Resultants and Elimination Theory

Pramana Saldin

April 18, 2024

1 Motivation and definitions

Let k be an algebraically closed field (e.g. $\mathbb{C}$). Suppose $f(x) = a_0 + a_1x + \cdots + a_rx^r$ and $g(x) = b_0 + b_1x + \cdots + b_sx^s$ are polynomials in $k[x]$ with roots $\{\alpha_i\}_{i=1}^r$ and $\{\beta_j\}_{j=1}^s$ respectively. Our goal is to find out if f and g share a common root $\alpha_i = \beta_j$. With resultants, we are able to check if f and g share a root *just using their coefficients.* Using these ideas, we can prove a more general theorem in algebraic geometry called the *fundamental theorem of elimination theory.* This introductory discussion was inspired by [M´].

1.1 Constructing the resultant

Since we are only concerned about the roots of f and g, we will temporarily ignore the leading coefficient and assume that f and g are monic.

Suppose f and g share a root $\alpha_1 = \beta_1$. The idea is to find some kind of linear dependence between f and g. Define the following two polynomials by removing the factor of their shared root:

$$f_1(x) \coloneqq \prod_{i=2}^{r}(x - \alpha_i), \qquad g_1(x) \coloneqq \prod_{j=2}^{s}(x - \beta_i).$$

Then

$$f_1(x)g(x) = \prod_{i=2}^{r}(x - \alpha_i)\prod_{j=1}^{s}(x - \beta_i) = \prod_{i=1}^{r}(x - \alpha_i)\prod_{j=2}^{s}(x - \beta_i) = f(x)g_1(x).$$

So

$$f_1g - fg_1$$

is the zero polynomial. By expanding f_1 and g_1 in terms of their coefficients, we see that this makes the set

$$\{f(x), xf(x), \ldots, x^{s-1}f(x), g(x), xg(x), \ldots, x^{r-1}g(x)\} \tag{1.1}$$

linearly dependent.

Conversely, suppose that the set in Equation 1.1 is linearly dependent. That is, there are numbers $c_0, \ldots, c_{s-1}, d_0, \ldots, d_{r-1} \in k$ such that

$$0 = \sum_{k=0}^{s-1} c_kx^kf(x) + \sum_{k=0}^{r-1} d_kx^kg(x) = \underbrace{\left(\sum_{k=0}^{s-1} c_kx^k\right)}_{(1)} f(x) + \underbrace{\left(\sum_{k=0}^{r-1} d_kx^k\right)}_{(2)} g(x).$$

Define the polynomial (1) as g_1 and (2) as f_1. The claim is that f and g share a root. Indeed, there are s roots of g (with multiplicity) and g_1 could have at most $s - 1$ of them. Therefore, f has at least one root of g.

With the basis $\{1, x, \dots, x^{r+s-1}\}$, the linear dependence of the set (1.1) is equivalent to the condition

$$\operatorname{rk}\begin{bmatrix} a_0 & a_1 & \cdots & a_r & & & \\ & a_0 & a_1 & \cdots & a_r & & \\ & & \ddots & \ddots & & \ddots & \\ & & & a_0 & a_1 & \cdots & a_r \\ b_0 & b_1 & \cdots & b_s & & & \\ & b_0 & b_1 & \cdots & b_s & & \\ & & \ddots & \ddots & & \ddots & \\ & & & b_0 & b_1 & \cdots & b_s \end{bmatrix} < r+s, \tag{1.2}$$

where there are s rows of a_i and r rows of b_i.

We check for linear dependence the common way we expect in a linear algebra class: by taking the determinant.

Definition 1.1 (Resultant)**.** Define the determinant of the matrix (1.2) as the **resultant** of f and g, which we denote $\operatorname{Res}(f, g)$. The resultant vanishes if and only if f and g share a common root.

Example 1.1. Let $f(z) = z^2 - 1$ and $g(z) = z - 1$ be polynomials in $\mathbb{C}[z]$. Then

$$\operatorname{Res}(f, g) = \det\begin{bmatrix} 1 & 0 & -1 \\ 1 & -1 & \\ & 1 & -1 \end{bmatrix} = 0.$$

So f and g share a root. Indeed, they share the root $\alpha = 1$.

Remark 1.2. Notice that the resultant depends on the degree of f and g. For example, $f(x) = 0x^2 + x - 1$ and $g(x) = 0x^2 + x + 1$ do not share any roots, but if we treat them as degree 2 polynomials, their resultant is

$$\det\begin{bmatrix} 0 & 1 & -1 & \\ & 0 & 1 & -1 \\ 0 & 1 & 1 & \\ & 0 & 1 & 1 \end{bmatrix} = 0.$$

Therefore, it would be more precise to write their degrees as parameters $\operatorname{Res}_{r,s}(f, g)$, but we omit them.

Let $R\colon (g_1, f_1) \mapsto fg_1 + gf_1$ be a linear map from polynomials of degree $\leq s-1$ and $\leq r-1$ respectively to polynomials of degree $\leq r+s-1$. We interpret the resultant as telling us whether the image of R has dimension less than $r+s$ (that is, whether R is surjective or not). This helps explain why the resultant in Remark 1.2 evaluated to zero.

1.2 The difference of roots form of the resultant

Another common formula used when introducing resultants is the following:

Theorem 1.3. Suppose $f(x) = a_0 + a_1 x + \cdots + a_r x^r$ and $g(x) = b_0 + b_1 x + \cdots + b_s x^s$ are polynomials in $k[x]$ with roots $\{\alpha_i\}_{i=1}^{r}$ and $\{\beta_j\}_{j=1}^{s}$ respectively. Then

$$\operatorname{Res}(f, g) = \pm a_r^s b_s^r \prod_{\substack{1 \leq i \leq r \\ 1 \leq j \leq s}} (\alpha_i - \beta_j).[1] \tag{1.3}$$

[1] Since we are mainly concerned with when the resultant *vanishes*, sign changes may safely be ignored.

Lemma 1.4 (Vandermonde matrix and determinant)**.** Let $x_0, \dots, x_n \in k$. The **Vandermonde matrix** is defined as

$$\operatorname{Vand}(x_0, x_1, \cdots, x_n) = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{bmatrix}.$$

Its determinant is given by

$$\det(\operatorname{Vand}(x_0, x_1, \cdots, x_n)) = \prod_{i<j}(x_j - x_i).$$

Proof. Let $V = \operatorname{Vand}(x_0, x_1, \cdots, x_n)$. By the Leibniz formula for the determinant, we should expect $f(x_0, \dots, x_n) \coloneqq \det V$ to be a polynomial with total degree $\frac{n(n+1)}{2}$ in the variables $x_0, \dots, x_n$. Since replacing x_j with x_i yields a zero determinant, $(x_j - x_i) \mid f(x_0, \dots, x_n)$ for all $j \neq i$. As a result,

$$\prod_{i<j}(x_j - x_i) \mid f(x_0, \dots, x_n),$$

so we may write

$$f(x_0, \dots, x_n) = Q(x_0, \dots, x_n) \prod_{i<j}(x_j - x_i), \qquad Q \in k[x_0, \dots, x_n].$$

Checking the degree of this polynomial, we find that Q is a nonzero constant. To show $Q = 1$, observe that the coefficient of $x_1 x_2^2 \cdots x_n^n$ in $f(x_0, \dots, x_n)$ is 1. □

Of course, if we have the determinant of the transpose of the Vandermonde matrix, we will also get the same determinant.

Proof of Theorem 1.3. Assume f and g are both monic ($a_r = b_s = 1$). Consider one of the roots of f, α_i. When we look at the resulting vector from

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_r & & & \\ & a_0 & a_1 & \cdots & a_r & & \\ & & \ddots & \ddots & & \ddots & \\ & & & a_0 & a_1 & \cdots & a_r \\ b_0 & b_1 & \cdots & b_s & & & \\ & b_0 & b_1 & \cdots & b_s & & \\ & & \ddots & \ddots & & \ddots & \\ & & & b_0 & b_1 & \cdots & b_s \end{bmatrix} \begin{bmatrix} 1 \\ \alpha_i \\ \alpha_i^2 \\ \vdots \\ \alpha_i^{r+s-1} \end{bmatrix},$$

we find the first entry is $a_0 + a_1\alpha_i + a_2\alpha_i^2 + \cdots + a_r\alpha_i^r = f(\alpha_i)$, the second is $a_0\alpha_i + a_1\alpha_i^2 + a_2\alpha_i^3 + \cdots + a_r\alpha_i^{r+1} = \alpha_i f(\alpha_i)$, and so on. Since α_i is a root of f, the first s entries are zero. For the next r entries, we get $b_0 + b_1\alpha_i + b_2\alpha_i^2 + \cdots + b_s\alpha_i^s = g(\alpha_i)$, $b_0\alpha_i + b_1\alpha_i^2 + b_2\alpha_i^3 + \cdots + b_s\alpha_i^{s+1} = \alpha_i g(\alpha_i)$, and so on. Combining these, we have that

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_r & & & \\ & a_0 & a_1 & \cdots & a_r & & \\ & & \ddots & \ddots & & \ddots & \\ & & & a_0 & a_1 & \cdots & a_r \\ b_0 & b_1 & \cdots & b_s & & & \\ & b_0 & b_1 & \cdots & b_s & & \\ & & \ddots & \ddots & & \ddots & \\ & & & b_0 & b_1 & \cdots & b_s \end{bmatrix} \begin{bmatrix} 1 \\ \alpha_i \\ \alpha_i^2 \\ \vdots \\ \alpha_i^{r+s-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ g(\alpha_i) \\ \alpha_i g(\alpha_i) \\ \vdots \\ \alpha_i^{s-1} g(\alpha_i) \end{bmatrix}$$

Similarly, for the roots β_j of g,

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_r & & & \\ & a_0 & a_1 & \cdots & a_r & & \\ & & \ddots & \ddots & & \ddots & \\ & & & a_0 & a_1 & \cdots & a_r \\ b_0 & b_1 & \cdots & b_s & & & \\ & b_0 & b_1 & \cdots & b_s & & \\ & & \ddots & \ddots & & \ddots & \\ & & & b_0 & b_1 & \cdots & b_s \end{bmatrix} \begin{bmatrix} 1 \\ \beta_j \\ \beta_j^2 \\ \vdots \\ \beta_j^{r+s-1} \end{bmatrix} = \begin{bmatrix} f(\beta_j) \\ \beta_j f(\beta_j) \\ \vdots \\ \beta_j^{r-1} f(\beta_j) \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Combining these two results into a single matrix, we have

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_r & & & \\ & a_0 & a_1 & \cdots & a_r & & \\ & & \ddots & \ddots & & \ddots & \\ & & & a_0 & a_1 & \cdots & a_r \\ b_0 & b_1 & \cdots & b_s & & & \\ & b_0 & b_1 & \cdots & b_s & & \\ & & \ddots & \ddots & & \ddots & \\ & & & b_0 & b_1 & \cdots & b_s \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 \\ \alpha_1 & \cdots & \alpha_r & \beta_1 & \cdots & \beta_s \\ \alpha_1^2 & \cdots & \alpha_r^2 & \beta_1^2 & \cdots & \beta_s^2 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{r+s-1} & \cdots & \alpha_r^{r+s-1} & \beta_1^{r+s-1} & \cdots & \beta_s^{r+s-1} \end{bmatrix}$$

$$= \left[\begin{array}{ccc|ccc} & & & f(\beta_1) & \cdots & f(\beta_s) \\ & & & \vdots & \ddots & \vdots \\ & & & \beta_1^{s-1} f(\beta_1) & \cdots & \beta_s^{s-1} f(\beta_s) \\ \hline g(\alpha_1) & \cdots & g(\alpha_r) & & & \\ \vdots & \ddots & \vdots & & & \\ \alpha_1^{r-1} g(\alpha_1) & \cdots & \alpha_r^{r-1} g(\alpha_r) & & & \end{array}\right]$$

Let A be the first matrix on the left-hand side (this is the matrix we want the determinant of). The second matrix on the left-hand side is a Vandermonde matrix, so we can compute its determinant using Lemma 1.4. The matrix on the right-hand side has blocks that are Vandermonde after factoring out the values on each column. Hence,

$$\begin{aligned} &\det A \cdot \det\left(\mathrm{Vand}(\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s)\right) \\ &\quad = -\left(g(\alpha_1) \cdots g(\alpha_r) \cdot \det\left(\mathrm{Vand}(\alpha_1, \dots, \alpha_r)\right) \cdot f(\beta_1) \cdots f(\beta_s) \cdot \det\left(\mathrm{Vand}(\beta_1, \dots, \beta_s)\right)\right). \end{aligned}$$

We have

$$\begin{aligned} \det\left(\mathrm{Vand}(\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s)\right) &= \left(\prod_{1 \le i_1 < i_2 \le r} (\alpha_{i_2} - \alpha_{i_1})\right) \left(\prod_{\substack{1 \le i \le r \\ 1 \le j \le s}} (\beta_j - \alpha_i)\right) \left(\prod_{1 \le j_1 < j_2 \le s} (\beta_{j_2} - \beta_{j_1})\right) \\ \det\left(\mathrm{Vand}(\alpha_1, \dots, \alpha_r)\right) &= \prod_{1 \le i_1 < i_2 \le r} (\alpha_{i_2} - \alpha_{i_1}) \\ \det\left(\mathrm{Vand}(\beta_1, \dots, \beta_s)\right) &= \prod_{1 \le j_1 < j_2 \le s} (\beta_{j_2} - \beta_{j_1}) \end{aligned}$$

by the Vandermonde matrix formula, and

$$\begin{aligned} g(\alpha_1) \cdots g(\alpha_r) &= \prod_{i=1}^{r} \prod_{j=1}^{s} (\alpha_i - \beta_j) \\ f(\beta_1) \cdots f(\beta_s) &= \prod_{j=1}^{s} \prod_{i=1}^{r} (\beta_j - \alpha_i), \end{aligned}$$

from writing f and g in terms of their factors. So

$$\begin{aligned}
\det A &= -\frac{g(\alpha_1)\cdots g(\alpha_r)\cdot \det\left(\mathrm{Vand}(\alpha_1,\dots,\alpha_r)\right)\cdot f(\beta_1)\cdots f(\beta_s)\cdot \det\left(\mathrm{Vand}(\beta_1,\dots,\beta_s)\right)}{\mathrm{Vand}(\alpha_1,\dots,\alpha_r,\beta_1,\dots,\beta_s)} \\
&= -\frac{\left(\prod_{\substack{1\le i\le r\\ 1\le j\le s}}(\alpha_i-\beta_j)\right)\left(\prod_{1\le i_1<i_2\le r}(\alpha_{i_2}-\alpha_{i_1})\right)\left(\prod_{\substack{1\le i\le r\\ 1\le j\le s}}(\beta_j-\alpha_i)\right)\left(\prod_{1\le j_1<j_2\le s}(\beta_{j_2}-\beta_{j_1})\right)}{\left(\prod_{1\le i_1<i_2\le r}(\alpha_{i_2}-\alpha_{i_1})\right)\left(\prod_{\substack{1\le i\le r\\ 1\le j\le s}}(\beta_j-\alpha_i)\right)\left(\prod_{1\le j_1<j_2\le s}(\beta_{j_2}-\beta_{j_1})\right)} \\
&= -\prod_{\substack{1\le i\le r\\ 1\le j\le s}}(\alpha_i-\beta_j) \\
&= \pm\prod_{\substack{1\le i\le r\\ 1\le j\le s}}(\beta_j-\alpha_i) \qquad \square
\end{aligned}$$

2 Application: solving systems in $k[x, y]$

Consider two polynomials $f(x, y)$ and $g(x, y)$. Suppose we want to solve the system

$$\begin{cases} f(x,y) = 0, \\ g(x,y) = 0. \end{cases}$$

If we fix y, then we can take the resultant of f and g with respect to x. Denote this $\mathrm{Res}_x(f, g)$. This becomes a polynomial in y, which only has finitely many roots $y_1, \dots, y_n$. These roots correspond to where $f(x, y_i)$ and $g(x, y_i)$ could possibly share a root. Now to find all solutions, we need to solve

$$\begin{cases} f(x,y_i) = 0, \\ g(x,y_i) = 0, \end{cases} \qquad 1 \le i \le n.$$

That is, the common roots of $f(x, y_i)$ and $g(x, y_i)$.

Example 2.1. Consider $f(x, y) = x^2 - y$ (a parabola) and $g(x, y) = x^2 + y^2 - 1$ (a circle).

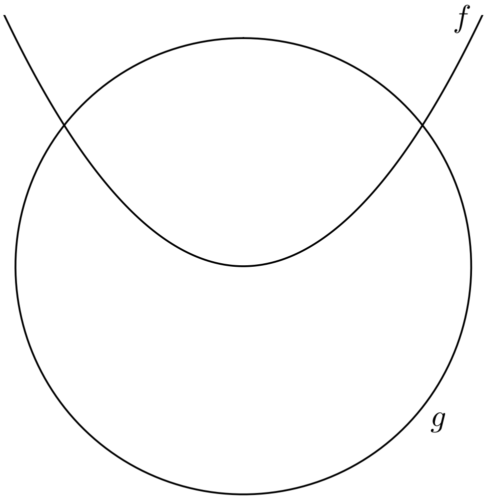

Figure 1: Plot of $f(x, y) = 0$ and $g(x, y) = 0$.

Their resultant is

$$\mathrm{Res}_x(f,g) = \det \begin{bmatrix} -y & 0 & 1 & \\ & -y & 0 & 1 \\ y^2-1 & 0 & x^2 & \\ & y^2-1 & 0 & x^2 \end{bmatrix} = y^4 + 2y^3 - y^2 - 2y + 1.$$

Noticing that this polynomial is $(y^2+y-1)^2$, we have that its roots are

$$y_1 = \frac{1}{2}(-1-\sqrt{5}), \quad y_2 = \frac{1}{2}(\sqrt{5}-1).$$

For y_1, $f(x,y_1)$ and $g(x,y_1)$ share the roots

$$x = \pm i\sqrt{\frac{1}{2}(1+\sqrt{5})},$$

and for y_2, $f(x,y_2)$ and $g(x,y_2)$ share the roots

$$x = \pm\sqrt{\frac{1}{2}(\sqrt{5}-1)}.$$

So the solutions to $f(x,y) = g(x,y) = 0$ are

$$\left\{ \left(-i\sqrt{\frac{1}{2}(1+\sqrt{5})}, \frac{1}{2}(-1-\sqrt{5})\right), \left(i\sqrt{\frac{1}{2}(1+\sqrt{5})}, \frac{1}{2}(-1-\sqrt{5})\right), \right.$$
$$\left. \left(-\sqrt{\frac{1}{2}(\sqrt{5}-1)}, \frac{1}{2}(\sqrt{5}-1)\right), \left(\sqrt{\frac{1}{2}(\sqrt{5}-1)}, \frac{1}{2}(\sqrt{5}-1)\right) \right\}.$$

Remark 2.2. Of course, we could have gotten the same solution by noticing we should plug in $x^2 = y$ into g. However, the power of using the resultants is that this method works for any system of two polynomials in two variables.

3 The algebro-geometric interpretation of the resultant

Let $\mathbb{A}^n$ be the set of all n-tuples of elements of k. We call this **affine space**. For any ideal $I \subseteq k[x_1,\dots,x_n]$, we can define a **variety** created from I as

$$\mathbf{V}(I) := \{(P_1,\dots,P_n) \in \mathbb{A}^n \mid f(P_1,\dots,P_n) = 0 \text{ for all } f \in I\}.$$

We define closed sets in the **Zariski topology** on $\mathbb{A}^n$ as zero sets of polynomials. For a set of polynomials $\{f_1,\dots,f_n\}$, the variety formed from these polynomials is the variety created from the ideal $(f_1,\dots,f_n)$, which we denote

$$\mathbf{V}(f_1,\dots,f_n) := \mathbf{V}((f_1,\dots,f_n)).$$

Let $S \subseteq \mathbb{A}^n$ be a set. Then

$$I(S) := \{f \in k[x_1,\dots,x_n] \mid f(P) = 0 \text{ for all } P \in S\}.$$

3.1 Projections and constructible sets

Let f, g be polynomials in $k[a_0,\dots,a_r,x]$. In this case, f and g cut out a variety $\mathbf{V}(f,g)$ in $\mathbb{A}^{(r+1)+1}$.

Consider the projection morphism

$$\pi\colon \mathbb{A}^{(r+1)+1} \to \mathbb{A}^{r+1},$$
$$(a_0,\dots,a_r,x) \mapsto (a_0,\dots,a_r).$$

We want to know what images of varieties in $\mathbb{A}^{(r+1)+1}$ look like.

Definition 3.1. Let X be a topological space. Consider the collection $\mathcal{C}$ of sets which consist of open sets in X, and finite intersections and unions of open sets. Sets in $\mathcal{C}$ are called **constructible**.

Example 3.1. A set being constructible in $\mathbb{A}^n$ translates to it consisting of the solutions and non-solutions to a finite number of polynomials in $k[x_0, \ldots, x_n]$ with "and" and "or" conjunctions, e.g.

$$\big\{(x, y) \mid x = 0 \text{ and } y = 0, \text{ or } x + y \neq 0 \text{ and } x^2 - 2 = 0\big\}$$

forms a constructible set in $\mathbb{A}^2$.

Example 3.2. To motivate the following proposition, we consider what images of closed sets under π look like.

(a) Let $f(x) = az + b$ be a linear equation in $\mathbb{C}[z]$. In $\mathbb{C}[a, b, z]$,

$$\mathbf{V}(az + b) = \{(a, b, z) \mid az + b = 0\},$$

and its projection under $\pi\colon \mathbb{A}^3 \to \mathbb{A}^2\colon (a, b, z) \mapsto (a, b)$ is the set

$$\pi(\mathbf{V}(az + b)) = \{(a, b) \mid \text{there exists } z \in k \text{ such that } az + b = 0\}.$$

A linear equation has a zero when it intersects the x-axis, which happens if and only if $a \neq 0$ or $a = b = 0$. The set

$$\{(a, b) \mid a \neq 0, \text{ or } a = 0 \text{ and } b = 0\}$$

is constructible.

(b) Let $f(z) = az^2 + bz + c$, $g(z) = dz + e$ be polynomials in $\mathbb{C}[z]$. Then

$$\operatorname{Res}(f, g) = \det \begin{bmatrix} a & b & c \\ d & e & \\ & d & e \end{bmatrix} = ae^2 - bde + cd^2.$$

Viewing f and g as polynomials in $\mathbb{C}[a, b, c, d, e, z]$,

$$\pi(\mathbf{V}(f, g)) = \big\{(a, b, c, d, e) \mid \text{there exists } z \in k \text{ such that } az^2 + bz + c = 0 \text{ and } dz + e = 0\big\}.$$

By the properties of the resultant,

$$\pi(\mathbf{V}(f, g)) \subseteq \mathbf{V}(\operatorname{Res}(f, g)).$$

This is not equality because, for example, with $(a, b, c, d, e) = (0, 0, 1, 0, 0)$,

$$\operatorname{Res}(0z^2 + 0z + 1, 0z + 0) = 0,$$

but there is no z such that $0z^2 + 0z + 1 = 0$.

To get the actual description of the image, we need to exclude cases based on the degrees of f and g. We notice that if $a \neq 0$ and $d \neq 0$, then $\operatorname{Res}(az^2 + bz + c, dz + e)$ precisely says when f and g share a root. Now, if $a = 0$ and $b \neq 0$ and $d \neq 0$, then $\operatorname{Res}(bz + c, dz + e)$ tells us when f and g share a root. The last case is where $(a, b, c, d, e) = (0, 0, 0, 0, 0)$, which trivially implies f and g share a root. In summary,

$$\begin{aligned}\pi(\mathbf{V}(f, g)) = \Big\{&(a, b, c, d, e) \mid a \neq 0 \text{ and } d \neq 0 \text{ and } \operatorname{Res}(az^2 + bz + c, dz + e) = 0 \\ &\text{or } a = 0 \text{ and } b \neq 0 \text{ and } d \neq 0 \text{ and } \operatorname{Res}(bz + c, dz + e) = 0 \\ &\text{or } a = b = c = d = e = 0\Big\}.\end{aligned}$$

Once again, the image of a closed set is constructible!

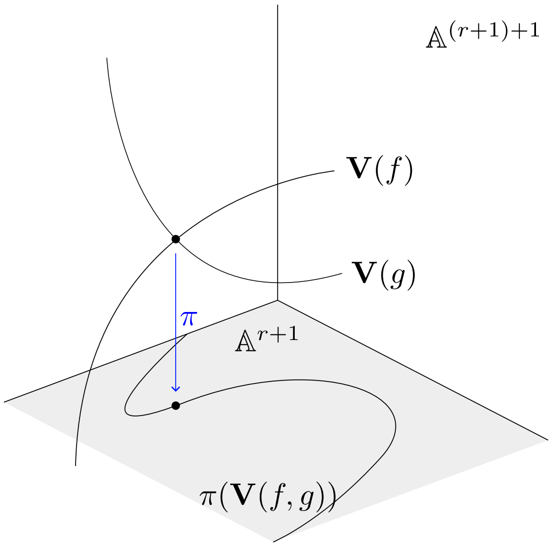

Figure 2: A point $(a_0, \dots, a_n) \in \pi(\mathbf{V}(f,g))$ represents f and g sharing a root.

We notice that each point of $\mathbb{A}^{r+1}$ corresponds to the coefficients of some polynomials, and it only appears in the projection of a variety if those polynomials share roots. This is visualized in Figure 2.

From these examples, we should be able to show that $\pi(\mathbf{V}(f_1, \dots, f_n))$ is constructible with resultants, following a similar proof (now we are testing for when $f_1, \dots, f_n$ share a root, which happens when the resultants for all pairs of f_i vanish). This leads us to the following proposition.

Proposition 3.3. π sends closed sets to constructible sets.

We omit the proof of the proposition. By extending the ideas in Example 3.2 to polynomials of arbitrary degree, this result follows.

3.2 Projectivizing

Working in projective space allows us to avoid issues where we may take a resultant of a higher degree than the polynomial.

As a reminder, the **projective space** $\mathbb{P}^n$ over a field k is defined as $(k^{n+1} \setminus \{\mathbf{0}\})/\sim$, where $\sim$ is the equivalence relation $(x_0, \dots, x_n) \sim (y_0, \dots, y_n)$ if $(x_0, \dots, x_n) = (\lambda y_0, \dots, \lambda y_n)$ for some nonzero $\lambda \in k$. We denote projective coordinates as $(x_0 : \dots : x_n)$.

We would like the resultant to also work over projective space $\mathbb{P}^n$ by using homogeneous polynomials. Let S_d be the set of homogeneous polynomials of degree d. Let $S = \bigoplus_{d=0}^{\infty} S_d$ be the graded ring of polynomials under polynomial addition and multiplication. We define the **Zariski topology** on $\mathbb{P}^n$ as zero sets of homogeneous polynomials.

The homogeneous resultant looks exactly like the non-homogeneous one. Indeed, if $f(x,y) = a_0 y^r + a_1 x y^{r-1} + \dots + a_r x^r$, $g(x,y) = a_0 y^s + a_1 x y^{s-1} + \dots + a_s x^s$ have degrees r and s respectively, and share a root, the set

$$\left\{y^{s-1} f(x,y), x y^{s-2} f(x,y), \dots, x^{s-1} f(x,y), y^{r-1} g(x,y), x y^{r-2} g(x,y), \dots, x^{r-1} g(x,y)\right\}$$

is linearly dependent, and using the basis $\left\{y^{r+s-1}, x y^{r+s-2}, \dots, x^{r+s-1}\right\}$, the matrix whose determinant detects when f and g share a root is precisely (1.2).

The projective resultant has the convenient property that it can detect if homogeneous polynomials f and g share a root “at infinity”.

Example 3.4. When introducing resultants, we showed $f(x) = 0x^2 + x - 1$ and $g(x) = 0x^2 + x + 1$ have a zero resultant but share no roots. Homogenizing (considering f and g as degree 2 polynomials), we get

$$f^h(x, y) = xy - y^2, \qquad g^h(x, y) = xy + y^2.$$

From this, we see that f and g *did* share a root, it was just at infinity! More precisely, $(x : y) = (1 : 0) \in \mathbb{P}^1$ is a root of f^h and g^h.

Example 3.5. One form of the projective resultant should be familiar to anyone who has taken a linear algebra class. Consider the linear homogeneous polynomials $f(x, y) = ax + by$ and $g(x, y) = cx + dy$. Their resultant is given by

$$\operatorname{Res}(f, g) = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc,$$

which is the standard way we are taught to check if two linear equations share a non-trivial solution.

We notice that the case work we had to do with the resultant based on the actual degree of the polynomials does not need to be applied here. Hence, f and g share a root in $\mathbb{P}^n$ *if and only if* their projective resultant is zero.

As with the affine case, we examine the projection

$$\pi_{\text{pr}} \colon \mathbb{A}^{r+1} \times \mathbb{P}^2 \to \mathbb{A}^{r+1},$$
$$(a_0, \ldots, a_r; x : y) \mapsto (a_0, \ldots, a_r).$$

Proposition 3.6. π_{pr} is a **closed morphism** (it sends Zariski closed sets to Zariski closed sets).

This result follows from the fundamental theorem of elimination theory, which we will prove in the next section.

4 Elimination theory

The goal of this section will be to generalize the result in Proposition 3.6. We will need to see how we can create something similar to the resultant for homogeneous polynomials in more than 2 variables.

We first use the following lemma to show that checking for a common root in $\mathbb{P}^n$ can be reframed as a question about ideals:

Lemma 4.1. Homogeneous polynomials $f_1, \ldots, f_m \in S$ have no common roots in $\mathbb{P}^n$ if and only if $(f_1, \ldots, f_m) \supseteq (x_0, \ldots, x_n)^N$ for some (sufficiently large) N. Equivalently, $(f_1, \ldots, f_m) \supseteq S_N$ for some N.

Proof. Suppose $f_1, \ldots, f_m$ share a common root in $\mathbb{P}^n$. This is true if and only if their zero set in $\mathbb{A}^{n+1}$ is more than the origin. Notice that $\mathbf{V}(x_0, \ldots, x_n) = \{(0, \ldots, 0)\}$, so this is equivalent to

$$\mathbf{V}(f_1, \ldots, f_m) \nsubseteq \mathbf{V}(x_0, \ldots, x_n).$$

By the order-reversing property of ideals, this is equivalent to

$$I(\mathbf{V}(f_1, \ldots, f_m)) \nsupseteq I(\mathbf{V}(x_0, \ldots, x_n)).$$

By Hilbert's Nullstellensatz (which states that $I(\mathbf{V}(\mathfrak{a}))$ is $\{f \mid f^n \in \mathfrak{a}\} =: \sqrt{\mathfrak{a}}$, the **radical** of the ideal $\mathfrak{a}$), we can rewrite this as

$$\sqrt{(f_1, \ldots, f_m)} \nsupseteq \sqrt{(x_0, \ldots, x_n)} = (x_0, \ldots, x_n).$$

This is true if and only if no powers of some generator of $(x_0, \dots, x_n)$ is in $(f_1, \dots, f_m)$. That is, there exists some $0 \le i \le n$ such that for all $N > 0$, $x_i^N \notin (f_1, \dots, f_m)$. This is true if and only if

$$(x_0, \dots, x_n)^N \nsubseteq (f_1, \dots, f_m)$$

for all N. Indeed, if $x_i^N \notin (f_1, \dots, f_m)$, then $(x_0, \dots, x_n)^N \nsubseteq (f_1, \dots, f_m)$. Conversely, if, for all i, $x_i^{N_i} \in (f_1, \dots, f_m)$ for some N_i, then for sufficiently large N (the precise value is $N = (\max_i N_i - 1)(n+1) + 1$), $(x_0, \dots, x_n)^N \subseteq (f_1, \dots, f_m)$.

The last statement follows from noticing that $(x_0, \dots, x_n)^N = (S_N)$. □

Example 4.2. For these examples, let $k = \mathbb{C}$.

(a) We will show that the system

$$\begin{cases} f(x,y,z) = x^2 + y^2 - z^2 = 0 \\ g(x,y,z) = x - y = 0 \\ h(x,y,z) = y = 0 \end{cases}$$

has no non-trivial solutions in $\mathbb{P}^2$. By Lemma 4.1, we need to find some N such that

$$(f, g, h) \supseteq S_N.$$

For $N = 1$, the only way to form a degree 1 polynomial as a linear combination of f, g, and h is

$$0f + c_1 g + c_2 h = c_1 x + (c_2 - c_1)y,$$

where $c_1, c_2 \in \mathbb{C}$. There is no way to make the monomial z, so $(f, g, h) \not\supseteq S_1$. Next, we consider S_2. When S_2 is written as a $\mathbb{C}$-vector space, we have

$$S_2 = \operatorname{span}_{\mathbb{C}} \left\{x^2, y^2, z^2, xy, xz, yz\right\}.$$

Combining this with the fact that we can obtain every basis element from f, g, and h:

$$y \cdot h(x,y,z) = y^2, \quad x \cdot h(x,y,z) = xy, \quad z \cdot h(x,y,z) = yz,$$

$$x \cdot g(x,y,z) + x \cdot h(x,y,z) = x^2, \quad z \cdot g(x,y,z) + z \cdot h(x,y,z) = xz,$$

$$-f(x,y,z) + x \cdot g(x,y,z) + (x+y) \cdot h(x,y,z) = z^2,$$

we conclude that $(f, g, h) \supseteq (S_2)$.

(b) Consider the system

$$\begin{cases} f(x,y,z) = x^2 - yz = 0 \\ g(x,y,z) = x^2 + yz = 0 \\ h(x,y,z) = x = 0 \end{cases}$$

By Lemma 4.1, this has a solution if and only if

$$(f, g, h) \not\supseteq S_N$$

for all N.

I claim $z^N \notin (f, g, h)$ for any $N \ge 1$. Indeed, suppose there were homogeneous polynomials p, q, r with degrees $N-2$, $N-2$, and $N-1$ respectively such that

$$pf + qg + rh = z^N.$$

But the highest power of z in any monomial in pf, qg, and rh individually is z^{N-1}, so this is impossible. So $(f, g, h) \not\supseteq S_N$ for all N and this system has a solution. Indeed, $(0 : 0 : 1)$ is a solution, as we can see in Figure 3.

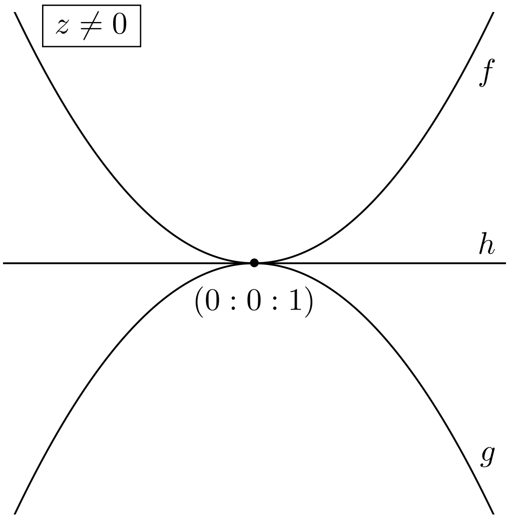

Figure 3: On the affine patch $\{z \neq 0\} \subseteq \mathbb{P}^2$ we can dehomogenize f, g, and h by setting $z = 1$ to get $f(x, y) = x^2 - y$, $g(x, y) = x^2 + y$, $h(x, y) = x$. Their zero sets are plotted here, where they visually have an intersection point.

Remark 4.3. Notice a key part in both of these proofs was related to proving the surjectivity of the map $(p, q, r) \mapsto pf + qg + rh$.

Both of these examples relied on linear algebra, and it turns out the main idea in the proof of the fundamental theorem of elimination theory is linear algebra.

Theorem 4.4 (The fundamental theorem of elimination theory)**.** The projection

$$\pi_{\text{pr}} \colon \mathbb{A}^{n+1} \times \mathbb{P}^m \to \mathbb{A}^{n+1},$$
$$(a_0, \ldots, a_n; x_0 : \cdots : x_m) \mapsto (a_0, \ldots, a_r),$$

is closed.

Proof. Let's generalize the ideas from the examples (this will follow the proof in [Vak]). Let $f_1, \ldots, f_m \in k[a_0, \ldots, a_n; x_0 : \cdots x_m]$ be polynomials that are homogeneous in $x_0, \ldots, x_m$. By Lemma 4.1, $f_1(\mathbf{x}) = f_2(\mathbf{x}) = \cdots = f_m(\mathbf{x}) = 0$ has a solution in $\mathbb{P}^n$ if and only if

$$(f_1, \ldots, f_m) \not\supseteq S_N$$

for all $N > 0$. This is equivalent to the existence of a polynomial $g \in S_N$ such that there *do not* exist homogeneous polynomials $p_1, \ldots, p_m$ satisfying

$$\sum_{i=1}^{m} p_i f_i = g.$$

Notice that each p_i must have degree $N - \deg f_i$, so this is equivalent to the linear map

$$S_{N - \deg f_1} \oplus \cdots \oplus S_{N - \deg f_m} \to S_N,$$
$$(p_1, \ldots, p_m) \mapsto \sum_{i=1}^{m} p_i f_i,$$

not being surjective, that is, the matrix representing this linear map not having full rank. This can be checked by seeing if all the determinants of the $\dim S_N \times \dim S_N$ sub-matrices inside this linear map vanish. By repeating this for all N, we get a (infinite) system of equations in $k[a_0, \ldots, a_r]$ that precisely tell us where $f_1, \ldots, f_m$ share a common root.

Let the determinants of the sub-matrices form the set $\{g_1, g_2, \ldots\} \subseteq k[a_0, \ldots, a_n]$. Despite the fact that this set is infinite, it still defines a closed set in $\mathbb{A}^{n+1}$, because $k[a_0, \ldots, a_n]$ is Noetherian, so $(g_1, g_2, \ldots)$ is finitely generated (in fact, this means that there is some bound on the N we need to check, but it may depend on the polynomials $f_1, \ldots, f_m$). □

Acknowledgements

I would like to thank Brian Lawrence for illustrating how to derive some of these proofs, giving insightful examples, and directing my studies in algebraic geometry.

References

[Ḿ] Pierre-Loïc Méliot. The Resultant of Two Polynomials.

[Vak] Ravi Vakil. MATH 216: Foundations of Algebraic Geometry.