4. Resultants

Theorem 4.1. (B.L.van der Waerden, “Algebra, vol.I”, p.102)
Let $a(x), b(x)$ *be two non-constant polynomials in* $K[x]$, K *a field. Then* a *and* b *have a non-constant common factor (i.e. a common root over the algebraic closure of* K*) if and only if there are polynomials* $p(x), q(x) \in K[x]$, *not both equal to 0, with* $\deg(p) < \deg(b), \deg(q) < \deg(a)$, *such that*

$$p(x)a(x) + q(x)b(x) = 0 \ . \qquad (*)$$

Proof: If a and b have the non-constant common factor c, then obviously we can write

$$(b/c) \cdot a - (a/c) \cdot b = 0 \ .$$

On the other hand, assume $(*)$. So we have

$$p(x)a(x) = -q(x)b(x) \ . \qquad (**)$$

We factor the left and right hand sides of $(**)$ into irreducible factors. All the irreducible factors of $a(x)$ must divide the right hand side at least as often as they divide $a(x)$. Yet they cannot divide $q(x)$ as often as they do $a(x)$ because of the degree restriction. Hence at least one irreducible factor of $a(x)$ occurs also in $b(x)$. □

How can we decide the existence of such polynomials p and q as in the previous theorem?

Let $m = \deg(a), n = \deg(b)$ and write

$$a(x) = \sum_{i=0}^{m} a_i x^i, \qquad b(x) = \sum_{i=0}^{n} b_i x^i \; .$$

Ansatz:

$$p(x) = \sum_{i=0}^{n-1} p_i x^i, \qquad q(x) = \sum_{i=0}^{m-1} q_i x^i \; .$$

Then

$$\begin{aligned}
& p \cdot a + q \cdot b = 0 \\
& \Longleftrightarrow \\
& \mathrm{coeff}(p \cdot a, x^i) + \mathrm{coeff}(q \cdot b, x^i) = 0 \quad \forall i \\
& \Longleftrightarrow \\
& \qquad p_{n-1}a_m + q_{m-1}b_n = 0 \\
& \qquad\qquad \vdots \\
& p_0 a_1 + p_1 a_0 + q_0 b_1 + q_1 b_0 = 0 \\
& \qquad p_1 a_0 + q_0 b_0 = 0 \\
& \Longleftrightarrow
\end{aligned}$$

$$(p_{n-1}, \ldots, p_0, q_{m-1}, \ldots, q_0) \cdot \begin{pmatrix} a_m & \cdots & a_0 & & \\ & \ddots & & \ddots & \\ & & a_m & \cdots & a_0 \\ b_n & \cdots & b_0 & & \\ & \ddots & & \ddots & \\ & & b_n & \cdots & b_0 \end{pmatrix} = (0, \ldots, 0)$$

This matrix we will call the determinant of a and b.

Definition 4.2. Let

$$a(x) = \sum_{i=0}^{m} a_i x^i, \qquad b(x) = \sum_{i=0}^{n} b_i x^i$$

be non-constant polynomials in $I[x]$ (I an integral domain) of degree m and n, respectively.
Let $\mathrm{Syl}_x(a,b)$ be the *Sylvester matrix* of a and b, i.e.

$$\mathrm{Syl}_x(a,b) \; = $$

$$\begin{pmatrix}
a_m & a_{m-1} & \cdots & \cdots & \cdots & a_1 & a_0 & 0 & \cdots & \cdots & \cdots & 0 \\
0 & a_m & a_{m-1} & \cdots & \cdots & \cdots & a_1 & a_0 & 0 & \cdots & \cdots & 0 \\
 & & & & \vdots & & & & & & & \\
0 & \cdots & \cdots & \cdots & 0 & a_m & a_{m-1} & \cdots & \cdots & \cdots & a_1 & a_0 \\
- & - & - & - & - & - & - & - & - & - & - & - \\
b_n & b_{n-1} & \cdots & \cdots & \cdots & b_1 & b_0 & 0 & \cdots & \cdots & \cdots & 0 \\
0 & b_n & b_{n-1} & \cdots & \cdots & \cdots & b_1 & b_0 & 0 & \cdots & \cdots & 0 \\
 & & & & \vdots & & & & & & & \\
0 & \cdots & \cdots & \cdots & 0 & b_n & b_{n-1} & \cdots & \cdots & \cdots & b_1 & b_0
\end{pmatrix}.$$

The lines of $\mathrm{Syl}_x(a,b)$ consist of the coefficients of the polynomials $x^{n-1}a(x), \ldots, xa(x), a(x)$ and $x^{m-1}b(x), \ldots, xb(x), b(x)$, i.e. there are n lines of coefficients of a and m lines of coefficients of b. The **resultant** of a and b is the determinant of $\mathrm{Syl}_x(a,b)$; i.e.

$$\mathrm{res}_x(a,b) := \det(\mathrm{Syl}_x(a,b).$$

The *resultant* $\text{res}_x(f, g)$ of two univariate polynomials $f(x), g(x)$ over an integral domain I is the determinant of the *Sylvester matrix* of f and g, consisting of shifted lines of coefficients of f and g. $\text{res}_x(f, g)$ is a constant in I. For $m = \deg(f), n = \deg(g)$, we have $\text{res}_x(f, g) = (-1)^{mn}\text{res}_x(g, f)$, i.e. the resultant is symmetric up to sign. If $a_1, \ldots, a_m$ are the roots of f, and $b_1, \ldots, b_n$ are the roots of g in their common splitting field, then

$$\text{res}_x(f, g) = \text{lc}(f)^n \text{lc}(g)^m \prod_{i=1}^{m} \prod_{j=1}^{n} (a_i - b_j).$$

The resultant has the important property that, for non–zero polynomials f and g, $\text{res}_x(f, g) = 0$ if and only if f and g have a common root, and in fact, f and g have a non-constant common divisor in $K[x]$, where K is the quotient field of I. If f and g have positive degrees, then there exist polynomials $a(x), b(x)$ over I such that $af + bg = \text{res}_x(f, g)$. The *discriminant* of $f(x)$ is

$$\text{discr}_x(f) = (-1)^{m(m-1)/2} \text{lc}(f)^{2(m-1)} \prod_{i \neq j} (a_i - a_j).$$

We have the relation $\text{res}_x(f, f') = (-1)^{m(m-1)/2}\text{lc}(f)\text{discr}_x(f)$, where f' is the derivative of f.

Also if $f(x), g(x)$ are polynomials over a field K, then

$$\text{res}_x(f, g) = p \cdot f + q \cdot g$$

for some $p(x), q(x) \in K[x]$.
(compare Cox,Little,O'Shea, "Ideals, Varieties, and Algorithms", p.152)

Lemma 4.3. (Lemma 4.3.1 in Winkler, “Computer Algebra”)
Let I, J be integral domains, ϕ a homomorphism from I into J. The homomorphism from $I[x]$ into $J[x]$ induced by ϕ will also be denoted ϕ, i.e. $\phi(\sum_{i=0}^{m} c_i x^i) = \sum_{i=0}^{m} \phi(c_i)x^i$. Let $a(x), b(x)$ be polynomials in $I[x]$. If $\deg(\phi(a)) = \deg(a)$ and $\deg(\phi(b)) = \deg(b) - k$, then $\phi(\mathrm{res}_x(a,b)) = \phi(\mathrm{lc}(a))^k \mathrm{res}_x(\phi(a), \phi(b))$.

Lemma 4.4. (Lemma 4.3.2 in Winkler, “Computer Algebra”)
Let $a(x_1, \ldots, x_r) = \sum_{i=0}^{m} a_i(x_1, \ldots, x_{r-1})x_r^i$, $b(x_1, \ldots, x_r) = \sum_{i=0}^{n} b_i(x_1, \ldots, x_{r-1})x_r^i$ be polynomials in $\mathbb{Z}[x_1, \ldots, x_r]$. Let $d = \max_{0 \le i \le m} \mathrm{norm}(a_i)$, $e = \max_{0 \le i \le n} \mathrm{norm}(b_i)$, α an integer coefficient in $\mathrm{res}_{x_r}(a,b)$. Then $|\alpha| \le (m+n)! d^n e^m$.

Are $a(x_1, \ldots, x_r), b(x_1, \ldots, x_r) \in \mathbb{Z}[x_1, \ldots, x_r]$, then the resultant of a and b w.r.t. the variable x_r can be computed by the following modular algorithm.

The subalgorithm RES_MODp computes multivariate resultants over $\mathbb{Z}_p$ by evaluation homomorphisms.

---

algorithm RES_MOD(**in:** a, b; **out:** c);
[$a, b \in \mathbb{Z}[x_1, \ldots, x_r]$, $r \geq 1$, a and b have positive degree in x_r;
$c = \mathrm{res}_{x_r}(a, b)$.]
(1) $m := \deg_{x_r}(a)$; $n := \deg_{x_r}(b)$;
$d := \max_{0 \leq i \leq m} \mathrm{norm}(a_i)$; $e := \max_{0 \leq i \leq n} \mathrm{norm}(b_i)$;
$P := 1$; $c := 0$; $B := 2(m+n)!d^n e^m$;
(2) **while** $P \leq B$ **do**
{$p :=$ a new prime such that $\deg_{x_r}(a) = \deg_{x_r}(a_{(p)})$ and
$\deg_{x_r}(b) = \deg_{x_r}(b_{(p)})$;
$c_{(p)} :=$ RES_MODp($a_{(p)}, b_{(p)}$);
$c :=$ CRA_2($c, c_{(p)}, P, p$);
[for $P = 1$ the output is simply $c_{(p)}$,
otherwise CRA_2 is actually applied to
the coefficients of c and $c_{(p)}$]
$P := P \cdot p$ };
return □

---

algorithm RES_MODp(**in:** a, b; **out:** c);
[$a, b \in \mathbb{Z}_p[x_1, \ldots, x_r]$, $r \geq 1$, a and b have positive degree in x_r;
$c = \mathrm{res}_{x_r}(a, b)$.]
(0) **if** $r = 1$ **then** { $c :=$ last element of PRS_SR(a, b); **return** };
(1) $m_r := \deg_{x_r}(a)$; $n_r := \deg_{x_r}(b)$;
$m_{r-1} := \deg_{x_{r-1}}(a)$; $n_{r-1} := \deg_{x_{r-1}}(b)$;
$B := m_r n_{r-1} + n_r m_{r-1} + 1$;
$D(x_{r-1}) := 1$; $c(x_1, \ldots, x_{r-1}) := 0$; $\beta := -1$;
(2) **while** $\deg(D) \leq B$ **do**
(2.1) {$\beta := \beta + 1$; [if $\beta = p$ stop and report failure]
if $\deg_{x_r}(a_{x_{r-1}=\beta}) < \deg_{x_r}(a)$ or $\deg_{x_r}(b_{x_{r-1}=\beta}) < \deg_{x_r}(a)$
then goto (2.1);
$c_{(\beta)}(x_1, \ldots, x_{r-2}) :=$ RES_MODp($a_{x_{r-1}=\beta}, b_{x_{r-1}=\beta}$);
$c := (c_{(\beta)}(x_1, \ldots, x_{r-2}) - c(x_1, \ldots, x_{r-2}, \beta))D(\beta)^{-1}D(x_{r-1})$
$+c(x_1, \ldots, x_{r-1})$;
[so c is the result of the Newton interpolation]
$D(x_{r-1}) := (x_{r-1} - \beta)D(x_{r-1})$ };
return □

Solving systems of algebraic equations by resultants

Theorem 4.5. (Theorem 4.3.3 in Winkler, “Computer Algebra”)
Let K be an algebraically closed field, let

$$a(x_1,\ldots,x_r)=\sum_{i=0}^{m} a_i(x_1,\ldots,x_{r-1})x_r^i,$$
$$b(x_1,\ldots,x_r)=\sum_{i=0}^{n} b_i(x_1,\ldots,x_{r-1})x_r^i$$

be elements of $K[x_1,\ldots,x_r]$ of positive degrees m and n in x_r, and let $c(x_1,\ldots,x_{r-1})=\mathrm{res}_{x_r}(a,b)$. If $(\alpha_1,\ldots,\alpha_r)\in K^r$ is a common root of a and b, then $c(\alpha_1,\ldots,\alpha_{r-1})=0$. Conversely, if $c(\alpha_1,\ldots,\alpha_{r-1})=0$, then one of the following holds:
(a) $a_m(\alpha_1,\ldots,\alpha_{r-1})=b_n(\alpha_1,\ldots,\alpha_{r-1})=0$,
(b) for some $\alpha_r\in K$, $(\alpha_1,\ldots,\alpha_r)$ is a common root of a and b.

This theorem suggests a method for determining the solutions of a system of algebraic, i.e. polynomial, equations over an algebraically closed field. Suppose, for example, that a system of three algebraic equations is given as

$$a_1(x,y,z)=a_2(x,y,z)=a_3(x,y,z)=0.$$

Let, e.g.,

$$\begin{aligned}
b(x)&=\mathrm{res}_z(\mathrm{res}_y(a_1,a_2),\mathrm{res}_y(a_1,a_3)),\\
c(y)&=\mathrm{res}_z(\mathrm{res}_x(a_1,a_2),\mathrm{res}_x(a_1,a_3)),\\
d(z)&=\mathrm{res}_y(\mathrm{res}_x(a_1,a_2),\mathrm{res}_x(a_1,a_3)).
\end{aligned}$$

In fact, we might compute these resultants in any other order. By Theorem 4.3.3, all the roots $(\alpha_1,\alpha_2,\alpha_3)$ of the system satisfy $b(\alpha_1)=c(\alpha_2)=d(\alpha_3)=0$. So if there are finitely many solutions, we can check for all of the candidates whether they actually solve the system.

Unfortunately, there might be solutions of b, c, or d, which cannot be extended to solutions of the original system, as we can see from the following example.

Example 4.6. Consider the system of algebraic equations

$$\begin{aligned} a_1(x,y,z) &= 2xy + yz - 3z^2 = 0, \\ a_2(x,y,z) &= x^2 - xy + y^2 - 1 = 0, \\ a_3(x,y,z) &= yz + x^2 - 2z^2 = 0. \end{aligned}$$

We compute

$$\begin{aligned} b(x) &= \mathrm{res}_z(\mathrm{res}_y(a_1,a_3), \mathrm{res}_y(a_2,a_3)) \\ &= x^6(x-1)(x+1)(127x^4 - 167x^2 + 4), \\ c(y) &= \mathrm{res}_z(\mathrm{res}_x(a_1,a_3), \mathrm{res}_x(a_2,a_3)) \\ &= (y-1)^3(y+1)^3(3y^2-1)(127y^4 - 216y^2 + 81)\cdot \\ &\qquad (457y^4 - 486y^2 + 81), \\ d(z) &= \mathrm{res}_y(\mathrm{res}_x(a_1,a_2), \mathrm{res}_x(a_1,a_3)) \\ &= 5184z^{10}(z-1)(z+1)(127z^4 - 91z^2 + 16). \end{aligned}$$

All the solutions of the system, e.g. $(1,1,1)$, have coordinates which are roots of b, c, d. But there is no solution of the system having y–coordinate $1/\sqrt{3}$. So not every root of these resultants can be extended to a solution of the system. □