5 &= \frac{6946037969540}{920569380731}\,x - \frac{740798266924}{70813029287} \\ A_6 &= \frac{208678923016350199413}{27774323234433100900}. \end{aligned} \tag{1}$$

As the example above suggests, the coefficients seem to grow exponentially. Let's try to understand why. Suppose $A_0(x) = \sum_{i=0}^{n} a_i x^i$ and $A_1(x) = \sum_{i=0}^{n-1} b_i x^i$ are integer polynomial with coefficients of bit-sizes L_0, L_1 resp. Then the ith coefficient in $A_2 := \texttt{rem}(A_0, A_1)$ is $(a_i - b_{i-1}a_n/b_{n-1})$. Thus the bit-sizes of the rationals involved in A_2 satisfy the recurrence $L_2 = L_0 + L_1$. If the degrees fall by exactly one at each euclidean step, then we see that the bit-sizes of the kth polynomial in the remainder sequence is roughly the kth Fibonacci number, i.e., *the bit-sizes grow exponentially* if we perform Euclid's algorithm in $\mathbb{Q}[x]$. Can we avoid this exponential growth? Is it possible to devise an algorithm that works only with integer polynomials and computes the gcd? Perhaps the gcd, assuming it is well defined, of the integer polynomials has coefficients of exponential bit-size [1] In this lecture we study these questions. We only focus on the univariate case in this lecture, since a multivariate polynomial in $R[x_1, \ldots, x_n]$, for some ring R, can be expressed as a univariate polynomial in x_n with coefficients in the ring $R[x_1, \ldots, x_{n-1}]$; the additional differences will be highlighted when necessary.

Throughout this lecture we will use the following notation:

$$A(x) = \sum_{i=0}^{m} a_i x^i, B(x) = \sum_{i=0}^{n} b_i x^i, m \geq n, a := \texttt{lead}(A), b := \texttt{lead}(B).$$

We willl assume that the coefficients are integers, however, all the results generalize to the case when the coefficient ring is some domain.

[1] This is unlike the integer case, where the bit-size is trivially bounded in the bit-size of the input numbers. Another point of difference between polynomials and integers.

1 Pseudo-Euclidean Polynomial Remainder Sequence

One approach to avoid working in the quotient ring $\mathbb{Q}[x]$ is based upon the following unravelling of the euclidean step: canceling the leading term of $A(x)$ using $B(x)$ gives us a polynomial whose leading term is $a^{(1)}x^{m-1}/b_n$, where $a^{(1)} := b_n a_{m-1} - b_{n-1}a_m \in \mathbb{Z}$; again canceling this term gives us a polynomial whose leading term is $a^{(2)}x^{m-1}/b_n^2$, where $a^{(2)}$ is recursively defined in terms of $a^{(1)}$; going on in this manner, we observe that the quotient is a polynomial of the form

$$Q(x) = \frac{a_m}{b_n}x^{m-n} + \frac{a^{(1)}}{b_n^2}x^{m-n-1} + \frac{a^{(2)}}{b_n^3}x^{m-n-2} + \cdots + \frac{a^{(m-n)}}{b_n^{m-n+1}},$$

i.e., the denominators of the coefficients divide b_n^{m-n+1}; the same observation applies to the remainder $R(x)$. In other words, the polynomials $b_n^{m-n+1}Q(x), b_n^{m-n+1}R(x)$ are integer polynomials, which implies that $\texttt{rem}(b_n^{m-n+1}A(x), B(x))$ is an integer polynomial. Define the **pseudoremainder** of two polynomials as

$$\texttt{prem}(A, B) := \texttt{rem}(\texttt{lead}(B)^{m-n+1}A, B) = \texttt{lead}(B)^{m-n+1}\texttt{rem}(A, B)?. \tag{2}$$

Similarly define the **pseudo-quotient**, $\texttt{pquo}(A, B) := \texttt{quo}(\texttt{lead}(B)^{m-n+1}A, B)$; note that the uniqueness of the pseudo-quotient and pseudoremainder follows from the uniqueness of the quotient and remainder in $\mathbb{Q}[x]$. Further define the **pseudo-Euclidean PRS** as $A_0 := A, A_1 := B$ and recursively

$$A_{i+1} := \texttt{prem}(A_{i-1}, A_i).$$

But what is the relation between the euclidean PRS over $\mathbb{Q}[x]$ and the pseudo-euclidean PRS? As we would expect, the pseudo-euclidean PRS is obtained from the euclidean PRS by clearing the denominators. In general, we will say **two polynomials** A, B **are similar**, $A \sim B$, if there are non-zero integers α, β such that $\alpha A = \beta B$. Thus the pseudo-euclidean PRS and euclidean PRS are similar sequences.

Does the pseudo-euclidean PRS resolve the exponential growth that we observed in (1)? The following computation shows that perhaps the answer is no:

$$\begin{aligned}
A_0 &:= x^8 + x^6 - 3\,x^4 - 3\,x^3 + 8\,x^2 + 2\,x - 5\\
A_1 &:= 2\,x^6 + 5\,x^4 - 4\,x^2 - 9\,x + 21\\
A_2 &:= 22\,x^4 + 12\,x^3 - 44\,x^2 - 38\,x + 86\\
A_3 &:= -42176\,x^3 + 58432\,x^2 + 126048\,x - 175776\\
A_4 &:= 143376257024\,x^2 - 4862984192\,x - 161945006080\\
A_5 &:= 16515934310457412139823923 20\,x - 22898619670252191772269608 96\\
A_6 &:= 29165011975794224984814075250533644773674517875478096214076319 3344.
\end{aligned} \tag{3}$$

We next show that this is indeed the case. For this we need to re-interpet the euclidean step in terms of matrices. *The key idea is that a euclidean step is a special case of Gaussian elimination.*

Given $A(x)$ and $B(x)$ consider the following $(m-n+2)\times(m+1)$ matrix in which the coefficients of B appear in the first $m-n+1$ rows and the coefficients of A in the last row:

$$M := \begin{bmatrix}
b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & & & & \\
 & b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & & & \\
 & & b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & & \\
 & & & \ddots & & & & \ddots & \\
 & & & & b_n & b_{n-1} & b_{n-2} & & b_0 \\
a_m & a_{m-1} & a_{m-2} & \cdots & & & & & a_0
\end{bmatrix} \tag{4}$$

The Euclidean step can be interpreted in terms of elementary row operations as follows: take the last row and subtract from it a multiple of a_m/b_n of the first row; this eliminates the first entry of the last row and is equivalent to the getting the leading entry of the quotient; next eliminate the second entry of the last row by multiplying with a suitable scaling of the second row; so on and so forth, we keep on modifying the last row; this can be done at most $m-n+1$ times;

the entries of the last row, which are now rational numbers, are clearly the coefficients of $\texttt{rem}(A, B)$, and the resulting matrix looks as follows

$$M' := \begin{bmatrix} b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & & & & \\ & b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & & & \\ & & b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & & \\ & & & \ddots & & & & \ddots & \\ & & & & b_n & b_{n-1} & b_{n-2} & & b_0 \\ 0 & 0 & 0 & \cdots & & c_{n-1} & c_{n-2} & & c_0 \end{bmatrix}, \tag{5}$$

where $c_i \in \mathbb{Q}$ are the coefficients of $\texttt{rem}(A, B)$. Associate a degree $(n-1)$ **determinantal polynomial** with M defined as follows:

$$\texttt{dpol}(M) := \det(M_1)x^{n-1} + \det(M_{n-2})x^{n-2} + \cdots + \det(M_{n+1}), \tag{6}$$

where M_i is the square-matrix obtained from M by taking the first $m-n+1$ columns and the $(m-n+1+i)$th column of M, where $i = m-n+2, \ldots, m+1$. Since elementary row operations do not change the determinant, it follows that

$$\texttt{dpol}(M) = \texttt{dpol}(M') = b^{m-n+1}\texttt{rem}(A, B) = \texttt{prem}(A, B). \tag{7}$$

The equation above gives us a very explicit description of the coefficients of the pseudoremainder in terms of the coefficients of A, B. It also implies the following relation: for two integers α, β

$$\texttt{prem}(\alpha A, \beta B) = \alpha\beta^{m-n+1}\texttt{prem}(A, B). \tag{8}$$

Based upon this relation, we can now show that the coefficient growth in pseudo-euclidean PRS is exponential.

¶1. Coefficient growth in pseudo-euclidean PRS Suppose $A_0, \ldots, A_k$ is the pseudo-euclidean PRS with degree sequence $(n_0, n_1, \ldots, n_k)$ and $\delta_i := n_{i-1} - n_i$. Let $\beta := \texttt{cont}(A_2)$ and $A_2' := \texttt{prim}(A_2)$. Then

$$A_3 := \texttt{prem}(A_1, A_2) = \texttt{prem}(A_1, \beta A_2').$$

Thus from (8) it follows that $\beta^{\delta_2+1} | A_3$. Similarly, let $A_3 = \beta^{\delta_2+1}A_3'$. Then we have

$$A_4 := \texttt{prem}(A_2, A_3) = \texttt{prem}(A_2, \beta^{\delta_2+1}A_3')$$

and from (8) it again follows that

$$\beta^{(\delta_2+1)(\delta_3+1)} | A_4.$$

Continuing in this manner we see that

$$\beta^{(\delta_2+1)(\delta_3+1)\ldots(\delta_k+1)} | A_k.$$

Since $\delta_i \geq 1$, it immediately follows that $\beta^{2^k} | A_k$, i.e., the bit-sizes of A_k are exponential in k.

2 Primitive PRS

As pseudoremainders do not solve the exponential coefficient growth, we have to look at other alternatives. Can we find other similar sequences that have polynomial growth? Perhaps one way to avoid the coefficient growth in (3) is to try to remove the part common to all the coefficients, or more precisely the content of the polynomial. That is, at each step we do the following:

$$A_{i+1} := \texttt{prim}(\texttt{prem}(A_{i-1}, A_i)).$$

Call such a PRS as the **primitive PRS**. The primitive PRS for our example is

$$\begin{aligned}
A_0 &:= x^8 + x^6 - 3\,x^4 - 3\,x^3 + 8\,x^2 + 2\,x - 5 \\
A_1 &:= 2\,x^6 + 5\,x^4 - 4\,x^2 - 9\,x + 21 \\
A_2 &= 11\,x^4 + 6\,x^3 - 22\,x^2 - 19\,x + 43 \\
A_3 &= -1318\,x^3 + 1826\,x^2 + 3939\,x - 5493 \\
A_4 &= 26299\,x^2 - 892\,x - 29705 \\
A_5 &= 3998585\,x - 5543863 \\
A_6 &= 1.
\end{aligned} \tag{9}$$

Clearly, the primitive PRS is similar to pseudo-euclidean PRS, but it seems to have the added advantage of small coefficient growth. We can show that the coefficient sizes are polynomially bounded in the input. We will next develop a more general framework to develop such bounds.

3 Subresultant PRS

Despite its polynomial coefficient growth, primitive PRS does have one drawback, namely at each step we have to do a multi-gcd computation, which becomes more prominent when we are working over the ring of multivariate polynomials since then the computation would proceed recursively in the dimension and so the cost increases dramatically. Thus *a desirable PRS algorithm should avoid computing multiple gcds*. This aim was successfully achieved by Collins subresultant PRS algorithm [1]. The algorithm reduces the coefficient growth by dividing by an integer at each step, i.e., it obtains a sufficiently large factor of the content at each step without doing any multiple gcds. In the remaining lecture we focus on this algorithm and show that it has nearly linear coefficient growth. .

All the modifications of the standard PRS can be captured by the following notion: A **generalized PRS** for two polynomials $A_0, A_1 \in \mathbb{Z}[x]$, based upon sequences $\{\alpha_i\}$ and $\{\beta_i\}$ is a sequence of polynomials A_i, $i = 1, \ldots, k$ such that

$$\beta_i A_{i+1} := \alpha_i A_{i-1} - Q_i A_i, \tag{10}$$

where $Q_i := \mathtt{pquo}(A_{i-1}, A_i)$. This immediately implies that

$$A_{i+1} \sim \mathtt{prem}(A_{i-1}, A_i)$$

and

$$\mathtt{GCD}(A_0, A_1) \sim A_k$$

In general we have the following chain of relations

$$\mathtt{GCD}(A_0, A_1) \sim \mathtt{GCD}(A_1, A_2) \sim \cdots \sim \mathtt{GCD}(A_{k-1}, A_k).$$

Let $n_i := \deg(A_i)$, and $\delta_i := n_{i-1} - n_i$.

For pseudo-euclidean PRS the base sequences are $\beta_i = 1$ and $\alpha_i = \mathtt{lead}(A_i)^{\delta_i+1}$; for primitive PRS the base sequences are $\beta_i = \mathtt{cont}(A_{i+1})$ and $\alpha_i = \mathtt{lead}(A_i)^{\delta_i+1}$. Thus pseudo-euclidean PRS and primitive PRS choose the extreme values for β_i. Collin's subsresultant PRS covers the ground in between by choosing a value for $\beta_i > 1$ without computing multiple gcds.

Our definition (4) helps us to understand one euclidean step. To understand a chain of such steps, we need a more general definition of $\mathtt{dpol}(M)$ where the coefficients of A and B are treated more equally. One way to express the matrix in (4) is as

$$\mathtt{dpol}(x^{m-n}B, x^{m-n-1}B, \ldots, B, A).$$

Based upon this formulation we have the following definition: The ith subresultant, $i = 0, \ldots, n$, of A and B is defined as

$$\mathtt{sres}_i(A, B) := \mathtt{dpol}(x^{n-i-1}A, x^{n-i-2}A, \ldots, A; x^{m-i-1}B, x^{m-i-2}B, \ldots, B). \tag{11}$$

For convenience, we will use the shorthand S_i for $\texttt{sres}_i(A, B)$. The underlying matrix has $(m + n - 2i)$ rows and $(m + n - i)$ columns and so $\deg(S_i) \leq i$; as the leading coefficients could vanish; see Figure 1 for illustration. We will say that S_i is **regular** if the degree is equal to i; otherwise, we will say it is **defective**. Note that

$$S_n = b_n^{m-n-1} B. \tag{12}$$

It is convenient to extend the definition to cover the cases $i = n + 1, \ldots, m$:

$$\begin{aligned} S_m &:= A \\ S_{m-1} &:= B \\ S_{m-2} = S_{m-3} &= \cdots = S_{n+1} := 0. \end{aligned} \tag{13}$$

The sequence

$$(S_m, \ldots, S_0) \tag{14}$$

is called the **subresultant chain**. Let $C_i := \texttt{lead}(S_i)$ be the nominal leading coefficient of S_i, also called the **principal subresultant coefficient**; *define* $C_m := 1$ *and not* $\texttt{lead}(A)$. If all the subresultants are regular then the chain is called regular.

Our main result would be along the following lines: for a generalized PRS $A_0, A_1, \ldots, A_k$

$$A_i \sim S_{n_{i-1}-1} \sim S_{n_i}, \quad i \leq k$$

and

$$S_j = 0, \text{ for } n_i < j < n_{i-1} - 1.$$

That is the subresultant chain captures all types of PRS's up to similarity. The following properties make them our focus of study:

1. $S_{n_{k-1}-1} \sim \texttt{GCD}(A_0, A_1)$.
2. $S_0 = 0$ iff $\deg(\texttt{GCD}(A_0, A_1)) > 0$.
3. They have polynomially sized coefficients.

Thus if we can construct a PRS that equals the corresponding subresultants then we have achieved our desired aim of a polynomial running time variant of Euclid's algorithm for rings. Collins [1] was the first to show this result.

Our first crucial observation is the following analogue of (7):

LEMMA 1. *For* $i = 0, \ldots, n$,

$$b^{(m-n+1)(n-i-1)} S_i = (-1)^{(m-i)(n-i)} \texttt{sres}_i(B, \texttt{prem}(A, B)). \tag{15}$$

Proof. The following steps constitute the proof of correctness (it helps to follow the illustration in Figure 1 simultaneously).

1. Multiply both sides of (11) $b^{(m-n+1)(n-i)}$. This gives us

   $$b^{(m-n+1)(n-i)} S_i = \texttt{dpol}(x^{n-i-1} b^{m-n+1} A, x^{n-i-2} b^{m-n+1} A, \ldots, b^{m-n+1} A; x^{m-i-1} B, x^{m-i-2} B, \ldots, B).$$

2. Now observe that there are $(m - n + 1 + n - i - 1)$ B-rows, i.e., for each A-row we have $(m - n + 1)$ rows to perform Gaussian elimination to get pseudoremainder. Proceeding with the elimination yields the matrix show in Figure 2(a), i.e., the matrix corresponding to

   $$\texttt{dpol}(x^{n-i-1}\texttt{prem}(A, B), x^{n-i-2}\texttt{prem}(A, B), \ldots, \texttt{prem}(A, B); x^{m-i-1} B, x^{m-i-2} B, \ldots, B).$$

3. Swapping the rows gives us

   $$S_i = (-1)^{(m-i)(n-i)} \texttt{dpol}(x^{n-i-2} B, x^{n-i-3} B, \ldots, B; x^{n-i-1}\texttt{prem}(A, B), x^{n-i-2}\texttt{prem}(A, B), \ldots, \texttt{prem}(A, B)).$$

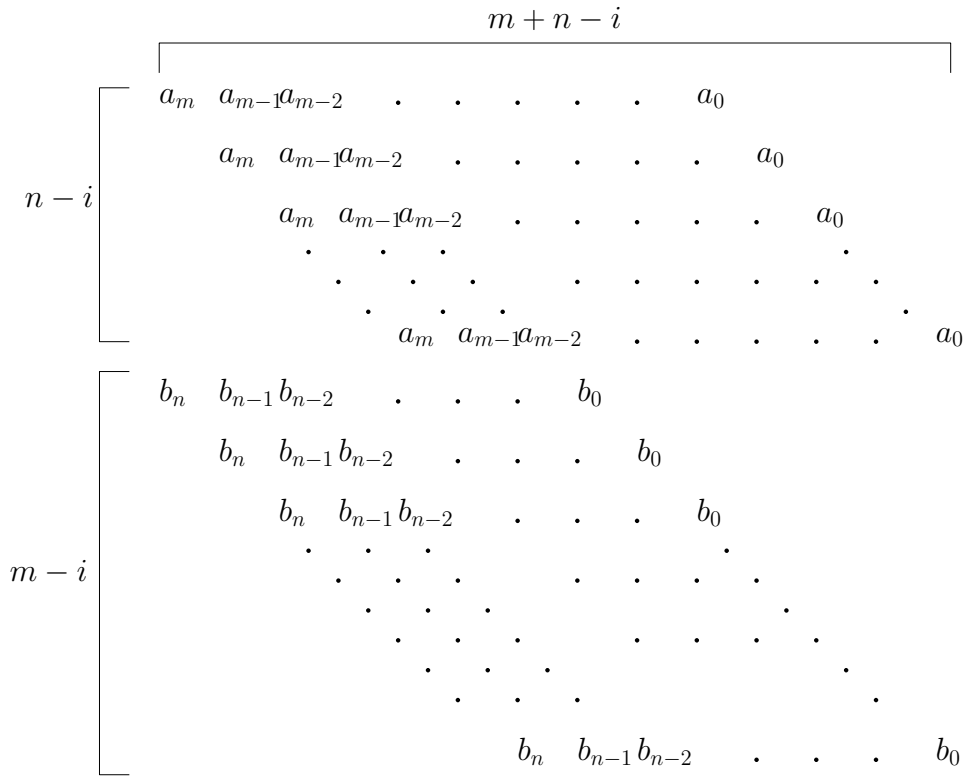


Figure 1: Illustrating $\texttt{sres}_i(A,B)$

4. As the main diagonal of the the first $(m-n+1)$ columns and rows contains b_n, we pullout a factor of b^{m-n+1} in the dpol to get

   $(-1)^{(m-i)(n-i)}b^{(m-n+1)(n-i-1)}S_i = \texttt{dpol}(x^{n-i-1}\texttt{prem}(A,B), x^{n-i-2}\texttt{prem}(A,B), \ldots, \texttt{prem}(A,B); x^{n-i-2}B, x^{n-i-3}B, \ldots$

   The resulting matrix is shown in Figure 2(b).

Q.E.D.

In particular, from (15) it follows that for $i = n-1$ we have

$$S_{n-1} = \texttt{prem}(A,B); \tag{16}$$

in Figure 2(b), if we substitute $i = n-1$ then all the B-rows vanish and only one row corresponding to $\texttt{prem}(A,B)$ survives. Thus so far we have that

$$S_m = A, S_{m-1} = B, S_{m-2} = \cdots = S_{n+1} = 0, S_n = b^{m-n+1}B, \text{ and } S_{n-1} = \texttt{prem}(A,B).$$

Call such a sequence of subsresultants a block. More precisely, a **block** is a sequence of polynomials $(T_1, \ldots, U_k)$ such that $T_1 \sim U_k$ and the intermediate polynomials are all zeros; T_1 is called the **top** and U_k the **base** of the block. In general, we will see the same phenomenon, i.e., the subresultant chain $(S_m, \ldots, S_0)$ can be partitioned into blocks $B_0, \ldots, B_k$.

We start with a simple case first, namely when $n = m-1$ and the subresultant chain is regular. In other words, we treat the coefficients purely symbolically. Let's call this interpretation as the symbolic resultant chain. Pictorially it is as shown in Figure **??**. What can we say about this symbolic chain? The following theorem gives us a handle on this chain.

THEOREM 2 (Habitch's Theorem). *Let $A(x)$ and $B(x)$ be degree m and degree $(m-1)$ polynomials and $(S_m, \ldots, S_0)$ the corresponding* symbolic subsresultant chain. *Then for all $j < m$,*

$$C_{j+1}^{2(j-i)}S_i = \texttt{sres}_i(S_{j+1}, S_j), \quad i = 0, \ldots, j-1. \tag{17}$$

Proof. The proof is by induction on j.

1. The base case is $j = m - 1$. Observe that when $j = m - 1$, (17) is just the definition of subresultants (15) as C_m was defined as 1, so (17) holds.

2. So suppose that (17) has been demonstrated for some j. Then we have for $i = j - 1$

$$C_{j+1}^2 S_{j-1} = \mathtt{sres}_{j-1}(S_{j+1}, S_j) = \mathtt{prem}(S_{j+1}, S_j), \tag{18}$$

   where the second equality follows from the same argument that was used to show (12) by substituting $n = j-1$, $A \leftarrow S_{j+1}$, and $B \leftarrow S_j$. Now we show that (17) holds for $j - 1$ and for $i < j - 1$.

3. Since (17) holds for j by assumption we know that

$$S_i = C_{j+1}^{-2(j-i)} \mathtt{sres}_i(S_{j+1}, S_j).$$

   Substituting $b = C_j$, $A = S_{j+1}$, $B = S_j$ in (15) we get that

$$\mathtt{sres}_i(S_{j+1}, S_j) = C_j^{-2(j-1-i)} \mathtt{sres}_i(S_j, \mathtt{prem}(S_{j+1}, S_j)).$$

   Thus we have for $i < j - 1$

$$\begin{aligned} S_i &= C_{j+1}^{-2(j-i)} C_j^{-2(j-1-i)} \mathtt{sres}_i(S_j, \mathtt{prem}(S_{j+1}, S_j)) \\ &= C_{j+1}^{-2(j-i)} C_j^{-2(j-1-i)} \mathtt{sres}_i(S_j, C_{j+1}^2 S_{j-1}) \\ &= C_j^{-2(j-1-i)} \mathtt{sres}_i(S_j, S_{j-1}). \end{aligned}$$

Q.E.D.

We now study the effect of removing the constraint that $n = m - 1$ and the chain is regular, i.e., we specialize the coefficients to their actual values. In particular, we want to understand what happens when we have defective subresultants.

THEOREM 3 (Subresultant Block Structure Theorem). *Let $j < m$. Suppose S_{j+1} is regular and S_j is defective with degree $d < j$. Then*

1. $S_{j-1} = S_{j-2} = \cdots S_{d+1} = 0$.
2. $C_{j+1}^{j-d} S_d = \mathtt{lead}(S_j)^{j-d} S_j$, *i.e.,* $S_d \sim S_j$ *and*
3. $C_{j+1}^{j-d+2} S_{d-1} \sim \mathtt{prem}(S_{j+1}, S_j)$.

This theorem gives us the block structure that we wanted. It explicitly gives us the coefficients of similarity between the top and the base of a block. Moreover, the third item gives the relation between two consecutive blocks. *Proof.* We consider the following three cases depending on how many leading zeros S_j has, i.e., how $j - r$ interacts with $j - i$. Each of the three cases give us the desired results respectively. See Figure 3 for an illustration of each case.

1. $d + 2j - 2i + 1 < 2j - i + 1$: In this case, the coefficients involved in the subresultant all vanish, because the matrices are upper triangular with zeros on the main diagonal. The condition is equivalent to $i > d$. See Figure 3(a).

2. $d + 2j - 2i + 1 = 2j - i + 1$: That is the coefficients of S_{j+1} and S_j are exactly on the main diagonal. From (17) it follows that

$$C_{j+1}^{2(j-i)} S_i = \mathtt{sres}_i(S_{j+1}, S_j) = C_{j+1}^{j-d} \mathtt{lead}(S_j)^{j-d} S_j.$$

   The condition is equivalent to $i = d$. See Figure 3(b).

3. $d + 2j - 2i + 1 = 2j - i + 2$: In this case, we expand the determinant along the first $j - d$ rows to get

$$\mathtt{sres}_{d-1}(S_{j+1}, S_j) = C_{j+1}^{j-d} \mathtt{dpol}(S_{j+1}, x^{j-d+1} S_j, \ldots, S_j) = C_{j+1}^{j-d} (-1)^{j-d+2} \mathtt{prem}(S_{j+1}, S_j)$$

   where the last equality follows from (15). The condition is equivalent to $i = d - 1$.

Q.E.D.

We now consider a specialization of (10) which gives us a direct approach to compute subresultants.

THEOREM 4 (Subresultant PRS Theorem). *Given A, B, consider the PRS obtained from (10) based upon the standard choice of α_i and*

$$\beta_{i+1} := -\mathtt{lead}(A_i)(-C_{n_i})^{\delta_{i+1}}, \quad \beta_1 := 1. \tag{19}$$

Then

$$A_i = S_{n_{i-1}-1}, \quad i = 1, \ldots, k.$$

That is A_i's are the top of the blocks in the subresultant chain.

Proof. The proof is by induction on i. The base case $i = 0, 1$ holds by definition since $A_0 := A = S_m = S_{n_0}$ and $A_1 := B = S_{m-1} = S_{n_0-1}$. Assume the theorem holds for A_i and A_{i+1}, i.e.,

$$A_i = S_{n_{i-1}-1} \text{ and } A_{i+1} = S_{n_i-1}.$$

Note that both $S_{n_{i-1}-1}$ and S_{n_i-1} are defective. Our assumption that $A_i = S_{n_{i-1}-1}$ implies that $S_{n_{i-1}-1}, \ldots, S_{n_i}$ forms a block. Applying Theorem 3(2) with $j + 1 = n_{i-1}$ and $d = n_i$, we obtain that

$$C_{n_{i-1}}^{\delta_i - 1} S_{n_i} = \mathtt{lead}(A_i)^{\delta_i - 1} A_i. \tag{20}$$

On comparing the leading coefficients on both sides we further get

$$C_{n_{i-1}}^{\delta_i - 1} C_{n_i} = \mathtt{lead}(A_i)^{\delta_i}. \tag{21}$$

Furthermore, substituting of $j + 1 = n_i$ and $d = n_{i+1}$, Theorem 3(3) gives us

$$(-C_{n_i})^{\delta_{i+1}+1} S_{n_{i+1}-1} = \mathtt{prem}(S_{n_i}, A_{i+1}). \tag{22}$$

Note that the choice of α_i implies that

$$\beta_{i+1} A_{i+2} = \mathtt{prem}(A_i, A_{i+1}).$$

Substituting A_i in terms of S_{n_i} using (20) gives us

$$\beta_{i+1} A_{i+2} = C_{n_{i-1}}^{\delta_i - 1} \mathtt{lead}(A_i)^{-(\delta_i - 1)} \mathtt{prem}(S_{n_i}, A_{i+1}).$$

From (22) we further obtain

$$\beta_{i+1} A_{i+2} = C_{n_{i-1}}^{\delta_i - 1} \mathtt{lead}(A_i)^{-(\delta_i - 1)} (-C_{n_i})^{\delta_{i+1}+1} S_{n_{i+1}-1}.$$

But from (21) we get that

$$\beta_{i+1} A_{i+2} = -\mathtt{lead}(A_i)(-C_{n_i})^{\delta_{i+1}} S_{n_{i+1}-1}.$$

But from (19) we know that the constant on the RHS is β_i. Thus

$$A_{i+2} = S_{n_{i+1}-1}.$$

Q.E.D.

¶2. Coefficient Growth of Subresultant PRS A straight forward application of Hadamard's bound shows us that the coefficients in subresultant PRS are $O(n(L + \log n))$, where n is the degree of the larger polynomial and L the maximum bit-length of the input coefficients.

4 Resultant

Given two polynomials $A(x), B(x) \in \mathbb{Z}[x]$, we want to "eliminate the variable x", i.e., find two polynomials $P, Q \in \mathbb{Z}[x]$ such that the linear combination $\texttt{res}(A, B) := AP + BQ$ is independent of x and only belongs to the coefficient ring $\mathbb{Z}$. The linear combination $\texttt{res}(A, B)$ is called the **resultant** or eliminant.[2]. In this section, we study one special resultant, namely Sylvester's resultant. We have already encoutenred it as the subresultant S_0, but let's see why it naturally arises as the resultant of A and B. The starting point is the following result.

LEMMA 5. *The* $\deg(\texttt{GCD}(A, B)) \geq k$ *iff there exists two polynomials* P *and* Q*, of degree at most* $n - k$ *and* $m - k$ *respectively such that* $PA + QB = 0$*.*

Proof. If $\deg(\texttt{GCD}(A, B)) \geq k$, we can choose $P := B/\texttt{GCD}(A, B)$ and $Q := A/\texttt{GCD}(A, B)$.

Conversely, suppose $PA + QB = 0$. Since $\deg(P) \leq n - k$, this implies that a certain factor of A of degree at least k must divide Q. Thus $\deg(\texttt{GCD}(A, B)) \geq k$. **Q.E.D.**

The lemma above states that A and B have a non-constant gcd iff there exists two polynomials P and Q of degree $n-1$ and $m-1$ resp. such that $PA + QB = 0$. Considering the coefficients of P and Q as variables, we want to find a solution to the $m + n - 1$ equations

$$\sum_{j=0}^{k} p_j a_{k-j} + \sum_{j=0}^{k} q_j b_{k-j} = 0$$

where $k = 0, \ldots, m + n - 1$. We can conveniently express these equations as solution to a system of linear equation

$$(p_{n-1}, \ldots, p_0, q_{m-1}, \ldots, q_0) \cdot \mathrm{Syl}(A, B)$$

where $\mathrm{Syl}(A, B)$ is the matrix shown in Figure 4. Clearly, the above system of linear equations has a solution iff $\det(\mathrm{Syl}(A, B)) = 0$. Thus

$$\texttt{res}(A, B) = \det(\mathrm{Syl}(A, B)). \tag{23}$$

is the desired resultant. Thus we have a direct proof of the following result

LEMMA 6. *The* $\texttt{res}(A, B)$ *is zero iff* A *and* B *have a non-constant gcd.*

Another surprising equivalent condition for the vanishing of $\texttt{res}(A, B)$ is iff there exists a common root of A and B in the algebraic closure $\overline{\mathbb{Z}}$ (not necessarily in $\mathbb{Z}$!). This relation is captured by the following theorem, which gives the beautiful correlation between the algebra of coefficients and the geometry of roots.

THEOREM 7. *Let* $\alpha_1, \ldots, \alpha_m \in \overline{\mathbb{Z}}$ *be the roots of* $A(x)$ *and* $\beta_1, \ldots, \beta_n \in \overline{\mathbb{Z}}$ *the roots of* $B(x)$*. The following are all equivalent expressions for* $\texttt{res}(A, B)$*:*

1. $a^n \prod_{i=1}^{m} B(\alpha_i)$*,*
2. $b^m \prod_{i=1}^{n} A(\beta_i)$*, and*
3. $a^n b^m \prod_{i=1}^{n} \prod_{j=1}^{m} (\alpha_i - \beta_j)$*.*

The result follows from recursive application of the following observation: for any $\alpha \in \overline{\mathbb{Z}}$

$$\texttt{res}((x - \alpha) \cdot A, B) = B(\alpha)\texttt{res}(A, B). \tag{24}$$

For convenience we will take $A(x)$ to be a degree $m - 1$ polynomial. Thus $(x - \alpha)A = \sum_{i=0}^{m-1} (a_i - \alpha a_{i+1}) x^i$. The

[2]The definition can be generalized to a system of more than one variables

resultant has the following form:

$$M := \begin{bmatrix}
a_{m-1} & a_{m-2}-\alpha a_{m-1} & a_{m-3}-\alpha a_{m-2} & \cdots & a_0-\alpha a_1 & -\alpha a_0 & & \\
 & a_{m-1} & a_{m-2}-\alpha a_{m-1} & a_{m-3}-\alpha a_{m-2} & \cdots & a_0-\alpha a_1 & -\alpha a_0 & \\
 & & & \ddots & & & \ddots & \\
 & & & a_{m-1} & a_{m-2}-\alpha a_{m-1} & a_{m-3}-\alpha a_{m-2} & \cdots & a_0-\alpha a_1 \\
b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & & & \\
 & b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & & \\
 & & b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & \\
 & & & \ddots & & & & \ddots \\
 & & & & b_n & b_{n-1} & b_{n-2} &
\end{bmatrix}$$

Do the following operations on M in the prescribed order: multiply α to column i and add it to column $(i+1)$, for $i = 1, \ldots, m+n-1$. Let M' be matrix obtained. To describe the entries in M' corresponding to the rows of B, the following notation will be useful: define

$$B|_{x^i} := B(x) \texttt{ quo } x^i, \text{ for } i \geq 0. \tag{25}$$

For a certain value of $x = \alpha$, $B|_{\alpha^i}$ means the $B|_{x^i}$ evaluated at $x = \alpha$; thus

$$B|_{x^n} = b_n, B|_{x^{n-1}} = b_n x + b_{n-1}, \ldots, B|_{x^0} = B(x), B|_{x^{-1}} = xB(x) \ldots.$$

With the notation above, we can conveniently write

$$M' = \begin{bmatrix}
a_{m-1} & a_{m-2} & a_{m-3} & \cdots & a_0 & 0 & & \\
 & a_{m-1} & a_{m-2} & a_{m-3} & \cdots & a_0 & 0 & \\
 & & & \ddots & & & \ddots & \\
 & & & a_{m-1} & a_{m-2} & a_{m-3} & \cdots & a_0 \\
B|_{\alpha^n} & B|_{\alpha^{n-1}} & B|_{\alpha^{n-2}} & \cdots & B|_{\alpha^0} & \alpha B(\alpha) & \cdots & \alpha^{m-2}B(\alpha) \\
 & B|_{\alpha^n} & B|_{\alpha^{n-1}} & B|_{\alpha^{n-2}} & \cdots & B|_{\alpha^0} & \alpha B(\alpha) & \cdots \\
 & & B|_{\alpha^n} & B|_{\alpha^{n-1}} & B|_{\alpha^{n-2}} & \cdots & B|_{\alpha^0} & \alpha B(\alpha)\cdots \\
 & & & \ddots & & & \ddots & \\
 & & & B|_{\alpha^n} & B|_{\alpha^{n-1}} & B|_{\alpha^{n-2}} & \cdots & B|_{\alpha^0} \\
 & & & & B|_{\alpha^n} & B|_{\alpha^{n-1}} & B|_{\alpha^{n-2}} & \cdots
\end{bmatrix}$$

$$\begin{matrix} 0 \\ 0 \\ \\ \alpha^{m-1}B(\alpha) \\ \alpha^{m-2}B(\alpha) \\ \alpha^{m-3}B(\alpha) \\ \\ \alpha B(\alpha) \\ B|_{\alpha^0} \end{matrix}$$

Now do the following operations starting from the last row to the $(n+1)$th row: subtract α times jth row to $(j-1)$th row, where $m+n \geq j > n+1$. This yields us the following matrix

$$M'' = \begin{bmatrix}
a_{m-1} & a_{m-2} & a_{m-3} & \cdots & a_0 & 0 & & \\
 & a_{m-1} & a_{m-2} & a_{m-3} & \cdots & a_0 & 0 & \\
 & & & \ddots & & & \ddots & \\
 & & & a_{m-1} & a_{m-2} & a_{m-3} & \cdots & a_0 \\
b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & 0 & \cdots & 0 \\
 & b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & 0 & \cdots \\
 & & b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & 0\cdots \\
 & & & \ddots & & & & \ddots \\
 & & & b_n & b_{n-1} & b_{n-2} & \cdots & b_0 \\
 & & & & B|_{\alpha^n} & B|_{\alpha^{n-1}} & B|_{\alpha^{n-2}} & \cdots
\end{bmatrix}$$

$$\begin{matrix} 0 \\ 0 \\ \\ 0 \\ 0 \\ 0 \\ \\ 0 \\ B|_{\alpha^0} \end{matrix}$$

Since the determinant does not change in doing these operations, it follows that

$$\det(M) = \det(M') = \det(M'').$$

Expanding M'' along the last column, which only contains one non-zero entry, namely $B(\alpha)$, it follows that $\det(M) = B(\alpha)\texttt{res}(A, B)$.

¶3. Computing the Resultant. A straightforward determinant computation will be costly. The key to computing resultants is based upon the following recursion:

$$\texttt{res}(A, B) = (-1)^{mn} b^{m-r} \texttt{res}(B, R)$$

where $R := \texttt{rem}(A, B)$ and $r := \deg(R)$.

Note: Given two multivariate polynomials $A(x_1, \ldots, x_n), B(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ we can still take their resultant w.r.t. the variable x_n by treating them as univariate polynomials in x_n with coefficients in $\mathbb{Z}[x_1, \ldots, x_{n-1}]$. We will represent this resultatnt as $\texttt{res}_{x_n}(A, B)$, which is a polynomial in $\mathbb{Z}[x_1, \ldots, x_{n-1}]$.

4.1 Algebraic Numbers and Algebraic Integers

We know that every degree n polynomial has n roots in $\mathbb{C}$. However, when our polynomials are integer polynomials, it is not necessary to go to $\mathbb{C}$ to get all the roots. A root of an integer polynomial is called an **algebraic number**. The set $\overline{\mathbb{Z}}$ containing all the algebraic numbers is called the **algebraic closure** of $\mathbb{Z}$. [3] What is the structure of this set? Is it a ring or a field?

Clearly, $\mathbb{Z}, \mathbb{Q} \subseteq \overline{\mathbb{Z}}$. Moreover, $\overline{\mathbb{Z}} \subset \mathbb{C}$, since there are numbers such as π, e that are not the roots of any integer polynomials; such numbers are called **transcendental numbers**. In fact, by Cantor's diagonalization argument it follows that most numbers are transcendental in nature. We will show that $\overline{\mathbb{Z}}$ is a field, and just as $\mathbb{Z}$ provides the underlying arithmetic structure for $\mathbb{Q}$, there is a set that acts in an analogous manner to $\overline{\mathbb{Z}}$.

THEOREM 8. *Let α, β be two algebraic numbers and let $A(x), B(x)$ be polynomials such that $A(\alpha) = B(\beta) = 0$. Then*

1. *1α is a root of $x^m A(1/x)$, if $\alpha \neq 0$.*
2. *$\beta \pm \alpha$ is a root of $\texttt{res}_y(A(y), B(x \mp y))$.*
3. *$\alpha\beta$ is a root of $\texttt{res}_y(A(y), y^n B(x/y))$.*

Proof.

1. Immediate.
2. $$\begin{aligned}\texttt{res}_y(A(y), B(x \mp y)) &= a^n \prod_{i=1}^{m} B(x \mp \alpha_i) \\ &= a^n b^m \prod_{i=1}^{m} \prod_{j=1}^{n} (x \mp \alpha_i - \beta_j).\end{aligned}$$
3. $$\begin{aligned}\texttt{res}_y(A(y), y^n B(x/y)) &= a^n \prod_{i=1}^{m} \alpha_i^n B(x/\alpha_i) \\ &= a^n b^m \prod_{i=1}^{m} \alpha_i^n \prod_{j=1}^{n} (\frac{x}{\alpha_i} - \beta_j) \\ &= a^n b^m \prod_{i=1}^{m} \prod_{j=1}^{n} (x - \alpha_i \beta_j).\end{aligned}$$

Q.E.D.

An **algebraic integer** is an algebraic number that is a root of a monic polynomial, i.e., a polynomial with leading coefficient one. From the theorem above, it follows that $\mathbb{O}$ forms a ring. Furthermore, the set of algebraic integers is denoted as $\mathbb{O}$ and it plays the same role as the integers in rationals.

[3] In general, the set containing all the roots of polynomials in $D[x]$, for some domain D, is denoted as $\overline{D}$.

LEMMA 9. *Every algebraic number $\alpha = \beta/a$, where β is an algebraic integer and $a \in \mathbb{Z}$.*

Proof. Suppose $A(\alpha) = 0$. Consider the polynomial

$$a^{n-1}A(x) = \sum_{i=0}^{n} a_i a^{n-1-i}(ax)^i = x^n + a_{n-1}(ax)^{n-1} + a_{n-2}a(ax)^{n-2} + \cdots + a_0 a^{n-1}.$$

Clearly, $a\alpha$ is a root of the polynomial. But the polynomial is monic, so $a\alpha$ must be an algebraic integer, say β. Thus $\alpha = \beta/a$. **Q.E.D.**

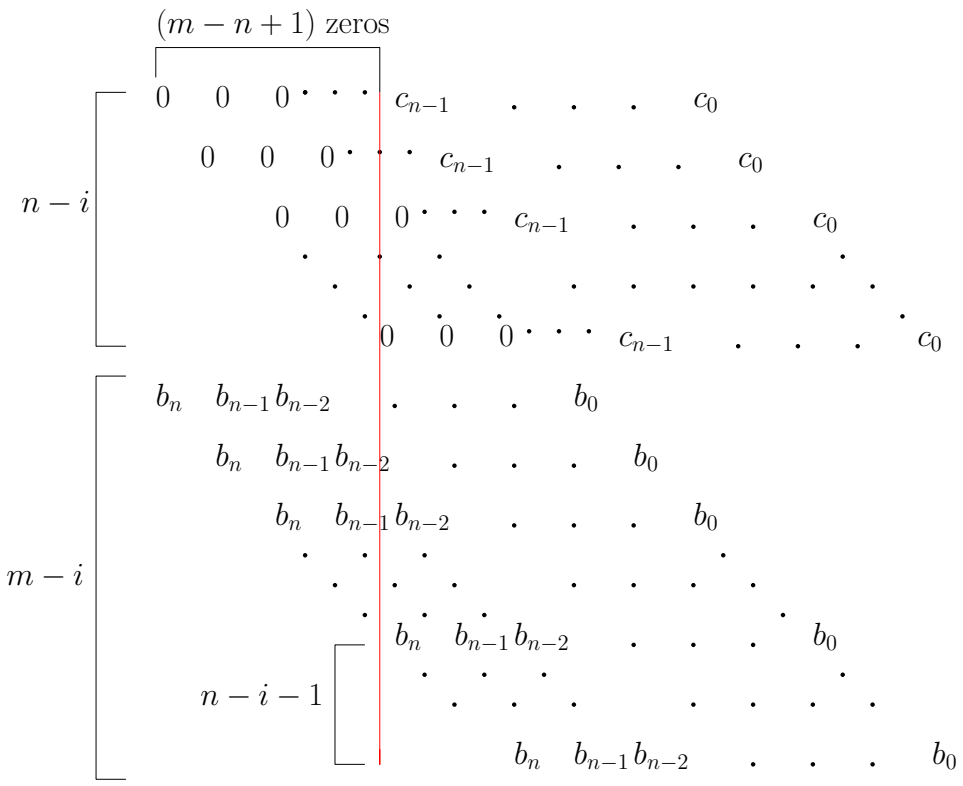
By definition, an algebraic number α is a root of an integer polynomial $A(x)$. It is often convenient to choose a specific polynomial as follows: we choose $A(x)$ to be a primitive polynomial that has the smallest degree, such a polynomial is called **the minimal polynomial** of α. They have the following property:

Let $A(x)$ be the minimal polynomial of α and suppose $B(x)$ is s.t. $B(\alpha) = 0$ then A divides B.

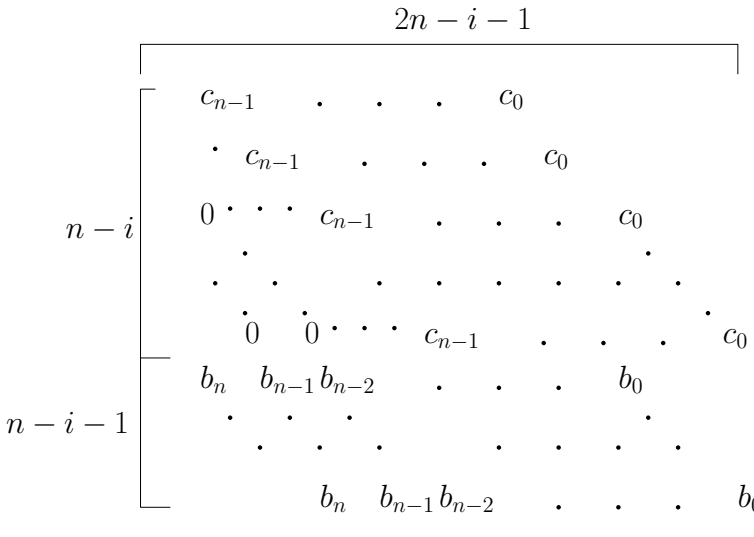
If not, then let $R(x)$ be the remainder, i.e., $B(x) = A(x)Q(x) + R(x)$; since both $A(\alpha) = B(\alpha) = 0$ it follows that $R(\alpha) = 0$; but $\deg(R) < \deg(A)$, which is a contradiction to the definition of minimal polynomial.

References

[1] G. E. Collins. Subresultants and reduced polynomial remainder sequences. *J. ACM*, 14:128–142, 1967.



(a)



(b)

Figure 2: After reducing the first $(n-i)$ rows to get the pseudoremainder $\sum_{i=0}^{n-1} c_i x^i$

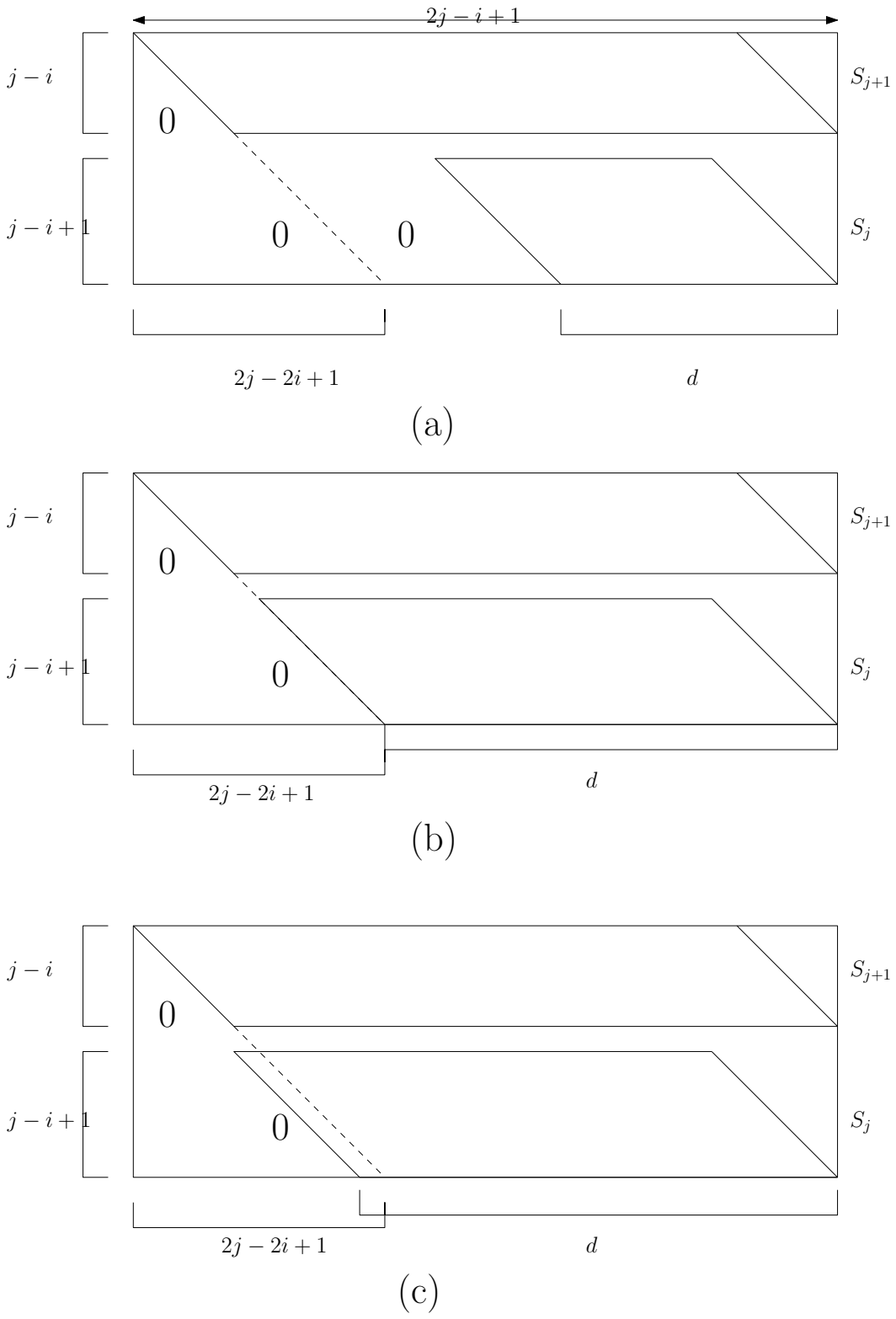


Figure 3: Illustration of proof of Subresultant Theorem

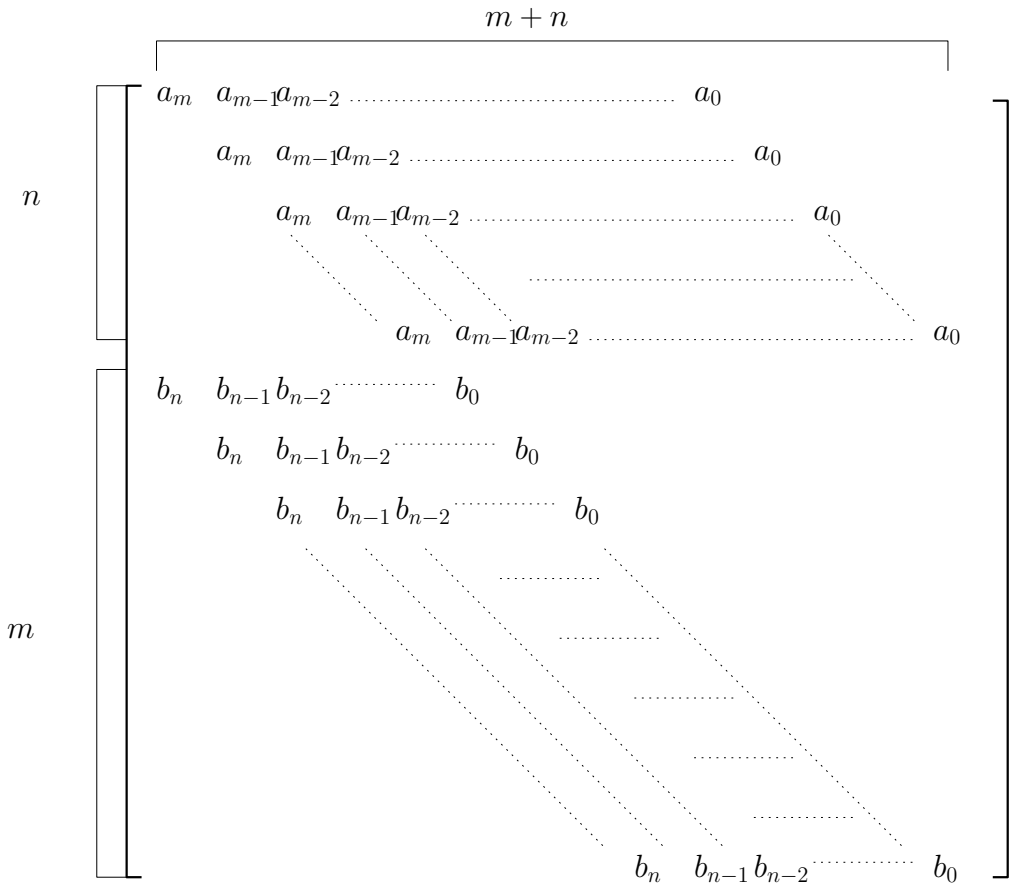


Figure 4: Sylvester matrix of two polynomials