Algorithms for polynomial systems: elimination & Gröbner bases

Lecture VI: Resultant and applications

June, 10th, 17th 2010. Part I: definition

Part II: Main formula

Part III: Applications

Xavier Dahan

Sylvester matrix

$$\begin{aligned} A(x) &= a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0, \quad a_m \neq 0, \quad m \geq 1 \\ B(x) &= b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0, \quad b_n \neq 0, \quad n \geq 1. \end{aligned}$$

$$\mathsf{Syl}(A,B) = \begin{pmatrix}
a_m & \cdots & a_1 & a_0 & \xleftrightarrow{\;n-1\;} & & & \\
 & a_m & & & a_0 & & & \\
{\scriptstyle n-1}\Big\updownarrow & & \ddots & & & \ddots & & \\
 & & & & a_m & \cdots & \cdots & a_0 \\
b_n & b_{n-1} & \cdots & \cdots & b_1 & b_0 & & \\
 & & \ddots & & & & \ddots & \\
\xleftrightarrow[\;m-1\;]{} & & b_n & \cdots & \cdots & \cdots & \cdots & b_0
\end{pmatrix}$$

$(n+m)$ lines

$(n+m)$ columns

Definition 1 *The* resultant *of A and B is the determinant of the Sylvester matrix of A and B:* $\boxed{\mathsf{Res}(A,B) = \det \mathsf{Syl}(A,B)}$.

Sylvester matrix as a linear map

Let A and B in $R[X]$ as in the previous slide.

$F = \mathrm{Frac}(R)$, field of fractions of R ($R = \mathbb{Z} \Rightarrow F = \mathbb{Q}$, $R = k[X] \Rightarrow F = k(X)$).

$R[X]_{<\ell}$ → poly. of degree strictly smaller than ℓ.

Let
$$\begin{array}{ccccc} \psi : R[X]_{<n} & \times & R[X]_{<m} & \rightarrow & R[X]_{<n+m} \\ (f & , & g) & \mapsto & Af + Bg \end{array} .$$

$$\mathcal{B} = \bigcup_{i=1}^{n} \{(X^{n-i}, 0)\} \bigcup_{j=1}^{m} \{(0, X^{m-j})\} = \{(X^{n-1}, 0), \ldots, (1, 0), (0, X^{m-1}) \ldots, (0, 1)\}$$

$\mathcal{B}$ → canonical basis of the *R-module* [a] $R[X]_{<n} \times R[X]_{<m}$.

$\mathcal{B}'$ → $(X^{n+m-1}, \ldots, X, 1)$ canonical basis of $R[X]_{<n+m}$.

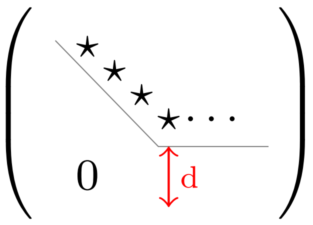
Claim: The matrix of the linear map ψ written in the bases $\mathcal{B}$ and $\mathcal{B}'$ is the *transpose*[b] of the Sylvester matrix of A and B. $\boxed{\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(\psi) = {}^t\mathsf{Syl}(A, B)}$

[a] or F-vector space $F[X]_{<n} \times F[X]_{<m}$, if the reader is not familiar with modules

[b] some authors do not take the transpose to define the Sylvester matrix

Sylvester matrix and GCD

Row echelon form:



d=number of zero lines, $\Rightarrow d = \dim\ker(\text{matrix})$. Either $\star \neq 0$, or $\star = 0$ and the first non-zero element of this line is on the right of the the first non-zero element of the *above* lines.

Gaussian elimination (without "pivot"): Every matrix over a field admits an equivalent[a] matrix in row echelon form.

Lemma 1 *The vector on the last non-zero line of the row echelon form of the Sylvester matrix of A and B corresponds to a gcd of A and B (in $F[X]$)*

Example: $A = 1x^4 - 1x^3 - 7x^2 + 2x + 3$ and $B = 1x^3 - 4x^2 + 2x + 3$.

$$\begin{pmatrix} 1 & -1 & -7 & 2 & 3 & 0 & 0 \\ 0 & 1 & -1 & -7 & 2 & 3 & 0 \\ 0 & 0 & 1 & -1 & -7 & 2 & 3 \\ 1 & -4 & 2 & 3 & 0 & 0 & 0 \\ 0 & 1 & -4 & 2 & 3 & 0 & 0 \\ 0 & 0 & 1 & -4 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 & -4 & 2 & 3 \end{pmatrix} \xrightarrow{\text{Gaussian elimination}} \begin{pmatrix} 1 & -1 & -7 & 2 & 3 & 0 & 0 \\ 0 & 1 & -1 & -7 & 2 & 3 & 0 \\ 0 & 0 & 1 & -1 & -7 & 2 & 3 \\ 0 & 0 & 0 & -14 & 45 & -3 & -18 \\ 0 & 0 & 0 & 0 & -\frac{5}{7} & \frac{12}{7} & \frac{9}{7} \\ 0 & 0 & 0 & 0 & 0 & \multicolumn{2}{c}{\boxed{\frac{1}{10} \quad -\frac{3}{10}}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The last non-zero line gives a gcdof A and B. $\gcd(A, B) = \frac{x}{10} - \frac{3}{10}$.

[a] represent the same linear application in different bases

Sylvester matrix and GCD

PROOF:Since GCDs are defined with coefficients in a *field*, and Gaussian elimination is done over a *field* we can work over $F = \mathrm{Frac}(R)$.

The last non-zero line has for coordinate the coefficients of a polynomial of minimal degree in the image of the map ψ. By definition (Cf. Lecture I),

$$\begin{aligned} h \text{ is s } \mathsf{gcd} \text{ of } A \text{ and } B &\Leftrightarrow \langle A, B\rangle = \langle h\rangle \text{ in } F[X] \\ &\Leftrightarrow h \text{ is of minimal degree among polynomials in } \langle A, B\rangle \end{aligned}$$

Finally, by the Bézout identity, a gcd of A and B is always in $\mathrm{Image}(\psi) \subset \langle A, B\rangle$ that permits to conclude the proof of Lemma 1. □

Corollary 1 *Let A and B be two polynomials in $F[X]$ (F a field). Then A and B have a (non-trivial) common factor iff $\mathsf{Res}(A, B) = 0$.*

PROOF: A and B have a common factor $\Leftrightarrow \deg \mathsf{gcd}(A, B) > 0$

$$\begin{aligned} &\Leftrightarrow \dim \ker(\mathsf{Syl}(A, B)) > 0 \\ &\Leftrightarrow \mathsf{Res}(A, B) = \det(\mathsf{Syl}(A, B)) = 0. \; \square \end{aligned}$$

Specialization of the resultant

Map of rings: Let R_1 and R_2 be 2 integral domains, and let $\phi : R_1 \to R_2$ be a ring morphism.

This map extends to $\phi : R_1[X] \to R_2[X]$:

$$\forall i,\ a_i \in R_1, \qquad \phi\left(\sum_i a_i X^i\right) = \sum_i \phi(a_i) X^i \ \in R_2[X].$$

And more generally to a map $\phi : R_1[X_1, \ldots, X_n] \to R_2[X_1, \ldots, X_n]$, or to a map $\phi : \mathrm{Mat}_{n\times m}(R_1) \to \mathrm{Mat}_{n\times m}(R_2)$.

Example: $\phi_p : \mathbb{Z}[X] \to \mathbb{F}_p[X]$ or for a prime p, or $\phi_a : k[X, Y] \to k[Y]$, $P(X, Y) \mapsto P(a, Y)$, for $a \in \bar{k}$.

Proposition 1 *Let $f, g \in R_1[X]$ and let $\phi : R_1 \to R_2$ a ring morphism.*

- *If $\phi(\mathrm{LC}(f))\phi(\mathrm{LC}(g)) \neq 0,$ then $\phi(\mathsf{Res}(f, g)) = \mathsf{Res}(\phi(f), \phi(g))$.*
- *If $\phi(\mathrm{LC}(f)) \neq 0,$ then $\phi(\mathsf{Res}(f, g)) = \phi(\mathrm{LC}(f))^{\deg(g) - \deg(\phi(g))}\mathsf{Res}(\phi(f), \phi(g))$.*

PROOF: *(of Prop. 1)* In the first case, we have $\deg(f) = \deg(\phi(f))$, and $\deg(g) = \deg(\phi(g))$, hence $\mathsf{Syl}(f, g) \in \mathrm{Mat}(R_1)$ and $\mathsf{Syl}(\phi(f), \phi(g)) \in \mathrm{Mat}(R_2)$ have same size, namely $\deg(f) + \deg(g)$.

It follows that $\phi\left(\mathsf{Syl}(f, g)\right) = \mathsf{Syl}(\phi(f), \phi(g))$. The determinant is defined by + and × operations only, hence $\phi(\det(\mathsf{Syl}(f, g))) = \det(\phi(\mathsf{Syl}(f, g)))$, which is equal to $\det(\mathsf{Syl}(\phi(f), \phi(g))) = \mathsf{Res}(\phi(f), \phi(g))$ as just seen.

In the second case, maybe $\phi(\textsc{lc}(g)) = 0$. This implies $\deg(\phi(g)) < \deg(g)$, and $\mathsf{Syl}(\phi(f), \phi(g))$ is a *smaller matrix* than $\phi(\mathsf{Syl}(f, g))$.

Let $f = f_m X^m + \cdots$ and $g = g_n X^n + \cdots$.

We denote $\phi(a) = \bar{a} \in R_2$, for $a \in R_1$.

Let $n' = \deg(\bar{g})$, ($n' < n$ in this case), so that:

$\phi(g) = \bar{g}_{n'} X^{n'} + \cdots$ and $\phi(f) = \bar{f}_m X^m + \cdots$.

The image by ϕ of the Sylvester matrix of f and g is written hereunder.

$$\phi(\mathsf{Syl}(f,g)) = \begin{pmatrix} & \bar f_m & \cdots & \cdots & \cdots & \bar f_1 & \bar f_0 & \xleftrightarrow{n-1} & \\ & & \ddots & & & & \ddots & & \\ & & & \bar f_m & \cdots & \cdots & \cdots & \cdots & \bar f_0 \\ \xleftrightarrow{n-n'} & \bar g_{n'} & \cdots & \bar g_1 & \bar g_0 & & & & \\ & & \ddots & & & \ddots & & & \\ & \xleftrightarrow[n-n'+m-1]{} & & & \bar g_{n'} & \cdots & \cdots & \bar g_0 & \end{pmatrix}$$

We compute the determinant along the first column:

$$\phi(\mathsf{Res}(f,g)) = \bar f_m \begin{vmatrix} & \bar f_m & \cdots & \cdots & \cdots & \bar f_1 & \bar f_0 & \xleftrightarrow{n-2} \\ & & \ddots & & & & \ddots & \\ & & & \bar f_m & \cdots & \cdots & \cdots & \bar f_0 \\ \xleftrightarrow[n-n'-1]{} & \bar g_{n'} & \cdots & \bar g_1 & \bar g_0 & & & \\ & & \ddots & & & \ddots & & \\ & \xleftrightarrow[n-n'+m-2]{} & & & \bar g_{n'} & \cdots & \cdots & \bar g_0 \end{vmatrix} = \bar f_m^2 \left| \ddots \right| = \bar f_m^{n-n'} \begin{vmatrix} \bar f_m & \cdots & \cdots & \bar f_1 & \bar f_0 & \xleftrightarrow{n'-1} \\ & \ddots & & & & \ddots \\ & & & & \cdots & \cdots \\ \bar g_{n'} & \cdots & \bar g_1 & \bar g_0 & & \\ & \ddots & & & \ddots & \\ \xleftrightarrow[m-1]{} & & \bar g_{n'} & \cdots & \cdots & \bar g_0 \end{vmatrix}$$

This last matrix is equal to $\mathsf{Syl}(\phi(f), \phi(g))$. This shows that

$$\phi(\mathsf{Res}(f,g)) = \bar f_m^{n-n'} \mathsf{Res}(\phi(f), \phi(g)).$$

We conclude by seeing that $\bar f_m = \phi(\textsc{lc}(f))$, and $n - n' = \deg(g) - \deg(\phi(g))$.□

Main Theorem

Let $\mathfrak{A}, \mathfrak{a}_1, \ldots, \mathfrak{a}_m$ and $\mathfrak{B}, \mathfrak{b}_1, \ldots, \mathfrak{b}_n$ be $n+m+2$ indeterminates.

And let $R = \mathbb{Z}[\mathfrak{A}, \mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{B}, \mathfrak{b}_1, \ldots, \mathfrak{b}_n]$ be the polynomial ring in these $n+m+2$ indeterminates.

Theorem 1 *Let A and B be polynomials in $R[X]$:*

$$\begin{aligned} A &= \mathfrak{A}(X-\mathfrak{a}_1)(X-\mathfrak{a}_2)\cdots(X-\mathfrak{a}_m) \\ B &= \mathfrak{B}(X-\mathfrak{b}_1)(X-\mathfrak{b}_2)\cdots(X-\mathfrak{b}_n), \end{aligned}$$

with a and b the leading coefficients in R. Then:

$$\begin{aligned} \mathsf{Res}(A,B) &\overset{(1)}{=} \mathfrak{A}^n\mathfrak{B}^m \prod_{\substack{1\le i\le m \\ 1\le j\le n}} (\mathfrak{a}_i - \mathfrak{b}_j) \overset{(2)}{=} (-1)^{mn}\mathfrak{B}^m \prod_{1\le j\le n} A(\mathfrak{b}_j) \\ &\overset{(3)}{=} \mathfrak{A}^n \prod_{1\le i\le m} B(\mathfrak{a}_i) \overset{(4)}{=} (-1)^{mn}\mathsf{Res}(B,A). \end{aligned}$$

Corollary 2 *Let $p_A(X)$ and $p_B(X)$ be two polynomials in $R_0[X]$, that are completely factorized in R: $p_A(X) = a(X_m - \alpha_1)\cdots(X - \alpha_m)$ and $p_B(X) = b(X - \beta_1)\cdots(X - \beta_n)$ with a and b the leading coefficients in R_0 as well. Then:*

$$\mathsf{Res}(p_A, p_B) = a^n b^m \prod\nolimits_{i,j}(\alpha_i - \beta_j) \tag{1}$$

Proof:We consider the ring morphism defined by:

$$\begin{aligned} \varphi : \mathbb{Z}[\mathfrak{A}, \mathfrak{a}, \mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{B}, \mathfrak{b}_1, \ldots, \mathfrak{b}_n] &\longrightarrow R_0 \\ \mathfrak{A} \text{ or } \mathfrak{B} \text{ or } \mathfrak{a}_i \text{ or } \mathfrak{b}_j &\longmapsto a \text{ or } b \text{ or } \alpha_i \text{ or } \beta_j. \end{aligned}$$

We notice that $\varphi(A) = p_A$ and $\varphi(B) = p_B$. By Theorem 1, we have:

$$\varphi\left(\mathsf{Res}(A, B)\right) = \varphi\left(\mathfrak{A}^n \mathfrak{B}^m \prod\nolimits_{i,j}(\mathfrak{a}_i - \mathfrak{b}_j)\right) = a^n b^m \prod\nolimits_{i,j}(\alpha_i - \beta_j).$$

Since $0 \neq a = \text{LC}(p_A) = \text{LC}(\varphi(A))$ and $0 \neq b = \text{LC}(p_B) = \text{LC}(\varphi(B))$, we are in the first ("good") case of the *specialization property* of the resultant, and it follows that $\varphi(\mathsf{Res}(A, B)) = \mathsf{Res}(\varphi(A), \varphi(B))$. □

Proof of the main theorem (1/4)

First, let us prove that the 4 equalities are equivalent.

- $\overset{(1)\Leftrightarrow(2)}{\bullet}$ $\prod_j A(\mathfrak{b}_j) = \prod_j \mathfrak{A} \prod_i (\mathfrak{b}_j - \mathfrak{a}_i) = (-1)^{mn}\mathfrak{A}^n \prod_{i,j}(\mathfrak{a}_i - \mathfrak{b}_j)$
- $\overset{(1)\Leftrightarrow(3)}{\bullet}$ Similar calculations as above.
- $\overset{(1)\Leftrightarrow(4)}{\bullet}$ $\mathsf{Res}(B,A) \overset{\text{by (1)}}{=} \mathfrak{B}^m\mathfrak{A}^n \prod_{j,i}(\mathfrak{b}_j - \mathfrak{a}_i) = (-1)^{mn}\mathfrak{A}^n\mathfrak{B}^m \prod_{i,j}(\mathfrak{a}_i - \mathfrak{b}_j) = (-1)^{mn}\mathsf{Res}(A,B)$.

Hence, we only need to prove Equality (1). Next we can assume $\mathfrak{A} = \mathfrak{B} = 1$. Indeed, if $A = \mathfrak{A}\tilde{A}$ and $B = \mathfrak{B}\tilde{B}$ ($\tilde{A}$ and $\tilde{B}$ are *monic*), then:

for $1 \le i \le n$ $\qquad$ i-th line of $\mathsf{Syl}(A,B) = \mathfrak{A}\times$ (i-th line of $\mathsf{Syl}(\tilde{A},\tilde{B})$)

for $1 \le j \le m$ $\qquad$ j-th line of $\mathsf{Syl}(A,B) = \mathfrak{B}\times$ (j-th line of $\mathsf{Syl}(\tilde{A},\tilde{B})$)

Since the determinant is *multilinear* with respect to the lines, it comes: $\mathsf{Res}(A,B) = \mathfrak{A}^n\mathfrak{B}^m\mathsf{Res}(\tilde{A},\tilde{B})$. Regarding the equality (1) we only need to prove $\boxed{\mathsf{Res}(A,B) = \prod_{i,j}(\mathfrak{a}_i - \mathfrak{b}_j)}$, where A and B are monic :$\mathfrak{A} = \mathfrak{B} = 1$

Proof of the main theorem (2/4)

Hence, we write $A = (X - \mathfrak{a}_1)\cdots(X - \mathfrak{a}_m)$ and $B = (X - \mathfrak{b}_1)\cdots(X - \mathfrak{b}_n)$.

Lemma 2 *The resultant* $\mathsf{Res}(A, B)$ *is a polynomial in* $\mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{a}_n, \mathfrak{b}_1, \ldots, \mathfrak{b}_m]$.

Proof:We have $A = X^m + \sum_{i=1}^{m} (-1)^i \mathfrak{s}_{i,m}(\mathfrak{a}_1, \ldots, \mathfrak{a}_m) X^{m-i}$, where $\mathfrak{s}_{i,m}$ is the i-th elementary symmetric polynomials in m variables.

$$\mathfrak{s}_{i,m}(\mathfrak{a}_1, \ldots, \mathfrak{a}_m) = \sum_{1 \leq \ell_1 < \ell_2 < \cdots < \ell_i \leq m} \mathfrak{a}_{\ell_1} \mathfrak{a}_{\ell_2} \cdots \mathfrak{a}_{\ell_i}. \tag{6}$$

Similarly, $B = X^n + \sum_{j=1}^{n} (-1)^j \mathfrak{s}_{j,n}(\mathfrak{b}_1, \ldots, \mathfrak{b}_n) X^{n-j}$.

Note that by Equation (6), $\mathfrak{s}_{i,m}$ and $\mathfrak{s}_{i,n}$ are in $\mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_n]$.

Hence the matrix $\mathsf{Syl}(A, B)$ has its entries in $\mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_n]$. Now since the determinant is a polynomial of $\mathbb{Z}[(n+m)^2$ entries of the matrix], it follows that $\mathsf{Res}(A, B) \in \mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_n]$. □

Proof of the main theorem (3/4)

Lemma 3 *In the polynomial ring* $\mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_n]$, *holds:*

$$\prod_{i,j} (\mathfrak{a}_i - \mathfrak{b}_j) \mid \mathsf{Res}(A, B).$$

Proof:Let $r(\mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_n) \in \mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_n]$ be a shorthand notation to denote the resultant: $\mathsf{Res}(A, B) = r$.

For each $1 \leq i \leq m$, let $R_i = \mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{a}_{i-1}, \mathfrak{a}_{i+1}, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_n]$, and $p_i \in R_i[\mathfrak{a}_i]$ be the univariate polynomial in $\mathfrak{a}_i$ so that $p_i(\mathfrak{a}_i) = r$.

Suppose that for some i, j, $\mathfrak{a}_i = \mathfrak{b}_j$. Then $X - \mathfrak{a}_i = X - \mathfrak{b}_j$ is a common factor of A and B and by Corollary 1, $\mathsf{Res}(A, B)|_{\mathfrak{a}_i = \mathfrak{b}_j} = 0$. This means:

$$r(\mathfrak{a}_1, \ldots, \mathfrak{b}_j, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_n) = p_i(\mathfrak{b}_j) = 0,$$

in R_i and $\mathfrak{a}_i - \mathfrak{b}_j | p_i(\mathfrak{a}_i)$ in $R_i[\mathfrak{a}_i] = \mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_n]$.

The i, j were arbitrarily chosen, so for each i, j, $\mathfrak{a}_i - \mathfrak{b}_j | p_i(\mathfrak{a}_i) = r = \mathsf{Res}(A, B)$ in $R_i[\mathfrak{a}_i] = \mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{b}_n]$, and hence $\prod_{i,j} \mathfrak{a}_i - \mathfrak{b}_j | \mathsf{Res}(A, B)$, as required. □

Let $s(\mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_n) := \prod_{i,j} (\mathfrak{a}_i - \mathfrak{b}_j)$ (as polynomials in $\mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{b}_n]$).

The previous Lemma shows that $\frac{r}{s} \in \mathbb{Z}[\mathfrak{a}_1, \ldots, \mathfrak{b}_n]$. The next Lemma shows that actually $\frac{r}{s} \in \mathbb{Z}$.

Lemma 4 *For* $1 \leq i \leq m$, *holds* $\deg_{\mathfrak{a}_i}(s) = \deg_{\mathfrak{a}_i}(r)$, *and for* $1 \leq j \leq n$, *holds* $\deg_{\mathfrak{b}_j}(s) = \deg_{\mathfrak{b}_j}(r)$.

Proof:Let $\mathfrak{a} = \mathfrak{a}_1, \ldots, \mathfrak{a}_m$ and $\mathfrak{b} = \mathfrak{b}_1, \ldots, \mathfrak{b}_n$. By Equation (6), we have: $\deg_{\mathfrak{a}_i}(\mathfrak{s}_{i,m}(\mathfrak{a})) = 1$ for $1 \leq i \leq m$, and $\deg_{\mathfrak{b}_j}(\mathfrak{s}_{j,n}(\mathfrak{b})) = 1$ for $1 \leq j \leq n$. Denote $\mathsf{Syl}_{i,j}$ be at the i-th line and j-th column of $\mathsf{Syl}(A, B)$.

$$\mathsf{Res}(A, B) = \sum_{\sigma \in \mathfrak{S}_{n+m}} (-1)^{\epsilon(\sigma)} \underbrace{\prod_{1 \leq i \leq n} \mathsf{Syl}_{i,\sigma(i)}}_{n\ first\ lines} \underbrace{\prod_{n+1 \leq j \leq m+n} \mathsf{Syl}_{j,\sigma(j)}}_{m\ last\ lines} \qquad (7)$$

Now if we look at the Sylvester matrix, we see that:

$$\mathsf{Syl}(A, B) = \begin{pmatrix} 1 & -\mathfrak{s}_{1,m}(\mathfrak{a}) & \cdots & (-1)^m \mathfrak{s}_{m,m}(\mathfrak{a}) & & \\ & \ddots & & & \ddots & \\ & & 1 & \cdots & & (-1)^m \mathfrak{s}_{m,m}(\mathfrak{a}) \\ 1 & -\mathfrak{s}_{1,n}(\mathfrak{b}) & \cdots & & (-1)^n \mathfrak{s}_{n,n}(\mathfrak{b}) & \\ & \ddots & & & & \ddots \\ & & 1 & \cdots & & (-1)^n \mathfrak{s}_{n,n}(\mathfrak{b}) \end{pmatrix}$$

Either $\mathsf{Syl}_{i,j} = 0$ or 1 or $\deg_{\mathfrak{a}_u}(\mathsf{Syl}_{i,j}) = 1\ \forall 1 \leq u \leq m$ or $\deg_{\mathfrak{b}_v}(\mathsf{Syl}_{i,j}) = 1\ \forall 1 \leq v \leq n$.

Hence for each permutation $\sigma \in \mathfrak{S}_{n+m}$ such that $\prod_{1\leq i\leq n} \mathsf{Syl}_{i,\sigma(i)} \prod_{m+1\leq j\leq m+n} \mathsf{Syl}_{j,\sigma(j)} \neq 0$, we have:

$$\forall 1 \leq i \leq m,\ \deg_{\mathfrak{a}_i}\left(\prod_{1\leq i\leq n} \mathsf{Syl}_{i,\sigma(i)} \prod_{m+1\leq j\leq m+n} \mathsf{Syl}_{j,\sigma(j)}\right) \leq n$$

$$\forall 1 \leq j \leq n,\ \deg_{\mathfrak{b}_j}\left(\prod_{1\leq i\leq n} \mathsf{Syl}_{i,\sigma(i)} \prod_{m+1\leq j\leq m+n} \mathsf{Syl}_{j,\sigma(j)}\right) \leq m$$

It follows that $\deg_{\mathfrak{a}_i}(\mathsf{Res}(A,B)) \leq n$ and $\deg_{\mathfrak{b}_j}(\mathsf{Res}(A,B)) \leq m$. But by definition, $\deg_{\mathfrak{a}_i}(s) = n$ and $\deg_{\mathfrak{b}_j}(s) = m$, and since $s|r$, the $\leq$ are actually "=", and the lemma follows. □

Let us write $r = C_0 s$, $C_0 \in \mathbb{Z}$. We must show that $\boxed{C_0 = 1}$ to conclude the proof of Theorem 1.

Let us put $\mathfrak{b}_1 = \cdots = \mathfrak{b}_m = 0$. We get: $s(\mathfrak{a}_1, \ldots, \mathfrak{a}_m, 0, \ldots) = (\prod_i \mathfrak{a}_i)^n$ and $r(\mathfrak{a}_1, \ldots, \mathfrak{a}_m, 0, \ldots) \overset{(\star)}{=} C_0 (\prod_i \mathfrak{a}_i)^n$. The Sylvester matrix can be rewritten as follows (where $A = X^m + a_{m-1}X^{m-1} + \cdots + A_1 X + a_0$), and we compute the determinant along the last line. Only one coefficient is not zero, hence:

$$\begin{vmatrix}
1 & \star & \cdots & a_1 & a_0 & & & \\
 & 1 & \star & \cdots & a_1 & a_0 & & \\
 & & \ddots & \ddots & \ddots & \ddots & & \\
 & & 1 & \star & \cdots & a_1 & a_0 & \\
1 & 0 & \cdots & \cdots & & & & \\
 & 1 & 0 & \cdots & \cdots & & & \\
 & & \ddots & \ddots & \ddots & & & \\
 & & & 1 & 0 & \cdots & &
\end{vmatrix}_{m\text{-th column}}
= \underset{\underset{(-1)^n}{\|}}{(-1)^{2m+n}}
\begin{vmatrix}
1 & \star & \cdots & a_2 & a_0 & & \\
 & 1 & \cdots & a_3 & a_1 & a_0 & \\
 & & \ddots & \ddots & \ddots & \ddots & \\
 & 1 & \star & \cdots & a_1 & a_0 & \\
1 & 0 & \cdots & \cdots & & & \\
 & & \ddots & \ddots & \ddots & & \\
 & & & 1 & 0 & \cdots &
\end{vmatrix}_{(m-1)\text{-th column}}
= (-1)^n \underset{\underset{(-1)^{2n}}{\|}}{(-1)^{2(m-1)+n}} \ldots$$

$$\begin{vmatrix} 1 & \star & \cdots & a_3 & a_0 & & & \\ & 1 & \cdots & a_4 & a_1 & a_0 & & \\ & & & \cdots & & \cdots & & \cdots \\ & & 1 & \star & \cdots & a_1 & a_0 & \\ 1 & 0 & \cdots & \cdots & & & & \\ & & & \cdots & \cdots & & & \\ & & & 1 & 0 & \cdots & & \end{vmatrix} = \underset{\substack{\| \\ (-1)^{3n}}}{(-1)^{2n}\,(-1)^{2(m-2)+n}} \begin{vmatrix} 1 & \star & \cdots & a_4 & a_0 & & & \\ & 1 & \cdots & a_5 & a_1 & a_0 & & \\ & & & \cdots & & \cdots & & \cdots \\ & & 1 & \star & \cdots & a_1 & a_0 & \\ 1 & 0 & \cdots & \cdots & & & & \\ & & & \cdots & \cdots & & & \\ & & & 1 & 0 & \cdots & & \end{vmatrix} =$$

$(m-2)$-th column $\qquad\qquad$ $(m-3)$-th column

$$= \underset{\substack{\| \\ (-1)^{4n}}}{(-1)^{3n}(-1)^{2(m-3)+n}} \begin{vmatrix} \cdots & & \\ \ddots & & \\ & & \cdots \end{vmatrix} = \cdots\cdots = (-1)^{(m-1)n} \begin{vmatrix} 1 & a_0 & & & & \\ 0 & a_1 & a_0 & & & \\ & \cdots & \cdots & & & \\ 0 & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ 1 & 0 & \cdots & & 0 & \end{vmatrix}$$

$$= (-1)^{(m-1)n}(-1)^n \begin{vmatrix} a_0 & & & \\ a_1 & a_0 & & \\ \vdots & & \ddots & \\ a_{n-1} & \cdots & a_1 & a_0 \end{vmatrix} = (-1)^{nm} a_0^n .$$

Finally, $r(\mathfrak{a}_1, \ldots, \mathfrak{a}_m, 0, \ldots) = (-1)^{mn} a_0^n$, where $a_0 = (-1)^m \mathfrak{s}_{m,m}(\mathfrak{a}) = (-1)^m \prod_i \mathfrak{a}_i$.

So $r(\mathfrak{a}_1, \ldots, \mathfrak{a}_m, 0, \ldots) = (-1)^{mn}(-1)^{mn}(\prod_i \mathfrak{a}_i)^n = (\prod_i \mathfrak{a}_i)^n \overset{(\star)}{\Rightarrow} \boxed{C_0 = 1}$. $\qquad\square$