Computational Algebraic Geometry

Laurent Busé

Computational Algebraic Geometry

Draft notes of Master2 MPA, 2020-22.

Laurent Busé
Université Côte d'Azur, Inria, France.
laurent.buse@inria.fr

November 30, 2021

Introduction

These notes are based on a course taught at the master of mathematics, 2nd year, of the University Côte d'Azur during the first semesters of the academic years 2020-21 and 2021-22. It covers some basic topics in computational algebraic geometry, with the goal to illustrate the use of computer algebra systems, such as `Macaulay2` [GS].

The content of these notes are based on the existing classical books on commutative algebra [Eis95] and algebraic geometry [EH00, Har77, Har92], including those with a more computational flavor [CLO07, CLO98, MS20], and in particular the book [Sch03] by Hal Schenck. The first chapter deals with algebraic varieties and primary decompositions, which provide effective methods for decomposing algebraic varieties. Computational definitions of dimension and degree of projective algebraic varieties are given in Chapter 2 by means of Hilbert polynomials. Then, finite free resolutions and regular sequences are introduced in Chapter 3 in order to demonstrate that dimension and degree, as defined by means of Hilbert polynomials, have the expected properties when slicing a projective algebraic variety by a hypersurface. In Chapter 4, Gröbner bases are presented; this is the main tool to perform computations in algebraic geometry. Some more specific tools and applications are treated in Chapter 5: the computation of projections, the Sylvester resultant, Bézout Theorem in the projective plane and the implicitization of parameterized algebraic plane curves by means of syzygies. The notes end with resultants over a projective space, a refined tool to perform elimination under suitable assumptions. For this part, we present the computational approach developed by Jean-Pierre Jouanolou in [Jou91, Jou97].

Contents

Chapter 1

Ideals and Varieties

1.1 Background on Rings and Modules

Ring. It is an abelian group $(+)$ with an associative multiplication $(.)$ which is distributive with respect to $+$. In these lectures, all rings with be commutative with unit (1).

Field. It is a ring such that every nonzero element has a multiplicative inverse.

Zero divisor. An element $a \neq 0$ is a nonzero divisor if there exists $b \neq 0$ such that $ab = 0$. A ring without zero divisors is a domain.

Modules. They are to rings what vector spaces are to fields. Let R be a ring. M is an R-module if M is an abelian group and there is a R-linear map $R \times M \to M$ such that

$$r_1(m_1 + m_2) = r_1m_1 + r_1m_2,\ (r_1 + r_2)m_1 = r_1m_1 + r_2m_1,$$

$$r_1(r_2m_1) = (r_1r_2)m_1,\ 1.m = m.$$

An R-module M is *finitely generated* if there exists $\{m_1, \dots, m_n\} \subset M$ such that for all $m \in M$, $m = \sum_{i=1}^{n} r_im_i$ with $r_i \in R$.

Examples of modules.

- A ring is a module over itself
- Ideals $I \subset R$ are submodules of the ring itself.
- The quotient ring R/I is an R-module, and also an R/I-module.
- An R-module M is *free* if $M \simeq \oplus_{i=1}^{n} R$. Not all modules are free.

Maps.

- A morphism of rings is a map $\phi : A \to B$ between two rings such that

$$\phi(ab) = \phi(a)\phi(b),\ \phi(a + b) = \phi(a) + \phi(b),\ \phi(1) = 1.$$

- An R-module morphism is a map $\psi : M_1 \to M_2$ of R-modules such that

$$\psi(m_1 + m_2) = \psi(m_1) + \psi(m_2), \ \psi(rm) = r\psi(m).$$

Notice that kernels, cokernels and images of such maps are all R-modules.

Some important types of ideals. Recall that an ideal $I \subset R$ is said to be a proper ideal if $I \neq (1)$.

- I is *principal* if I can be generated by a single element.
- A proper ideal I is *prime* if $fg \in I \Rightarrow f \in I$ or $g \in I$.
- A proper ideal I is *maximal* if there is no proper ideal J such that $I \subsetneq J$.
- A proper ideal I is *primary* if $fg \in I \Rightarrow f \in I$ or $g^m \in I$ for some m.
- A proper ideal I is *irreducible* if there do not exist ideals J_1 and J_2 such that $I = J_1 \cap J_2$ with $I \subsetneq J_i$.
- I is *radical* if the following property holds: $f^m \in I$ for some $m \in \mathbb{N} \Rightarrow f \in I$.

Exercise 1.1.1. A local ring is a ring with a unique maximal ideal $\mathfrak{m}$. Prove that in a local ring, if $f \notin \mathfrak{m}$ then f is a unit.

Exercise 1.1.2. Prove that we have the following implications for an ideal I in R:

$$I \text{ maximal } \Rightarrow I \text{ prime } \begin{array}{l} \Rightarrow I \text{ is radical} \\ \Rightarrow I \text{ is primary.} \end{array}$$

These implications are not reversible. Nevertheless, an ideal which is primary and radical is prime. Also, every intersection of prime ideals is radical. (See for instance [MS20, Proposition 1.6] for proofs).

1.2 Ideals and Varieties

Let k be a field and define the polynomial ring $R = k[x_1, \ldots, x_n]$.

Affine varieties. The *affine n-space* over k is defined as

$$\mathbb{A}^n_k := \{(a_1, \ldots, a_n) \in k^n\}.$$

An *affine variety* is the common zero locus of a collection of finitely many polynomial $f_1, \ldots, f_m \in R$. It is denoted by

$$V_k(f_1, \ldots, f_m) := \{\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{A}^n_k : f_1(\mathbf{a}) = \cdots = f_m(\mathbf{a}) = 0\} \subset \mathbb{A}^n_k$$

(whenever there is no ambiguity, we will omit the index k). It is important to notice that $V(f_1, \ldots, f_m)$ only depends of the ideal $I = (f_1, \ldots, f_m)$ in R; thus it is also denoted by $V(I)$.

Example 1.2.1. Let $I = (x^2 - y^2 - 3, 2x^2 + 3y^2 - 11) \subset \mathbb{R}[x, y]$; what is $V(I) \subset \mathbb{A}^2_{\mathbb{R}}$?

A first option is to set $u = x^2$ and $v = y^2$ and solve for u and v. A second option is to use the use the first equation to write $x^2 = y^2 + 3$ and to substitute it in the second equation, so that

$$\begin{aligned} V(I) &= V(x^2 - y^2 - 3, 2(y^2+3) + 3y^2 - 11) \\ &= V(x^2 - y^2 - 3, 5y^2 - 5) = V(x^2 - y^2 - 3, y^2 - 1). \end{aligned}$$

Now, in the same way we use the second equation to simplify the first one and we get $V(I) = V(x^2 - 2, y^2 - 1)$. In the next lectures we will see a systematic way to find such "good" generators for the ideal I.

Ideal of a set. Given an ideal J in R we have just defined the geometric object $V(J)$. Conversely, given $S \subseteq \mathbb{A}^n$, we define $I(S)$ as the set of all polynomials in R that vanishes on S. This is an ideal of R (prove it!).

Lemma 1.2.2. *The following implications hold:*

$$J_1 \subseteq J_2 \Rightarrow V(J_2) \subseteq V(J_1), \quad S_1 \subseteq S_2 \Rightarrow I(S_2) \subseteq I(S_1).$$

Remark 1.2.3. If $S = V(J)$ for some ideal J, then we do not always have $J = I(V(J))$ (take for instance $J = (x^2) \subset k[x]$, in which case $I(V(J)) = (x)$). However, one always has $J \subseteq I(V(J))$.

Remark 1.2.4. Show that $X = V(I(X))$ if X is a variety, that is to say if $X = V(J)$ for some ideal J ($X \subseteq V(I(X))$ obviously; then, $X = V(J)$ so $J \subseteq I(X)$ and hence $V(I(X)) \subseteq V(J) = X$).

Irreducible varieties. A natural thing to do to study varieties is to break them into simpler parts. In this direction, we introduce irreducible varieties.

Definition 1.2.5. *A non-empty variety V is irreducible if it is not the union of two proper subvarieties:*

$$V \neq V_1 \cup V_2 \ \textit{ for any } V_i \ : \ V_i \subsetneq V.$$

Theorem 1.2.6. *V is irreducible if and only if $I(V)$ is a prime ideal.*

Proof. ($\Leftarrow$): Assume that $I(V)$ is prime and that $V = V_1 \cup V_2$ is reducible. Set $I_1 = I(V_1)$ and $I_2 = I(V_2)$. We claim that there exist $p \in V_2$ and $f \in I_1$ such that $f(p) \neq 0$. Indeed, if this is not the case then $I_1 \subseteq I_2$ and hence $V_2 = V(I_2) \subseteq V_1 = V(I_1)$, which is not possible. By the same argument, there exist $q \in V_1$ and $g \in I_2$ such that $g(p) \neq 0$. Now, $fg \in I(V_1 \cup V_2) = I(V)$ but $f \notin I(V)$ and $g \notin I(V)$, which is in contradiction with the fact that $I(V)$ is a prime ideal.

($\Rightarrow$): For the converse, assume that $I(V)$ is not prime, i.e. there exist $f, g \notin I(V)$ such that $fg \in I(V)$. Define $V_1 = V(I(V) + (f))$ and $V_2 = V(I(V) + (g))$. We have $V_1 \subsetneq V(I(V))$ because $f \notin I(V)$ and similarly, $V_2 \subsetneq V(I(V))$ because $g \notin I(V)$. It follows that $V_1 \cup V_2 \subseteq V$. Moreover, $V \subseteq V_1 \cup V_2$ because $fg \in I(V)$ (for any $p \in V$, $f(p)g(p) = 0$ so f or g, or both, vanish at p). Therefore, we deduce that $V = V_1 \cup V_2$ and V is reducible. □

Exercise 1.2.7. Let I, J be two ideals, then $I+J$, IJ and $I \cap J$ are ideals. Show that $V(I+J) = V(I) \cap V(J)$ and $V(IJ) = V(I \cap J) = V(I) \cup V(J)$.

1.3 Hilbert Basis Theorem

Definition 1.3.1. *A ring is Noetherian if it contains no infinite ascending (infinite proper inclusions) chains of ideals $I_1 \subsetneq I_2 \subsetneq \cdots$.*

Lemma 1.3.2. *A ring is Noetherian if and only if every ideal is finitely generated.*

Proof. Exercise! □

Theorem 1.3.3 (Hilbert Basis Theorem)**.** *If a ring R is Noetherian then the polynomial ring $R[x]$ is also Noetherian.*

Proof. First we introduce the following terminology: if $f = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x_1 + a_0 \in R[x]$, with $a_n \neq 0$, we define the *initial term* of f to be $a_n x^n$ and the *initial coefficient* of f to be a_n.

Let I be an ideal in $R[x]$. We aim to show that I is finitely generated. For that purpose, choose a sequence of elements $f_1, f_2, \ldots$ in I as follows: f_1 is a nonzero element of least degree in I. For all $i \geq 1$, if $(f_1, \ldots, f_i) \subsetneq I$ then choose f_{i+1} to be an element of least degree in $I \setminus (f_1, \ldots, f_i)$. If $(f_1, \ldots, f_i) = I$ then we stop choosing elements.

For all j, denote by a_j the leading coefficient of f_j. Since R is Noetherian, the ideal $J = (a_1, a_2, \ldots)$ is finitely generated. Let m be the smallest integer such that $J = (a_1, \ldots, a_m)$. We claim that $I = (f_1, \ldots, f_m)$. Indeed, if this is not true then the above process select a nonzero element f_{m+1} and there exist element $u_j \in R$ such that $a_{m+1} = \sum_{i=1}^{m} u_j a_j$. As the degree of f_{m+1} is greater or equal to the degree of $f_1, \ldots, f_m$ by construction, we can define the polynomial

$$g = \sum_{i=1}^{m} u_j x^{\deg(f_{m+1}) - \deg(f_j)} f_j \in (f_1, \ldots, f_m)$$

that has the same degree and the same initial term as f_{m+1}. The polynomial $f_{m+1} - g$ belongs to I and not to $(f_1, \ldots, f_m)$, but its degree is strictly less than the degree of f_{m+1}, which gives a contradiction. □

Straightforward applications of this result shows that if I is an ideal in a Noetherian ring R, then R/I is also Noetherian. Also, using an induction we deduce that any polynomial ring over a Noetherian ring is Noetherian. As a consequence of all this we also get that any finitely algebra R over a Noetherian ring R_0, is Noetherian (see [Eis95, §I.1.4]).

The above definition and properties can be extended to modules; see [Eis95, §I.1.4].

We notice that if k is a field then $k[x_1, \ldots, x_n]$ is a Noetherian ring and hence all its ideals are finitely generated. In the next lectures we will develop techniques to find nice generating sets for such ideals.

1.4 Primary Decomposition

In this part, we aim to decompose an ideal into the intersection of simpler ones (remember that geometrically, the intersection of ideals corresponds to the union of varieties).

Exercise 1.4.1.

- Prove that the ideal $(x^2 - 4, y^2 - 1)$ can be written as the intersection of four maximal ideals in $\mathbb{R}[x, y]$.
- Prove that $(x^2 - x, xy) = (x) \cap (x - 1, y)$, which is the intersection of a prime ideal and a maximal ideal.
- Is the ideal (x^2) in $k[x]$, k a field, can be written as the intersection of prime ideals?

We are now going to prove that in a Noetherian ring, any ideal can be written as a finite intersection of primary ideals. We recall the following definitions:

- A proper ideal I is *primary* if $fg \in I \Rightarrow f \in I$ or $g^m \in I$ for some m.
- A proper ideal I is *irreducible* if there do not exist ideals J_1 and J_2 such that $I = J_1 \cap J_2$ with $I \subsetneq J_i$.
- I is *radical* if the following property holds: $f^m \in I$ for some $m \in \mathbb{N} \Rightarrow f \in I$.

Proposition 1.4.2. *Let I be an ideal in a Noetherian ring R, then there exist finitely many primary ideals $q_1, \dots, q_s$ in R such that*

$$I = q_1 \cap q_2 \cap \cdots \cap q_s.$$

Proof. We first prove that in a Noetherian ring, any ideal is a finite intersection of irreducible ideals. Suppose not and let I_1 be an ideal which is not a finite intersection of irreducible ideals. I_1 is hence reducible: $I_1 = J_1 \cap J_2$ with ideals J_1 and J_2 both strictly larger than I_1. If J_1 and J_2 are finite intersections of irreducible ideals, then so is I_1, so assume $I_2 := J_1$ is not. We have $I_1 \subsetneq I_2$. And repeating the above construction we get an ascending chain of ideals $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots$; this is in contradiction with our Noetherian assumption, so I_1 must be a finite intersection of irreducible ideals.

We next prove that, in a Noetherian ring, any irreducible ideal is a primary ideal. So, let I be an irreducible ideal. Let $f, g \in R$ such that $fg \in I$ but $f \notin I$; we shall show that, for some positive integer m, $g^m \in I$, equivalently $g^m = 0$ in the quotient ring $A = R/I$.

Recall that the *annihilator* of an element $r \in A$ is the ideal $\mathrm{ann}_A(r) = \{a \in A : ar = 0\}$ of R. Since A is Noetherian (because R is Noetherian), the ascending chain of ideals

$$0 \subseteq \mathrm{ann}_A(g) \subseteq \mathrm{ann}_A(g^2) \subseteq \mathrm{ann}_A(g^3) \subseteq \cdots$$

becomes stationary, which means that there exists a positive integer n such that $\mathrm{ann}_A(g^n) = \mathrm{ann}_A(g^{n+1})$. We claim that $(0) = (f) \cap (g^n)$. Indeed, let $a \in (f) \cap (g^n)$. Since $a \in (f)$, $ag = 0$ (because we assumed $fg \in I$). Moreover, $a \in (g^n)$, so $a = bg^n$

for some $b \in A$ and hence $(bg^n)g = bg^{n+1} = 0$. But $\operatorname{ann}_A(g^n) = \operatorname{ann}_A(g^{n+1})$, so it follows that $bg^n = 0$, hence that $a = 0$. To conclude the proof, we observe that, by our assumptions, the ideal (0) is an irreducible ideal in A and $f \neq 0$ in A. So, from the equality $(0) = (f) \cap (g^n)$ we deduce that $g^n = 0$ in A. □

We have just proved the existence of primary decompositions of ideals in Noetherian rings. The first natural question is now to ask about the uniqueness of such decompositions. The following result suggest grouping the primary ideals by their radicals.

Definition 1.4.3. *The* radical ideal *of an ideal I in a ring R, denoted $\sqrt{I}$, is the set of all $f \in R$ such that $f^m \in I$ for some integer m. The ideal I is said to be radical if $I = \sqrt{I}$.*

Lemma 1.4.4. *If q is a primary ideal then $p := \sqrt{q}$ is a prime ideal. Moreover p is the unique smallest prime ideal containing q.*

Proof. Let $f, g \in R$ such that $fg \in \sqrt{q}$. This implies that $f^m g^m \in q$ for some m. Since q is primary, either $f^m \in q$, i.e. $f \in \sqrt{q}$, or either $g^{mm'} \in q$ for some m', i.e. $g \in \sqrt{q}$. So $\sqrt{q}$ is a prime ideal, as claimed. To conclude the proof, let p' a prime ideal such that $q \subseteq p'$. By taking radicals, we get $p = \sqrt{q} \subseteq \sqrt{p'} = p'$ (recall that a prime ideal is radical). □

A primary ideal q whose radical is the prime ideal p is called *p-primary.*

Remark 1.4.5. Not all ideal whose radical is prime is primary. Consider $I = (x^2, xy) \in \mathcal{C}[x, y]$, then $\sqrt{I} = (x)$ but I is not a primary ideal ($xy \in I$ but $x \notin I$ and $y^m \notin I$ for any integer m).

Example 1.4.6. We notice that it is not true in general that the power of a prime ideal is a primary ideal, although this may hold in many cases; see [MS20, Exercise 3.13] for an example (computations with Macaulay2 recommended).

Now, we focus on p-primary ideals for a fixed prime ideal p.

Lemma 1.4.7. *If q_1 and q_2 are p-primary ideals, then so is $q_1 \cap q_2$.*

Proof. Set $q = q_1 \cap q_2$. It is easy to check that $\sqrt{q} = \sqrt{q_1} \cap \sqrt{q_2}$ (this property actually holds for any couple of ideals), so that $\sqrt{q} = p$. To prove that q is primary, assume that $fg \in q$ and $f \notin q$. Then $f \notin q_1$ or $f \notin q_2$, say $f \notin q_1$. But $fg \in q_1$ and q_1 is primary, so $g \in \sqrt{q_1} = p = \sqrt{q}$. It follows that $g^m \in q$ for some integer m, which concludes the proof. □

The above lemmas suggest to gather primary ideals having the same radical in a primary decomposition. This leads to the following definitions.

Definition 1.4.8. *A* minimal primary decomposition *of an ideal I in a Noetherian ring R is a decomposition*

$$I = q_1 \cap q_2 \cap \cdots \cap q_s$$

where the q_i's are primary ideals that have pairwise distinct radicals and such that the intersection is irredundant, meaning $\cap_{i \neq j} q_i \neq I$ (equivalently $\cap_{i \neq j} q_i$ is not contained in q_j) for all j.

The prime ideals $p_i = \sqrt{q_i}$, with $i = 1, \ldots, s$, are called the associated primes *of the ideal I; they are all distincts. The set of associated primes of the ideal I is usually denoted by* $\mathrm{Ass}(I)$.

The minimal (with respect to inclusion) elements of Ass(I) are the minimal primes *of the ideal I. Associated primes that are not minimal are called* embedded primes.

Some comments and examples are in order:

- Minimal primary decompositions may not be unique; here is an example:

$$(x^2, xy) = (x) \cap (x, y)^2 = (x) \cap (x^2, y) \subset \mathcal{C}[x, y].$$

- In the above primary decomposition, (x) is a minimal associated prime; geometrically it corresponds to the line of equation $x = 0$. The associated prime (x, y) is an embedded prime; geometrically it corresponds to the "embedded" point $x = y = 0$ which is supported on the line $x = 0$.
- The ideal $(x^2 - x, xy) = (x) \cap (x-1, y)$ has two minimal primes and no embedded primes (exercise: prove that this is a minimal primary decomposition).
- The radical of I is the intersection of the minimal primes of I (exercise: prove it). For instance, $\sqrt{(x^2, xy)} = (x)$. In other words, the radical "remove" embedded primes.

Our next goal is to show that while minimal primary decompositions are not unique, associated primes are unique (i.e. independent of the choice of the minimal primary decomposition).

Definition 1.4.9. *Let I, J be two ideals in a ring R. The ideal quotient of I by J is the ideal*

$$(I : J) = \{r \in R \ : \ rJ \subseteq I\}.$$

Lemma 1.4.10. *Let q be a p-primary ideal and let $f \in R$.*

i) If $f \in q$ then $(q : f) = R$.

ii) If $f \notin q$ then $(q : f)$ is p-primary.

iii) If $f \notin p$ then $(q : f) = q$.

Proof. i) is clear. *iii)*: The inclusion $q \subseteq (q : f)$ is obvious. Let $g \in (q : f)$. Then $gf \in q$ but $f^m \notin q$ for some m (otherwise $f \in p = \sqrt{q}$), so $g \in q$.

Next, we prove *ii)*. If $g \in (q : f)$ then $g^m \in q$ for some m because $f \notin q$ and q is a primary ideal. This implies that $g \in \sqrt{q} = p$. Therefore, we have the inclusions

$$q \subseteq (q : f) \subseteq p$$

from we deduce, by taking radicals, that $\sqrt{(q:f)} = p$. To show that $(q : f)$ is primary, let $ab \in (q : f)$. If $a \in (q : f)$ then we are done. Otherwise, $a \notin (q : f)$, i.e. $af \notin q$, but $ab \in (q : f)$ implies that $(fa)b \in q$, which implies (q is primary) that $b^m \in q \subseteq (q : f)$ for some m. □

Exercise 1.4.11. Prove the following properties.

- If a prime ideal $p = I_1 \cap I_2$ then $p = I_1$ or $p = I_2$.
- $(I_1 \cap I_2) : f = (I_1 : f) \cap (I_2 : f)$.
- $\sqrt{I_1 \cap I_2} = \sqrt{I_1} \cap \sqrt{I_2}$.

Corollary 1.4.12. *The associated primes of an ideal I are independent of the choice of a (minimal) primary decomposition.*

Proof. Let $I = q_1 \cap \cdots \cap q_s$ be a minimal primary decomposition. Since the decomposition is irredundant, for all i we have $\cap_{j \neq i} q_j \not\subset q_i$. Therefore, for all i one can choose $f_i \notin q_i$ such that $f_i \in q_j$ for all $j \neq i$. By Lemma 1.4.10, we deduce that $(I : f_i) = (q_i : f_i)$ is p_i-primary, where $p_i := \sqrt{q_i} = \sqrt{(q_i : f_i)}$. It follows that the primes p_i's belong to the set

$$\mathcal{S} := \{\sqrt{(I:f)} \text{ such that } f \in R\}.$$

Now, let p be a prime ideal of the form $\sqrt{(I : f)}$ for some $f \in R$. We have (use Exercise 1.4.11)

$$p = \sqrt{(I:f)} = \sqrt{\cap_j (q_j : f)} = \cap \sqrt{(q_j : f)} = \cap_j p_j$$

and hence $p = p_j$ for some j. In conclusion, the associated primes of the ideal I are the prime ideals in S. □

1.5 The Nullstellensatz and Zariski topology

Given two varieties X, Y, it is natural to consider maps between them, so we need a topology. In what follows k denotes a field (that will be very often assumed to be algebraically closed).

Definition 1.5.1 (Topology). *A topology on a set X is a collection $\mathcal{U}$ of subsets of X such that*

- *$\emptyset$ and X are in $\mathcal{U}$,*
- *$\mathcal{U}$ is closed under finite intersections,*
- *$\mathcal{U}$ is closed under arbitrary unions.*

The members of $\mathcal{U}$ are called the open sets. Closed sets *are the complements of open sets.*

Zariski topology. The closed sets in k^n are affine varieties $V(I)$, $I \subset k[x_1, \dots, x_n]$ (closed under finite union). The notation $\mathbb{A}^n_k$ is used instead of k^n in order to emphasis that we use this topology (called the Zariski topology).

If X is a variety in $\mathbb{A}^n_k$, then X is naturally equipped with the subspace topology: the open sets in X are the open sets in k^n intersected with X.

An open set of the form $\mathcal{C}V(f)$, $f \in k[x_1, \dots, x_n]$ is called a *distinguished open set* and often denoted bu U_f. Notice that every Zariski open set can be written as a union of distinguished open sets.

The Nullstellensatz. We have seen that for any ideal J, $J \subseteq I(V(J))$ and that this can be a proper containment. Notice also that from the definition of the radical, we have

$$J \subseteq \sqrt{J} \subseteq I(V(J)).$$

To get a precise relation between J and $I(V(J))$, a first question to consider is the following: when is the variety of an ideal empty? It is clear that if $1 \in J$ then $V(J) = \emptyset$. On the other hand, $(x^2+1) \subset \mathbb{R}[x]$ is a proper ideal and $V_{\mathbb{R}}(x^2+1) = \emptyset$.

Theorem 1.5.2 (Weak Hilbert Nullstellensatz)**.** *If k is an algebraically closed field and $V(J)$ is empty, then $1 \in J$.*

Proof. We refer to [Eis95, Chapter 4, §4.5], but also to [MS20, Theorem 6.1] (this proof uses Gröbner basis, that we will see later on). See also [Har77, Theorem 1.3A] for some other references. □

Theorem 1.5.3 (Strong Hilbert Nullstellensatz)**.** *Assume k is an algebraically closed field and let J be an ideal in $R = k[x_1, \dots, x_n]$. If $f \in I(V(J)) \subset R$, then $f^m \in J$ for some integer m, i.e. $\sqrt{J} = I(V(J))$.*

Proof. Let $J = (f_1, \dots, f_s)$ and $f \in I(V(J))$. Consider the ideal $J' = (J, 1 - yf)$ in the ring $R[y]$. Since $V(J') = \emptyset$ we deduce from the weak Nullstellensatz that

$$1 = \sum_{i=1}^{s} a_i f_i + g(1 - yf) \in R[y].$$

Substituting formally y by $1/f$ in this equality and cleaning denominators, we get the claimed property. □

As a consequence, over an algebraically closed field varieties are in correspondence with radical ideals.

Algebraic closure. Let $S \subset k^n$ be a set. Then $V(I(S))$ is the smallest variety containing S; it is called the Zariski closure and denoted by $\bar{S} = V(I(S))$. Here is an interesting consequence of Hilbert Nullstellensatz.

Proposition 1.5.4. *Let I, J be two ideals in $R = k[x_1, \dots, x_n]$, then*

$$\overline{V(I) \setminus V(J)} \subseteq V(I : J).$$

Moreover, if k is algebraically closed and I is radical, then this is an equality.

Proof. To prove the first claim, we need to show that

$$(I : J) \subseteq I(V(I) \setminus V(J)).$$

Pick an element $f \in (I : J)$ and a point $p \in V(I) \setminus V(J)$. There exists a polynomial $g \in J$ such that $g(p) \neq 0$. But $fg \in I$ so $f(p)g(p) = 0$ and hence $f(p) = 0$.

Now, assume that k is an algebraically closed field and that $I = \sqrt{I}$. Pick a point $p \in V(I : J)$ and let $f \in I(V(I) \setminus V(J))$; we want to show that f vanishes at p. For any $g \in J$, fg vanishes on $V(I)$, which implies that $fg \in \sqrt{I}$ by the Nullstellensatz, and that $fg \in I$ by our assumption. It follows that $f \in (I : J)$ and hence that $f(p) = 0$. □

Example 1.5.5. Consider the set

$$S = \{p_1, p_2, p_3, p_4\} = \{(0,0), (0,1), (1,0), (1,1)\} \subseteq \mathbb{A}^2_k.$$

We have

$$I(S) = \cap_{i=1}^4 I(p_i) = (x^2 - x, y^2 - y).$$

One can remove the points lying on the line $x = y$ as follows:

$$(x^2 - x, y^2 - y) : (x - y) = (x + y - 1, y^2 - y).$$

Chapter 2

Graded Objects and Projective Geometry

Algebraically closed fields can be used to ensure that a univariate polynomial $f(x)$ has always as many roots as its degree, counted with multiplicity. Analogously, projective spaces are the right place to count intersections between algebraic objects. For instance, two lines in $\mathbb{A}^2$ can be disjoint (distinct and parallel), whereas in the projective plane $\mathbb{P}^2$ they will always intersect.

2.1 Projective Spaces and Varieties

Let k be an infinite field. The projective space of dimension n is the space of lines passing through the origin in the affine space of dimension $n+1$. More precisely:

$$\mathbb{P}_k^n := \left(\mathbb{A}_k^{n+1} \setminus \{0\}\right) / (a_0, \dots, a_n) \sim (b_0, \dots, b_n)$$

where $(a_0, \dots, a_n) \sim (b_0, \dots, b_n)$ if and only if there exists $\lambda \in k^* := k \setminus \{0\}$ such that $(a_0, \dots, a_n) = \lambda(b_0, \dots, b_n)$ (points on the same line through the origin are identified).

Coordinates in $\mathbb{P}_k^n$, which are usually called *homogeneous coordinates*, are denoted by $(a_0 : \dots : a_n)$, which are hence defined up to multiplication by a nonzero constant:

$$(a_0 : \dots : a_n) = (\lambda a_0 : \dots : \lambda a_n) \text{ for all } \lambda \in k^*.$$

We notice that by definition, $(0 : \dots : 0)$ is not valid.

Depending on notations, the hyperplane corresponding to points such that $a_0 = 0$ is called the *hyperplane at infinity*. It is a $\mathbb{P}_k^{n-1}$. Its complement if the affine part of $\mathbb{P}_k^n$: its points are such that $a_0 \neq 0$ and hence

$$(a_0 : a_1 : \dots : a_n) = \left(1 : \frac{a_1}{a_0} : \dots : \frac{a_n}{a_0}\right)$$

and they form a $\mathbb{A}_k^n$. In conclusion $\mathbb{P}_k^n = \mathbb{A}_k^n \cup \mathbb{P}_k^{n-1}$.

To define a variety over a projective space, we need polynomials that vanish on lines through the origin, i.e.

$$p \in \mathbb{A}_k^{n+1} \;:\; f(p) = 0 \Rightarrow f(\lambda p) = 0 \;\forall \lambda \in k^*.$$

A polynomial is *homogeneous* if all its monomials are of the same total degree. Thus, a homogeneous polynomial defines a variety in $\mathbb{P}_k^n$ (but also in $\mathbb{A}_k^{n+1}$). The following celebrated result is a good motivation to consider projective varieties.

Theorem 2.1.1 (Bézout Theorem)**.** *Let k be an algebraically closed field and let f, g be two homogeneous polynomials in $k[x, y, z]$ of degree d, e, respectively, with no common factor. Then, the projective curves $V(f)$ and $V(g)$ in $\mathbb{P}_k^2$ meet in de points, counted with multiplicities.*

Proof. See [Har77, Corollary 7.8]. We will also see a proof of this theorem by means of the Sylvester resultant later on. □

Remark 2.1.2. As a consequence of Bézout Theorem, we notice that two lines in $\mathbb{A}_k^2$ may not intersect, but they always intersect in $\mathbb{P}_k^2$.

Definition 2.1.3.

- *A* homogeneous ideal *is an ideal generated by homogeneous elements.*
- *A variety in a projective space, called a* projective variety*, is a variety which is defined by a homogeneous ideal.*
- *The* Zariski topology *on $\mathbb{P}_k^n$ is defined by making projective varieties the closed sets.*

2.2 Graded Rings and Modules, Hilbert Functions

The algebraic counterpart of projective varieties are graded rings and modules.

A $\mathbb{Z}$-graded ring R is a ring which can be decomposed into homogeneous pieces:

$$R = \oplus_{i\in\mathbb{Z}} R_i \quad \text{(abelian group)}$$

such that if $r_i \in R_i$ and $r_j \in R_j$ then $r_i r_j \in R_{i+j}$.

A R-module M is graded if $M = \oplus_{i\in\mathbb{Z}} M_i$ and if $R_i M_j \subseteq M_{i+j}$ for all i, j. Elements in R_i and M_i are called *homogeneous elements of degree i*.

Example 2.2.1.

- $R = k[x_1, \ldots, x_n]$ is graded by setting $\deg(x_i) = 1$ (usual grading). Thus, $R_0 = k$, $R_1 = \langle x_1, \ldots, x_n\rangle_1$ (linear forms), etc.
- The quotient of a graded ring R by a homogeneous ideal I is a graded ring:

$$R/I = \oplus_{j\in\mathbb{Z}} R_j/I_j.$$

From now on we will assume that k is an infinite field and we consider the graded polynomial rings of the form $R = k[x_1, \ldots, x_n]$. The graded pieces R_i of R are k-vector spaces and one can consider their dimension as such.

Example 2.2.2.

- $\dim_k(k[x]_i) = 1$ for all i.
- $\dim_k(k[x,y]_i) = i+1$ for all i.
- $\dim_k(k[x_0,\dots,x_n]_i) = \binom{n+i}{i}$ for all i.

We observe that all the above formulas are polynomials in the variable i.

Definition 2.2.3 (Hilbert function)**.** *The Hilbert function of a finitely generated graded R-module M is defined by*

$$\mathrm{HF}_M(i) := \dim_k(M_i).$$

Remark 2.2.4. In the above definition, M_i is a $R_0 = k$-module so the finiteness follows from our assumption that M is finitely generated.

Example 2.2.5. If $R = k[x,y,z]$ then

$$\mathrm{HF}_R(i) = \binom{i+2}{2} = \frac{(i+2)(i+1)}{2}.$$

This number is the dimension if the vector space of forms of degree i in $\mathbb{P}^2$. For instance, conics in $\mathbb{P}^2$ form a $\mathbb{P}^5$.

Notation 2.2.6 (Shift in grading)**.** Let R be a graded ring and $m \in \mathbb{Z}$, the notation $R(m)$ is used to shift the grading by m; more precisely, $R(m)$ is the graded ring such that $R(m)_i = R_{i+m}$ for all $i \in \mathbb{Z}$.

Example 2.2.7. Let $R = k[x,y]$, then

i	0	1	2	3	4	$\cdots$
$\mathrm{HF}_R(i)$	1	2	3	4	5	$\cdots$
$\mathrm{HF}_{R(-2)}(i)$	0	0	1	2	3	$\cdots$

Example 2.2.8. Let $R = k[x,y,z]$ and consider the homogeneous ideal $I = (x^3 + y^3 + z^3)$. We have

$$\dim(R/I)_i = \dim(R_i/I_i) = \dim R_i - \dim I_i.$$

Moreover, since I is generated by a single homogeneous polynomial of degree 3, we deduce that

$$\dim(I_i) = \dim(R_{i-3}) = \dim(R(-3)_i).$$

Thus,

i	0	1	2	3	4	$\cdots$
$\mathrm{HF}_R(i)$	1	3	6	10	15	$\cdots$
$\mathrm{HF}_{R(-3)}(i)$	0	0	0	1	3	$\cdots$
$\mathrm{HF}_{R/I}(i)$	1	3	6	9	12	$\cdots$

Actually, for all $i \geq 1$,

$$\mathrm{HF}_{R/I}(i) = \dim R_i - \dim R_{i-3} = \binom{i+2}{2} - \binom{i-1}{2} = 3i.$$

(we observe that this is a polynomial in i).

Now, add a linear form, say x, to the ideal I; consider the ideal $J = I + (x)$ (geometrically, this corresponds to the intersection of the plane curve of equation $x^3 + y^3 + z^3 = 0$ with the line of equation $x = 0$). We have $R/J \simeq k[y, z]/(y^3 + z^3)$ and we get

i	0	1	2	3	4	$\cdots$
$\mathrm{HF}_{R/J}(i)$	1	2	3	3	3	$\cdots$

Thus, we observe that the Hilbert function of R/J stabilizes to the constant value 3. Moreover, by Bézout theorem, 3 is precisely the number of intersection points between the plane curve of equation $x^3 + y^3 + z^3 = 0$ with the line of equation $x = 0$.

Hilbert functions can be encoded as series.

Definition 2.2.9 (Hilbert Series)**.** *The Hilbert series of a finitely generated and graded R-module M is defined as*

$$\mathrm{HS}_M(t) = \sum_{i \in \mathbb{Z}} \mathrm{HF}_M(i) t^i.$$

Later on, we will prove that if $R = k[x_1, \ldots, x_n]$, then $\mathrm{HS}_M(t) = P(t)/(1-t)^n$, where $P(t) \in \mathbb{Z}[t, t^{-1}]$. This explains the usefulness of Hilbert series.

Exercise 2.2.10. Show that $\mathrm{HS}_{k[x]}(t) = 1 + t + t^2 + \ldots = 1/(1-t)$. Then, show that $\mathrm{HS}_{k[x_1,\ldots,x_n]}(t) = 1/(1-t)^n$ (hint: proceed by induction).

Another illustration of the usefulness of Hilbert series is the following. A ring is *Artinian* if there is no infinite proper descending chains of ideals (submodules). Suppose we have a graded ring R such that $R_i \neq 0$ for $i \gg 0$. Then,

$$(R_1) \supsetneq (R_2) \supsetneq \cdots$$

is an infinite descending chain of ideals. From this observation, we deduce that if R is a polynomial ring, M a finitely generated graded R-module, then M is Artinian if and only if $M_i = 0$ for all $i \gg 0$. It follows that M is Artinian if and only if $\mathrm{HS}_M(t) \in \mathbb{N}[t, t^{-1}]$.

2.3 Hilbert Polynomial

We observed that the Hilbert function of a polynomial ring coincide with a polynomial function in sufficiently high degrees. This is actually a general property that we will discuss in this section.

A morphism of graded R-modules $\phi : M \to N$ is *graded* if $\phi(M_i) \subseteq N_i$ for all i. An important motivation to consider graded maps is to rely on linear algebra by taking graded slices, in which case we get maps between vector spaces.

Example 2.3.1. Let R be a graded ring and let $f \in R_i$ for some $i > 0$. Then, the multiplication map $R \xrightarrow{\times f} R$ is not a graded map. However, the map $R(-i) \xrightarrow{\times f} R$ is a graded map. Write down its graded slices in degree $i, i+1, \ldots$.

A sequence of vector spaces and linear maps

$$V_\bullet \ : \ \cdots V_{j+1} \xrightarrow{\phi_{j+1}} V_j \xrightarrow{\phi_j} V_{j-1} \to \cdots$$

is a *complex* (of vector spaces) if $\mathrm{Im}(\phi_{j+1}) \subseteq \mathrm{Ker}(\phi_j)$ for all j. The complex $V_\bullet$ is said to be *exact at position* j if $\mathrm{Im}(\phi_{j+1}) = \mathrm{Ker}(\phi_j)$. It is called an *exact complex* if it is exact everywhere. The homology of the complex $V_\bullet$ is defined by

$$H_j(V_\bullet) := \mathrm{Ker}(\phi_j)/\mathrm{Im}(\phi_{j+1}), \ \forall j.$$

Exercise 2.3.2 (Euler characteristic)**.** Given a complex

$$V_\bullet : 0 \to V_n \to \cdots \to V_0 \to 0$$

of finite dimensional vector spaces, one has

$$\chi(V_\bullet) := \sum_{i=0}^{n} (-1)^i \mathrm{dim}(V_i) = \sum_{i=0}^{n} (-1)^i \mathrm{dim}(H_i(V_\bullet)).$$

This quantity is called the Euler characteristic of the complex $V_\bullet$. We observe that if $V_\bullet$ is an exact complex, then $\chi(V_\bullet) = 0$.

The above definitions we have made for sequences of vector spaces generalize in a straightforward way to sequences of modules and to sequences of graded modules with graded maps.

Theorem 2.3.3. *If M is a finitely generated graded R-module, then there exits a polynomial $f(x) \in \mathbb{Q}[x]$ such that*

$$\mathrm{HF}_M(i) = f(i) \textit{ for all } i \gg 0.$$

The polynomial f is called the Hilbert polynomial *of M and denoted by* $\mathrm{HP}_M(i)$.

Proof (sketch of). The proof is by induction on the number of variables in the ring over which M is defined. If there is no variables, i.e. $R = R_0 = k$, then M is a finite dimensional vector space and $\mathrm{HP}_M(i) = 0$ for all $i \gg 0$.

Suppose that the claimed property is true for $n-1$ variables. We build the exact sequence

$$0 \to K \to M(-1) \xrightarrow{\times x_n} M \to C \to 0.$$

K and C are finitely generated and since x_n "kills" them, they are actually finitely generated other the polynomial ring in $n-1$ variables. Therefore,

$$\mathrm{HF}_M(i) - \mathrm{HF}_M(i-1) \in \mathbb{Q}[i], \text{ for all } i \gg 0.$$

Form here, the conclusion follows from this general property (see for instance [Har77, Chapitre I, Proposition 7.3] for a proof): given a function $P : \mathbb{N} \to \mathbb{Z}$ such that $\Delta P(i) := P(i) - P(i-1)$ is a polynomial with rational coefficients (for $i \gg 0$), P is itself a polynomial with rational coefficients and has degree one greater than ΔP. □

Exercise 2.3.4 (Hilbert polynomial of a set of points in $\mathbb{P}^n_k$)**.**

1. Let p be a points in $\mathbb{P}^n_k$ and let $I(p)$ be its defining ideal. Compute the Hilbert polynomial of its coordinate ring $R/I(p)$ (where $R = k[x_0, \dots, x_n]$ is the coordinate ring of $\mathbb{P}^n_k$).
2. Let p_1, p_2 be two distinct points in $\mathbb{P}^n_k$; prove the exactness of the following sequence
$$0 \to I(p_1) \cap I(p_2) \xrightarrow{\psi} I(p_1) \oplus I(p_2) \xrightarrow{\phi} I(p_1) + I(p_2) \to 0$$
where $\psi(h) = (h, h)$ and $\phi(f, g) = f - g$.
3. Compute the Hilbert polynomial of $R/(I(p_1)+I(p_2))$ (what is the variety defined by $I(p_1) + I(p_2)$?).
4. Let $p_1, \dots, p_d$ be d distinct points in $\mathbb{P}^n_k$. Prove by induction that the Hilbert polynomial of $R/(\cap_j I(p_j))$ is a constant polynomial which is equal to d.

2.4 Dimension and Degree

The Hilbert polynomial of a graded quotient R/I contains useful information about the projective variety $V(I)$. In the next chapters we will prove that $\mathrm{HP}_{R/I}(i)$ is of the form
$$\frac{a_m}{m!} i^m + \frac{a_{m-1}}{(m-1)!} i^{m-1} + \cdots, \; a_i \in \mathbb{Z}, a_m > 0.$$

Definition 2.4.1. *For a homogeneous ideal $I \subseteq R = k[x_0, \dots, x_n]$ with*
$$\mathrm{HP}_{R/I}(i) = \frac{a_m}{m!} i^m + \frac{a_{m-1}}{(m-1)!} i^{m-1} + \cdots, a_m \neq 0$$
we define:

- *the dimension of $V(I) \subseteq \mathbb{P}^n_k$ as m,*
- *the codimension of $V(I) \subseteq \mathbb{P}^n_k$ as $n - m$,*
- *the degree of $V(I) \subseteq \mathbb{P}^n_k$ as a_m.*

The ideas of dimension and degree are obtained by slicing with hyperplanes. We will prove this in the next chapter.

Example 2.4.2. Taking again Example 2.2.8, check that the dimension and degree of I and J are the expected ones.

Exercise 2.4.3. Let $R = k[x_1, \dots, x_n]$ and $f \in R_d$ with $d \geq 1$. Compute the Hilbert series of $M = R/(f)$.

Exercise 2.4.4 (From Hilbert series to Hilbert polynomial)**.**

1. Let $R = k[x_1, \dots, x_n]$ and M be a positively graded R-module, i.e. $M = \oplus_{\nu \geq 0} M_\nu$. Prove that $\mathrm{HS}_M(t) = Q(t)/(1-t)^n$ where $Q(t) \in \mathbb{Z}[t]$.
2. Simplifying $\mathrm{HS}_M(t)$, one gets $\mathrm{HS}_M(t) = G(t)/(1-t)^s$ where $0 \leq s \leq n$ and $G(t) = \sum_{j=0}^d g_j t^j \in \mathbb{Z}[t]$ with $g_d \neq 0$ and $G(1) \neq 0$. Give the Hilbert polynomial of M in terms of the g_i's, in particular its leading term.

3. Use the above results to compute the Hilbert polynomial of the module $M = R/(f)$ as defined in Example 2.4.3. Check your results with Macaulay2.

Exercise 2.4.5 (Using Macaulay2)**.** Consider the variety in $\mathbb{P}^3$ defined by the ideal

$$I = \left(-y\,w + w^2, x\,w - 3\,z\,w, x^2y - y^2z - 9\,z^2w + z\,w^2, x^3 - 3\,x^2z - x\,y\,z + 3\,y\,z^2\right)$$

in $R = k[x, y, z]$.

1. Compute the Hilbert polynomial of $M = R/I$ with Macaulay2.What is its dimension and degree?
2. Compute the irreducible components of $V(I)$. Does the results agree with the previous computations?
3. Slice $V(I)$ with a general hyperplane of equation $l = 0$. What do you expect to obtain? What is the Hilbert polynomial of R/J, where $J = I + (l)$?

Exercise 2.4.6 (Using Macaulay2)**.** Let $R = k[x, y, z]$. and $I = (x^2 - xz, y^3 - yz^2)$. What is this variety? Draw a picture. Verify that Bézout theorem holds.

Exercise 2.4.7 (Using Macaulay2)**.** Consider the two curves in the plane defined by $y^2 - xz = 0$ and $x = 0$. How many intersection points have these two curves? How many points are expected by Bézout theorem? What is the Hilbert polynomial of the ideal corresponding to the intersection of these two curves? Discuss these observations.

Chapter 3

Free Resolutions and Regular Sequences

In this chapter, our goal is to use free modules in order to represent any finitely generated module. Indeed, free modules are the nicest possible modules and we have simple formulas for their Hilbert series.

3.1 Projective Modules

Free modules fit into the broader class of projective modules, that we briefly introduce. In what follows, R denotes a commutative ring with unit.

Definition 3.1.1. *A R-module P is* projective *if for any surjective map $f : A \to B$ and any map $g : P \to B$ of R-modules, there exists a map $h : P \to A$ such that the following diagram commutes (i.e. $g = f \circ h$).*

$$\begin{tikzcd}
 & P \arrow[d, "g"] \arrow[dl, dashed, "h"'] & \\
A \arrow[r, "f"] & B \arrow[r] & 0.
\end{tikzcd}$$

Here are key properties of projective modules

Lemma 3.1.2. *The following properties are equivalent:*

i) P is a projective R-module,

ii) Every exact sequence

$$0 \to N \to M \xrightarrow{f} P \to 0$$

splits (i.e. there exits $h : P \to M$ such that $f \circ h$ is the identity map),

iii) There exits a R-module K such that $P \oplus K \simeq F$ for some free R-module F.

Proof. *i)* ⇒ *ii)* follows from the definition of a projective module, by lifting the identity map $P \to P$ through the map f.

To prove that *ii*) implies *iii*) we first recall the following classical property (exercise):

$$0 \to N \to M \to P \to 0 \text{ splits } \Leftrightarrow M \simeq P \oplus N.$$

Now, consider the exact sequence

$$0 \to \operatorname{Ker}(\pi) \to \oplus_{p \in P} R \xrightarrow{\pi} P \to 0,$$

then $\oplus_{p \in P} R \simeq P \oplus \operatorname{Ker}(\pi)$.

We next prove that *iii*) implies *i*). Let F be such that $P \oplus K \simeq F$ and suppose given a surjective map $f : A \to B$ and a map $g : P \to B$; we have to show that the map g can be lifted to a map $h : P \to A$. For that purpose, we write $F = P \oplus K \simeq \oplus_i R b_i$ (free module with basis the b_i's) and consider the commutative diagram

$$\begin{tikzcd}
 & F \arrow[d] \arrow[ddl, dashed, "h"'] & \\
 & P \arrow[d, "g"] & \\
A \arrow[r, "f"] & B \arrow[r] & 0
\end{tikzcd}$$

where the maps are defined as follows: each b_i is mapped canonically to P, then to $g(b_i) \in B$. Now, since f is surjective, each $g(b_i)$ has a preimage via f that we denote by m_i. The map h is then defined by sending b_i to m_i for all i. Finally, the map $h_{|P} : P \to A$ is the lifting of g we wanted to prove that P is projective. □

Projective modules are very interesting because they allow to get commutative diagrams and to lift through surjective maps. It is important to notice that (see [Eis95, Chapter 4, Exercise 4.11]):

- over a local ring, a projective module is a free module,
- *a finitely generated graded projective module over a polynomial ring $k[x_1, \ldots, x_n]$, over a field k, is a graded free module.*

3.2 Free Resolutions

In what follows R denotes a polynomial ring over a field k. We begin with an illustrative example.

Example 3.2.1. Let $R = k[x, y, z]$ and consider the homogeneous ideal $I = (x^3 + y^3 + z^3)$. In Example 2.2.8 we have seen that

$$\mathrm{HF}_{R/I}(i) = \mathrm{HF}_R(i) - \mathrm{HF}_{R(-3)}(i).$$

Actually, this property follows from the following graded exact sequence:

$$0 \to R(-3) \xrightarrow{\times (x^3+y^3+z^3)} R \to R/I \to 0.$$

What happens if we slice the projective curve of equation $x^3 + y^3 + z^3 = 0$ with the line $x = 0$, i.e. if we consider the ideal $J = I + (x)$? It is not difficult to check that we have the following exact sequence

$$0 \to R(-4) \xrightarrow{\begin{pmatrix} x^3+y^3+z^3 \\ -x \end{pmatrix}} R(-1) \oplus R(-3) \xrightarrow{(x, x^3+y^3+z^3)} R \to R/J \to 0$$

(exercise: check exactness of this sequence.) From here, we deduce that

$$\begin{aligned} \mathrm{HP}_{R/J}(i) &= \mathrm{HP}_R(i) - \mathrm{HP}_{R(-1)}(i) - \mathrm{HP}_{R(-3)}(i) + \mathrm{HP}_{R(-4)}(i) \\ &= \binom{i+2}{2} - \binom{i+1}{2} - \binom{i-1}{2} - \binom{i-2}{2} = 3. \end{aligned}$$

It turns out that 3 is precisely the number of intersection points between the projective curve of equation $x^3 + y^3 + z^3 = 0$ and the line $x = 0$ by Bézout Theorem.

Exercise 3.2.2. Let $R = k[x, y, z]$ and consider $f(x, y, z)$ and $g(x, y, z)$ two homogeneous polynomials in R of degree d, e respectively. Assuming that f and g have no common factor, show that we have and exact sequence of the form

$$0 \to R(-d-e) \to R(-d) \oplus R(-e) \to R \to R/(f, g) \to 0$$

and deduce that $\mathrm{HP}_{R/(f,g)}(i) = de$, as expected by Bézout Theorem.

The above example shows the usefulness of having a *resolution* of a quotient ring by free modules, particularly to deduce the Hilbert polynomial, hence dimension and degree. It turns out that such a resolution always exists over R; this is a celebrated result of Hilbert.

Theorem 3.2.3 (Hilbert Syzygy Theorem)**.** *Let M be a finitely generated graded module over $R = k[x_1, \dots, x_n]$, k a field, then there exists a graded exact sequence of modules*

$$0 \to F_n \to F_{n-1} \to \cdots \to F_1 \to F_0 \to M \to 0$$

where the F_i's are finitely generated free modules.

Proof. See [Eis95, Chapter I, Theorem 1.13] and references therein. □

The sequence $0 \to F_n \to \cdots \to F_0$ is called a *finite free resolution* of the module M. We notice that the existence of free resolutions is easy to prove: since M is finitely generated, there exists a surjective map $R^{n_0} \xrightarrow{\phi_0} M \to 0$. Now, as R is Noetherian, the kernel of ϕ_0 is also finitely generated and hence one gets an exact sequence of the form

$$R^{n_1} \xrightarrow{\phi_1} R^{n_0} \xrightarrow{\phi_0} M \to 0.$$

Continuing this way, we obtain a free resolution. So, the key point of Hilbert Syzygy Theorem is to prove that M has a *finite* free resolution over a polynomial ring, which is moreover of length $\leq n$ (recall that n is the number of variables of R).

Example 3.2.4. Let $T = k[x]/(x^2)$, then a free resolution of $(x) \subset T$ is given by

$$\ldots \to T(-2) \xrightarrow{\times x} T(-1) \xrightarrow{\times x} (x) \to 0.$$

We notice that it is infinite.

Notation 3.2.5. Of particular importance in a finite free resolution are the shifts in grading, as illustrated in Example 3.2.1. In `Macaulay2`, these shifts are obtained as *Betti tables*. For instance, the betti table of the free resolution

$$F_0 = R \leftarrow F_1 = R(-1) \oplus R(-3) \leftarrow R(-4)$$

(`Macaulay2` writes resolutions from the left side) is the following

total:	1	2	1
0 :	1	1	·
1 :	·	·	·
2 :	·	1	1

The first row gives the ranks of the F_i's, from F_0 to F_2. The other rows gives the shifts in grading that have to be added to the expected shift $F_i(-i)$ for all i (see below).

Minimal free resolutions. A free resolution is called *minimal* if there are no constant terms in any of the maps (all entries belongs to $\oplus_{i \geq 1} R_i$). Indeed, if a constant term appears in a map, then it can be simplified; for instance

$$0 \to R(-3) \xrightarrow{\begin{pmatrix} y \\ -1 \end{pmatrix}} R(-2) \oplus R(-3) \xrightarrow{(x^2, yx^2)} I \to 0,$$

can be simplified to

$$0 \to R(-2) \xrightarrow{(x^2)} I \to 0.$$

In the other direction, the exact sequence $0 \to R \xrightarrow{(1)} R \to 0$ can be added to any other exact sequence $\cdots F_i \xrightarrow{d_i} F_{i-1} \to \cdots$ to get

$$\cdots \to F_{i+1} \to F_i \oplus R \xrightarrow{\begin{pmatrix} d_i & 0 \\ 0 & 1 \end{pmatrix}} F_{i-1} \oplus R \to F_{i-2} \to \cdots$$

As an important fact, matrices in free resolutions are not unique (as choice of generators for ideals), but the free modules that appear in a minimal free resolution are unique; see [Eis95, Chapter 20, Theorem 20.2].

Hilbert series. Let us see some consequences of the existence of FFR on Hilbert series. Let $R = k[x_1, \ldots, x_n]$ and let M be a graded R-module with a FFR (Finite Free Resolution)

$$0 \to F_n \to F_{n-1} \to \cdots \to F_1 \to F_0$$

where $F_k \simeq \oplus_{i=1}^{r_k} R(-a_{k,i})$, $r_k = \operatorname{rank}(F_k)$.

The fact that the Hilbert function of M becomes a polynomial in sufficiently large degrees is because this is obviously true for free modules. More precisely:

$$\begin{aligned} \mathrm{HP}_M(i) &= \sum_{j=0}^{n} (-1)^j \mathrm{HP}_{F_j}(i) \\ &= \sum_{j=0}^{n} (-1)^j \sum_{i=1}^{r_j} \binom{n-1+i-a_{j,i}}{n-1}. \end{aligned}$$

Another important and easy consequence of the FFR if the following. We have seen that $\mathrm{HS}_R(t) = 1/(1-t)^n$. It follows that $\mathrm{HS}_{R(-a)} = t^a/(1-t)^n$ and hence

$$\mathrm{HS}_{F_k}(t) = \frac{\sum_{i=1}^{r_k} t^{a_{k,i}}}{(1-t)^n}.$$

Thus, we have just proved that $\mathrm{HS}_M(t) = P_M(t)/(1-t)^n$ where $P_M(t) \in \mathbb{Z}[t, t^{-1}]$.

3.3 Slicing by a hyperplane

Our next goal is to understand what happens when one slices a variety with a hyperplane, or more generally with a hypersurface.

Lemma 3.3.1. *Let $I \subseteq R$ be a homogeneous ideal and $f \in R_d$. Then, we have the following graded exact sequence*

$$0 \to R(-d)/(I:f) \to R/I \to R/(I+(f)) \to 0.$$

Proof. The canonical sequence (we use the notation (I,f) for the ideal $I+(f)$)

$$0 \to (I,f)/I \to R/I \to R/(I,f) \to 0$$

is clearly exact. Now, the multiplication by the homogeneous polynomial f:

$$R(-d) \xrightarrow{\times f} (I,f)/I \to 0$$

has kernel equals to $(I:f)$, so we deduce that $R(-d)/(I:f) \simeq (I,f)/I$ (graded isomorphism). □

Let $f \in R$; f is a *nonzero divisor* on M if $f.m \neq 0$ for all $m \neq 0$ in M. Therefore, f is a nonzero divisor on R/I if and only if $(I:f) = I$. Suppose that

$$\mathrm{HP}_{R/I}(i) = \frac{a_m}{m!} i^m + \cdots .$$

Thus, if f is a homogeneous linear form which is not a zerodivisor on R/I, then by Lemma 3.3.1 we deduce that

$$\mathrm{HP}_{R/(I,f)}(i) = \mathrm{HP}_{R/I}(i) - HP_{R/I}(i-1) = \frac{a_m}{(m-1)!} i^{m-1} + \cdots .$$

In other words, the dimension drops by one and degree is left unchanged, as we claimed in the previous chapter. Moreover, by repeating this slicing process, we arrive at a constant Hilbert polynomial.

The previous argument shows what we expected, but the problem is about the existence of a nonzero divisor. Indeed, f is a nonzero divisor if and only if $(I : f) = I$. Let $I = \cap_{i=1}^{r} q_i$ be a minimal primary decomposition of I. We have seen that if $f \notin p_i = \sqrt{q_i}$ then $(q_i : f) = q_i$, so since $(I : f) = \cap(q_i : f)$ we might have an issue if I has an $\mathfrak{m}$-primary component, where $\mathfrak{m} = (x_1, \ldots, x_n)$, because then $f \in \mathfrak{m}$. For instance, if f is a linear form then $(\mathfrak{m}^2 : f) = \mathfrak{m}$. Nevertheless, this difficulty can be overcome as follows.

If I as a $\mathfrak{m}$-primary component, define I' to be the same ideal but remove this component: $I = \cap_{i=1}^{r} q_i$ and q_r $\mathfrak{m}$-primary, then $I' = \cap_{i=1}^{r-1} q_i$. Now, $HP_{R/I} = \mathrm{HP}_{R/I'}$ because of the exact sequence

$$0 \to R/(I \cap J) \to R/I \oplus R/J \to R/(I + J) \to 0$$

that holds for any ideals I, J, and because $\mathrm{HP}_{R/q_r}(i) = 0$ (show this! - the component q_r, which is such that $\sqrt{q_r} = \mathfrak{m}$, is geometrically irrelevant in the projective space defined by R).

Finally, to validate our previous argument, we claim that there exists a linear form $f \in R_1$ such that $f \notin \cup_{i=1}^{r-1} p_i$, where $p_i = \sqrt{q_i}$. This implies that f is a nonzero divisor of R/I' and we are done. The existence of this linear form is a consequence of the following lemma.

Lemma 3.3.2 (Prime avoidance). *If $I \subseteq \cup_{i=1}^{n} p_i$, with p_i prime ideals, then $I \subseteq p_i$ for some i.*

Proof. We prove that if $I \not\subseteq p_i$ for all i, then $I \not\subseteq \cup_{i=1}^{n} p_i$. We proceed by induction on n.

The case $n = 1$ is trivial. Suppose that $I \not\subseteq p_i$ for all i and $I \subseteq \cup_{i=1}^{n} p_i$. By our inductive assumption, $I \not\subseteq \cup_{j \neq i} p_j$ for all i. This means that for all i there exists $x_i \in I$ such that $x_i \notin \cup_{j \neq i} p_j$. One may actually assume that $x_i \in p_i$, so that $x_i \in I \cap p_i$, because otherwise we get a contraction since we assumed that $I \subseteq \cup_{i=1}^{n} p_i$.

Now, consider the element $x = \sum_{i=1}^{n} x_1 \cdots \widehat{x_i} \cdots x_n$. By construction, $x \in I$. Let us fix an integer k. Then $x_1 \cdots \widehat{x_k} \cdots x_n \notin p_k$ because for all $j \neq k$, $x_j \notin \cup_{i \neq j} p_i \supset p_k$ and p_k is prime. It follows that $x \notin p_k$ because all the other monomials in x belong to p_k. In conclusion, for any k, $x \notin p_k$, so $x \notin \cup_{k=1}^{n} p_k$, which is a contradiction. □

Indeed, $\mathfrak{m} \not\subseteq \cup_{i=1}^{r-1} p_i$ (union of associated primes of I') so there must be a linear from $f \in \mathfrak{m}_1$ such that $f \notin \cup_{i=1}^{r-1} p_i$.

3.4 Regular sequences

We conclude this chapter with the concept of regular sequences, which roughly correspond to the slicing iterative process we have just considered, but taking the whole space as a starting point.

Definition 3.4.1. *Let M be a graded R-module. A* regular sequence *on M is a sequence of homogeneous polynomials $\{f_1, \ldots, f_m\}$ such that*

- *f_1 is a nonzero divisor on M,*
- *f_i is a nonzero divisor on $M/(f_1, \ldots, f_{i-1})M$, for all $i \geq 1$.*

Example 3.4.2. The sequence $\{x_1, \ldots, x_n\}$ is a regular sequence in $R = k[x_1, \ldots, x_n]$

See Exercise 3.5.3 to discover some useful properties of regular sequences.

When one computes a free resolution for an ideal generated by a regular sequence, one should get only the trivial relations. For instance, if $I = (f_1, f_2, f_3)$, then one should get

$$0 \to R \xrightarrow{d_2} R^3 \xrightarrow{d_1} R^3 \xrightarrow{d_0} R \to R/I \to 0,$$

where

$$d_2 = \begin{pmatrix} f_3 \\ -f_2 \\ f_1 \end{pmatrix}, \; d_1 = \begin{pmatrix} -f_2 & -f_3 & 0 \\ f_1 & 0 & -f_3 \\ 0 & f_1 & f_2 \end{pmatrix}, \; d_0 = \begin{pmatrix} f_1 & f_2 & f_3 \end{pmatrix}.$$

This type of complex is known as a *Koszul complex.*

Koszul Complex. Let A be a ring. Given a sequence $\mathbf{x} := (x_1, \ldots, x_n)$ of n elements, its Koszul complex, denoted by $K_\bullet(x)$, is defined as follows: let $K_i(\mathbf{x})$ be the exterior power $\wedge^i(A^n)$. Then, if $\{e_1, \ldots, e_n\}$ denotes the canonical basis of A^n, $K_0(\mathbf{x}) = A$ and for all $p \in \mathbb{N}^*$

$$K_p(\mathbf{x}) = \bigoplus_{1 \leq i_1 < \cdots < i_p \leq n} Ae_{i_1} \wedge \cdots \wedge e_{i_p}.$$

Moreover, the differential map $d_p : K_p(\mathbf{x}) \to K_{p-1}(\mathbf{x})$ sends a basis element $e_{i_1} \wedge \cdots \wedge e_{i_p}$ to

$$d_p(e_{i_1} \wedge \cdots \wedge e_{i_p}) := \sum_{k=1}^{p} (-1)^{k+1} x_{i_k} e_{i_1} \wedge \cdots \wedge \widehat{e_{i_k}} \wedge \cdots \wedge e_{i_p}.$$

It is immediate to check that this defines a complex, that is to say that $d_{p-1} \circ d_p = 0$ for all p.

If M is a A-module, then we define the homological Koszul complex of the sequence $\mathbf{x}$ over M by $K_\bullet(\mathbf{x}; M) := K_\bullet(\mathbf{x}) \otimes_A M = K_\bullet(\mathbf{x}; A) \otimes_A M$. For all integer p we will denote by $H_p(\mathbf{x}; M)$ the pth homology A-module of the Koszul complex $K_\bullet(\mathbf{x}; M)$.

Proposition 3.4.3. *With the above notation,*

(i) *The ideals* $\mathrm{ann}_A(M)$ *and* $(\mathbf{x})$ *of A annihilates all the homology modules of the Koszul complex $K_\bullet(\mathbf{x}; M)$.*

(ii) *If $\mathbf{x}$ is a M-regular sequence, then $H_p(\mathbf{x}; M) = 0$ for all $p \geq 1$.*

Proof. For the first point, it suffices to check that for all integers $p \geq 0$ and $j = 1, \ldots, n$, and all $x \in K_p(\mathbf{x}; M)$ we have

$$d_{p+1}\sigma_p^j(x) + \sigma_{p_1}^j d_p(x) = x_j x,$$

where the map $\sigma_p^j : K_p(\mathbf{x}; M) \to K_{p+1}(\mathbf{x}; M)$ sends the basis element $e_{i_1} \wedge \cdots \wedge e_{i_p}$ to the element $e_j \wedge e_{i_1} \wedge \cdots \wedge e_{i_p}$.

To prove the second statement requires some homological algebra (see for instance [Eis95, Appendix 3]). We proceed by induction on n. If $n = 1$, then we have $H_1(x_1; M) = \mathrm{Ker}(M \xrightarrow{\times x1} M) = 0$. Now, assume that we have proved (ii) for all integer $1, \dots, t-1$ and put $\mathbf{x}' := (x_1, \dots, x_{n-1})$. It is easy to check that we have the following exact sequence of complexes:

$$0 \to K_\bullet(\mathbf{x}'; M) \hookrightarrow K_\bullet(\mathbf{x}; M) \xrightarrow{\pi} K_\bullet(\mathbf{x}'; M)[-1] \to 0$$

where $K_\bullet[-1]$ is the "left translation" of $K_\bullet$ (i.e. $K_p[-1] := K_{p-1}$ and $d_p[-1] := d_{p-1}$) and the A-linear map π sends a basis element $e_{i_1} \wedge \cdots \wedge e_{i_p}$ to $e_{i_1} \wedge \cdots \wedge e_{i_{p-1}}$ if $i_p = n$, or 0 otherwise. This exact sequence gives rise to the long exact sequence of homology groups (we leave to the reader the explicitation of the connecting map)

$$\cdots \to H_p(\mathbf{x}'; M) \xrightarrow{\times(-1)^p x_n} H_p(\mathbf{x}'; M) \to H_p(\mathbf{x}; M) \to H_{p-1}(\mathbf{x}; M) \to \cdots$$

which immediately shows, with the inductive hypothesis, that $H_p(\mathbf{x}; M) = 0$ for all $p > 1$. To finish the proof, we examine the right end of the long exact sequence:

$$0 = H_1(\mathbf{x}'; M) \to H_1(\mathbf{x}; M) \to H_0(\mathbf{x}'; M) \xrightarrow{\times x_n} H_0(\mathbf{x}'; M) \to \cdots$$

Since $\mathbf{x}$ is assumed to be a M-regular sequence, then the map on the right is injective and it follows that $H_1(\mathbf{x}; M) = 0$. □

Remark 3.4.4. The statement (ii) becomes an equivalence in the graded or local case. More precisely, if either

- A is a graded ring, M is a graded A-module of finite type and all the x_i's are homogeneous element with positive degree,
- A is a local noetherian ring $(A, \mathfrak{m})$ and for all $i = 1, \dots, n$ we have $x_i \in \mathfrak{m}$,

then $\mathbf{x}$ is a M-regular sequence if and only if $H_p(\mathbf{x}; M) = 0$ for all $p \geq 1$, if and only if $H_1(\mathbf{x}; M) = 0$. As a corollary, this proves that, under the same assumptions, $\mathbf{x}$ is a regular sequence independently of the order of its elements.

Note that if A is a graded ring, then the Koszul complex $K_\bullet(\mathbf{x}; M)$ inherits straightforwardly of this grading. For instance, if A is a $\mathbb{Z}$-graded ring and the elements $x_1, \dots, x_n$ are homogeneous of degree $d_1, \dots, d_n$, respectively, then the Koszul complex is graded by $K_0(\mathbf{x}; A) = A(0)$ and, for all $p \geq 1$,

$$K_p(\mathbf{x}; A) = \bigoplus_{1 \leq i_1 < \cdots < i_p \leq n} A(-d_{i_1} - \cdots - d_{i_p}).$$

3.5 Exercises

Exercise 3.5.1. Let $f(x_0, x_1, x_2, x_3)$ and $g(x_0, x_1, x_2, x_3)$ be two homogeneous polynomials in $R = \mathcal{C}[x_0, x_1, x_2, x_3]$ of degree 3 and 2 that define a cubic surface $\mathcal{H}$ and a quadratic surface $\mathcal{Q}$ in $\mathbb{P}^3$, respectively.

1. We assume that $\mathcal{H}$ and $\mathcal{Q}$ intersect in a curve $\mathcal{C}$. Show that this implies that (f, g) is a regular sequence in R.
2. Give a minimal graded finite free resolution of R/I.
3. Compute the Hilbert polynomial of the intersection curve $\mathcal{C}$. What is the degree of this curve?

Solution.

1. The polynomial f being nonzero, it is a nonzero divisor in R. Since $\mathcal{H}$ and $\mathcal{Q}$ cut out a curve then f and g have no common factors (otherwise $\mathcal{H}$ and $\mathcal{Q}$ would have a surface as a common component of dimension 2). Now, if h and k are polynomials in R such that $hf + kg = 0$ we deduce that f divides k, which means that g is not a zero divisor in $R/(f)$.
2. Since (f, g) is a regular sequence in R, its associated Koszul complex is a F.F.R. of R/I:

$$0 \to R(-5) \xrightarrow{\begin{pmatrix} -g \\ f \end{pmatrix}} R(-3) \oplus R(-2) \xrightarrow{(f,g)} R.$$

   It is clearly a minimal resolution.
3. Using the F.F.R. of R/I we get:

$$\begin{aligned} \mathrm{HP}(R/I, t) &= \mathrm{HP}(R, t) - \mathrm{HP}(R, t-2) - \mathrm{HP}(R, t-3) + \mathrm{HP}(R, t-5) \\ &= \binom{t+3}{3} - \binom{t+1}{3} - \binom{t}{3} + \binom{t-2}{3} \\ &= 6t - 3. \end{aligned}$$

   The curve is of degree 6.

Exercise 3.5.2. Let $R = \mathcal{C}[x_1, \ldots, x_n]$ and let f and g be two homogeneous polynomials of positive degree d and e respectively. We assume that f and g have no common factor in R and we denote by I the ideal generated by f and g, i.e. $I = (f, g) \subset R$.

1. Show that R/I has a finite free resolution of the form

$$0 \to F_2 \to F_1 \to F_0 \to R/I \to 0.$$

   Describe explicitly the graded free R-modules F_i and the maps in this finite free resolution.
2. What is the Hilbert series of R? What are the Hilbert series of F_0, F_1 and F_2?
3. Deduce that the Hilbert series of R/I is of the form $P(t)/(1-t)^{n-2}$ where $P(t) \in \mathbb{Z}[t]$ is such that $P(1) \neq 0$.
4. Finally, deduce that $V(I)$ is of dimension $n-2$ and degree $P(1)$. What is the value of $P(1)$ in terms of d and e? (hint: $1/(1-t)^{n-2}$ is the Hilbert series of a polynomial ring in $n-2$ variables).

Solution.

1. $F_0 = R$ and $F_1 = R(-d) \oplus R(-e)$ as the first map is given by the generators, that is to say
$$\partial_1 : F_1 \to F_0 : (p, q) \mapsto pf + qg.$$
The kernel of ∂_1 corresponds to couples p, q such that $pf + qg = 0$. Since f and g have no common factors and that R is a UFD we deduce that f divides q, i.e. $q = fq'$, and g divides p, i.e. $p = gp'$. In addition we have $gp'f + fg'g = 0$ from we deduce that $p' + q' = 0$. Therefore, the kernel of ∂_1 corresponds to elements of the form $h(-g, f)$ where h is any homogeneous polynomials; it is hence isomorphic to R. Taking into account the grading we get $F_2 = R(-d-e)$ and the map
$$F_2 \to F_1 : h \mapsto h(-g, f)$$
is injective, so that the resolution stops at the second step.
2. From the definition: $\mathrm{HS}(R, t) = \mathrm{HS}(F_0, t) = 1/(1-t)^n$, $\mathrm{HS}(F_1, t) = (t^d + t^e)/(1-t)^n$ and $\mathrm{HS}(F_2, t) = t^{d+e}/(1-t)^n$.
3. Applying Hilbert series to the exact sequence obtained in the first question we get
$$\begin{aligned}\mathrm{HS}(R/I, t) &= \mathrm{HS}(F_0, t) - \mathrm{HS}(F_1, t) + \mathrm{HS}(F_2, t)\\ &= \frac{1 - t^d - t^e + t^{d+e}}{(1-t)^n} = \frac{(1-t^d)(1-t^e)}{(1-t)^n}\\ &= \frac{(1 + t + \cdots + t^{d-1})(1 + t + \cdots + t^{e-1})}{(1-t)^{n-2}} =: \frac{P(t)}{(1-t)^{n-2}}.\end{aligned}$$
We have $P(1) = de \neq 0$.
4. We know that
$$\mathrm{HS}(\mathcal{C}[x_1, \ldots, x_{n-2}], t) = \frac{1}{(1-t)^{n-2}} = (1 + (n-2)t + \cdots + \binom{n-3+i}{n-3} t^i + \cdots)$$
where $\binom{n-3+i}{n-3} = \frac{(i+n-3)(i+n-4)\cdots(i+1)}{(n-3)!}$ is a polynomial in i of degree $n-3$ and leading coefficient equal to $1/(n-3)!$. Now, if $P(t) := \sum_{j=0}^{l} c_j t^j$ then the coefficient of t^i in $\mathrm{HS}(R/I, t)$ is equal to
$$\sum_{j=0}^{l} c_j \binom{n-3+i-j}{n-3},$$
assuming that i is sufficiently high. This is a polynomial of degree $n-3$ in i and its leading coefficient is equal to
$$\sum_{j=0}^{l} c_j \times \frac{1}{(n-3)!} = \frac{P(1)}{(n-3)!} = \frac{de}{(n-3)!}.$$
The Hilbert polynomial $\mathrm{HP}(R/I, i)$ of R/I is hence a polynomial of degree $n-3$ and leading coefficient $de/(n-3)!$. It follows that $V(I) \subset \mathbb{P}^{n-1}$ is of dimension $n-3$ (codimension 2) and of degree de.

Exercise 3.5.3. A sequence $\{a_1, \dots, a_s\}$ of elements in a commutative ring R is called a *regular sequence* if

i) $(a_1, \dots, a_s) \neq R$

ii) a_i is a nonzerodivisor in $R/(a_1, \dots, a_{i-1})$ for all $i = 1, \dots, s$.

An ideal generated by a regular sequence in called a *complete intersection ideal.*

1. Show that the property of being a regular sequence depends on the order of the elements. For that purpose, one can consider the ring $R = k[x, y, z]$ and the two sequences $(x, y(1-x), z(1-x))$ and $(y(1-x), z(1-x), x)$.
2. Let $\{a_1, \dots, a_s\}$ be a regular sequence in R. Show that the sequence obtained by permutation of a_i and a_{i+1} is regular if and only if a_{i+1} is not a zerodivisor in $R/(a_1, \dots, a_{i-1})$.
3. (Graded Nakayama Lemma) Let $R = \oplus R_i$ a graded ring and M a graded R-module such that $M_i = 0$ for i sufficiently negative. Set $R_+ := \oplus_{i>0} R_i$. Show that if $R_+ M = M$ then $M = 0$.
4. Let $\{a_1, \dots, a_s\}$ be a regular sequence of homogeneous elements in a graded ring R. Then, show that this sequence remains regular after any permutation of its elements (hint: use previous exercises).
5. Let $\{a_1, \dots, a_s\}$ be a regular sequence in a commutative ring R. Show that for all
$$g \in \mathrm{Syz}(a_1, \dots, a_s) := \{(h_1, \dots, h_s) \in R^s \text{ such that } \sum h_i a_i = 0\}$$
there exists a skew-symmetric matrix M with coefficients in R such that
$$g = M \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_s \end{pmatrix}$$

Solution.

1. The sequence $(x, y(1-x), z(1-x))$ is regular whereas $(y(1-x), z(1-x), x)$ is not regular in R.
2. It is enough to prove that a_i is not a zero divisor in $R/(a_1, \dots, a_{i-1}, a_{i+1})$. Let $a \in R$ such that
$$a a_i = b_1 a_1 + \cdots + b_{i-1} a_{i-1} + b_{i+1} a_{i+1}$$
where the b_j's are in R. Since a_{i+1} is not a zero divisor in $R/(a_1, \dots, a_i)$, we deduce that there exists $c_j \in R$ such that $b_{i+1} = c_1 a_1 + \cdots + c_i a_i$. It follows that $a_i(a - c_i a_{i+1})$ belongs to the ideal $(a_1, \dots, a_{i-1})$, and hence that $a - c_i a_{i+1}$ belongs to $(a_1, \dots, a_{i-1})$ because a_i is not a zero divisor in $R/(a_1, \dots, a_{i-1})$. Therefore $a = 0$ in $R/(a_1, \dots, a_{i-1}, a_{i+1})$.
3. Just notice that $R_+ M_i \subset Mi + 1$ for any i, so there is no smallest integer i_0 such that $M_{i_0} \neq 0$, otherwise we get a contradiction. So $M = 0$.
4. It is enough to prove the claim for a permutation of two successive elements, so using question 2., it is enough to prove that a_{i+1} is not a zero divisor in $R/(a_1, \dots, a_{i-1})$. Let $a \in R$ such that $a a_{i+1} \in (a_1, \dots, a_{i-1})$. We have $a \in M :=$

$\mathrm{ann}_{R/(a_1,\ldots,a_{i-1})}(a_{i+1})$, which is graded module. We will prove that $M \subset (a_i)M$, which implies that $M = 0$ by question 3. Now,

$$aa_{i+1} \in (a_1, \ldots, a_{i-1}) \subset (a_1, \ldots, a_i).$$

This implies that $a = b_1a_1 + \ldots + b_{i-1}a_{i-1} + b_ia_i$ because a_{i+1} is not a zero divisor in $R/(a_{,} \ldots, a_i)$. Multiplying this equality by a_{i+1} we deduce that $b_ia_ia_{i+1} \in (a_1, \ldots, a_{i-1})$. But since a_i is not a zero divisor in $R/(a_1, \ldots, a_{i-1})$, we get that $b_ia_{i+1} \in (a_1, \ldots, a_{i-1})$, i.e. $b_i \in M$. As $a = a_ib_i$ in $R/(a_1, \ldots, a_{i-1})$, the claimed property is proved.

5. We proceed by induction on s. Consider the case $s = 2$. Let $g = (g_1, g_2)$ such that $g_1a_1 + g_2a_2 = 0$. We have $g_2a_2 = 0$ in $R/(a_1)$, so $g_2 = 0$ in $R/(a_1)$ and hence $g_2 = g_2'a_1$. It follows that $a_1(g_1 + g_2'a_2) = 0$, hence $g_1 + g_2'a_2 = 0$, i.e. $g_1 = -g_2'a_2$. In summary

$$\begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = \begin{pmatrix} 0 & -g_2' \\ g_2' & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

Now, assume that the property holds for $s-1$. Let $g = (g_1, \ldots, g_s)$ be such that $\sum_{i=1}^s g_ia_i = 0$. As $g_sa_s \in (a_1, \ldots, a_{s-1})$, we deduce that $g_s \in (a_1, \ldots, a_{s-1})$, i.e.

$$g_s = h_1a_1 + \cdots + h_{s-1}a_{s-1}.$$

We get

$$\sum_{i=1}^s g_ia_i = 0 = (g_1 + h_1a_s)a_1 + \cdots + (g_{s-1} + h_{s-1}a_s)a_{s-1}.$$

By our inductive assumption, there exists a skew-symmetric matrix N such that

$$\begin{pmatrix} g_1 + h_1a_s \\ \vdots \\ g_{s-1} + h_{s-1}a_s \end{pmatrix} = N \begin{pmatrix} a_1 \\ \vdots \\ a_{s-1} \end{pmatrix},$$

which can be rewritten as

$$\begin{pmatrix} g_1 \\ \vdots \\ g_{s-1} \end{pmatrix} = N \begin{pmatrix} a_1 \\ \vdots \\ a_{s-1} \end{pmatrix} - a_s \begin{pmatrix} h_1 \\ \vdots \\ h_{s-1} \end{pmatrix}.$$

Therefore, we deduce that

$$\begin{pmatrix} g_1 \\ \vdots \\ g_{s-1} \\ g_s \end{pmatrix} = \left[\left(\begin{array}{ccc|c} & & & 0 \\ & N & & 0 \\ & & & \vdots \\ \hline 0 & \ldots & 0 & 0 \end{array} \right) + \left(\begin{array}{ccc|c} & & & -h_1 \\ & 0 & & \vdots \\ & & & -h_{s-1} \\ \hline h_1 & \ldots & h_{s-1} & 0 \end{array} \right) \right] \begin{pmatrix} a_1 \\ \vdots \\ a_{s-1} \\ a_s \end{pmatrix},$$

which concludes the proof.

Chapter 4

Gröbner Bases

The goal of this chapter is to provide algorithms that allow us to perform computations over a polynomial ring. In what follows, k denotes a field.

4.1 Multivariable Division Algorithm

Euclidian division. Let $f, g \in k[x]$, $g \neq 0$, then there exists a unique couple of polynomials $q, r \in k[x]$ such that $f = qg + r$ and $r = 0$ or $\deg(r) < \deg(g)$.

This very classical result can be turned into the following algorithm. Given a univariate polynomial $p = a_0 x^m + \cdots + a_m$ with $a_0 \neq 0$, we define the leading term of p as $\mathrm{LT}(p) = a_0 x^m$.

Algorithm 1 Euclidian division

```
  $q := 0$, $r := f$
  while $r \neq 0$ and $\mathrm{LT}(g) | \mathrm{LT}(r)$ do
      $q := q + \mathrm{LT}(r)/\mathrm{LT}(g)$
      $r := r - \frac{\mathrm{LT}(r)}{\mathrm{LT}(g)} g$
  end while
```

Exercise 4.1.1. Show that $k[x]$ is principal: any ideal is generated by a single element.

Monomial Orders. To generalize the previous Euclidian division to the multivariate setting, it is necessary to have a way to compare monomials in a polynomial ring. Given a monomial $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$, we set $\alpha := (\alpha_1, \ldots, \alpha_n)$ and $|\alpha| = \sum_{i=1}^{n} \alpha_i$.

Definition 4.1.2. *A monomial order on the polynomial ring $R = k[x_1, \ldots, x_n]$ is a relation $>$ such that*

- *$>$ is a total order (for any α, β, one always have $\alpha > \beta$, or $\alpha < \beta$ or $\alpha = \beta$),*
- *for any γ, $\alpha > \beta$ implies that $\alpha + \gamma > \beta + \gamma$,*

- *any non-empty subset has a smallest element.*

Here are three classical monomial orders:

- *Lexicographic:* $\alpha > \beta$ if the leftmost nonzero entry of $\alpha - \beta$ is positive.
- *Graded lexicographic:* $\alpha > \beta$ if $|\alpha| > |\beta|$ or $|\alpha| = |\beta|$ and the leftmost nonzero entry of $\alpha - \beta$ is positive.
- *Graded reverse lexicographic:* $\alpha > \beta$ if $|\alpha| > |\beta|$ or $|\alpha| = |\beta|$ and the rightmost nonzero entry of $\alpha - \beta$ is negative.

Example 4.1.3. Let $R = k[x, y, z]$. Then,

$$x >_{lex} yz^2 \text{ and } yz^2 >_{grlex} x,$$

$$x^3y^5z^2 >_{grlex} x^2y^7z \text{ and } x^2y^7z >_{grevlex} x^3y^5z^2.$$

Multivariate Division. We begin with some notation. Let $R = k[x_1, \ldots, x_n]$ and $>$ be a monomial order. If $f = \sum_\alpha c_\alpha x^\alpha$ then we define the *leading term* of f as $\mathrm{LT}(f) := c_\alpha x^\alpha$, where x^α is the biggest monomial such that $c_\alpha \neq 0$; we also define the *leading monomial* as $\mathrm{LM}(f) := x_\alpha$ and the *leading coefficient* as $\mathrm{LC}(f) := c_\alpha$. Observe that $\mathrm{LT}(fg) = \mathrm{LT}(f)\mathrm{LT}(g)$.

Proposition 4.1.4. *Let $\{f_1, \ldots, f_m\}$ be a set of polynomials in $R = k[x_1, \ldots, x_n]$, ordered with $>$. For any polynomial $f \in R$, one has*

$$f = a_1f_1 + a_2f_2 + \cdots + a_mf_m + r$$

where $a_1, \ldots, a_m, r \in R$ are such that

- *for all $i = 1, \ldots, m$, $a_if_i = 0$ or $\mathrm{LM}(f) \geq \mathrm{LM}(a_if_i)$,*
- *$r = 0$ or no monomial in r is divisible by $\mathrm{LM}(f_1), \mathrm{LM}(f_2), \ldots,$ or $\mathrm{LM}(f_m)$.*

The polynomial r is called the remainder of the division of f by $\{f_1, \ldots, f_m\}$.

Proof. The proof of this proposition can be given under the form of an algorithm; see Algorithm 2. □

It turns out that the remainder of the multivariate polynomial division depends on the order of the family F of polynomials, and moreover, if $r = 0$ obviously implies that $f \in (f_1, \ldots, f_m)$, the contrary is not true. We illustrate these facts with the following example.

Example 4.1.5. Let $R = k[x, y]$, $>$ be the lexicographic order and consider the polynomials $f = xy^2 - x$, $f_1 = xy + 1$ and $f_2 = y^2 - 1$. If $F := \{f_1, f_2\}$ then we get $f = yf_1 + 0f_2 + (-x - y)$. However, if $F := \{f_2, f_1\}$ then we get $f = xf_2 + 0f_1 + 0$.

To fix the above issues, we need to characterize "good" sets of polynomials F.

Algorithm 2 Multivariate polynomial division

```
$a_1 := 0, a_2 := 0, \dots, a_m := 0$, $r := 0$, $p := f$
while $p \neq 0$ do
  $i := 1$, div := false
  while $i \leq m$ and div = false do
    if $\mathrm{LT}(f_i) | \mathrm{LT}(p)$ then
      $a_i := a_i + \mathrm{LT}(p)/\mathrm{LT}(f_i)$
      $p := p - (\mathrm{LT}(p)/\mathrm{LT}(f_i)) f_i$
      div := true
    else
      $i := i + 1$
    end if
  end while
  if div = false then
    $r := r + \mathrm{LT}(p)$ and $p := p - \mathrm{LT}(p)$
  end if
end while
```

4.2 Gröbner Bases

Definition 4.2.1. *A subset $\{g_1, \dots, g_m\}$ of an ideal I in $R = k[x_1, \dots, x_n]$ is a* Gröbner basis *of I if*

$$(\mathrm{LT}(g_1), \mathrm{LT}(g_2), \dots, \mathrm{LT}(g_m)) = (\mathrm{LT}(I)),$$

where $\mathrm{LT}(I) = \{cx^\alpha$ such that $\exists f \in I \ : \ \mathrm{LT}(f) = cx^\alpha\}$.

Corollary 4.2.2.

i) Every ideal has a Gröbner basis.

ii) If $\{g_1, \dots, g_m\}$ is a Gröbner basis of I then $(g_1, \dots, g_m) = I$.

Proof. i) follows from Noetherianity [1].

To prove *ii*), first notice that the inclusion $(g_1, \dots, g_m) \subset I$ is obvious. Now, let $f \in I$ and apply the division algorithm: $f = \sum_{i=1}^m a_i g_i + r$. If $r \neq 0$ then the monomials of r are not divisible by any $\mathrm{LT}(g_i)$. But $r = f - \sum_{i=1}^m a_i g_i \in I$, so $\mathrm{LT}(r) \in (\mathrm{LT}(I)) = (\mathrm{LT}(g_1), \dots, \mathrm{LT}(g_m))$; a contradiction. □

Proposition 4.2.3. *Let $G = \{g_1, \dots, g_m\}$ be a Gröbner basis of I and let $f \in R = k[x_1, \dots, x_n]$. Then, there exists a unique polynomial $r \in R$ such that*

- *any term in r is not divisible by any $\mathrm{LT}(g_i)$,*
- *there exists $g \in I$ such that $f = g + r$.*

In particular, r is the remainder of the division of f by the set of polynomials $\{g_1, \dots, g_m\}$, independently of its ordering; it is denoted by $r = \overline{f}^G$.

[1] It could also be proved via Dickon's lemma; see [MS20, CLO07]

Proof. The existence of r follows from multivariate division; see Proposition 4.1.4. Let $f = g + r = g' + r'$ be two such decompositions. We deduce that $r - r' = g - g' \in I$. If $r - r' \neq 0$, then $\mathrm{LT}(r - r') \in \mathrm{LT}(I) = (\mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_m))$, so there exists an integer i such that $\mathrm{LT}(g_i)$ divides $\mathrm{LT}(r - r')$, which is impossible. □

Corollary 4.2.4. *If $G = \{g_1, \ldots, g_m\}$ is a Gröbner basis of I, then $f \in I$ if and only if $\overline{f}^G = 0$.*

Exercise 4.2.5. Let I be an ideal in $R = k[x_1, \ldots, x_n]$, k an algebraically closed field, $f \in I$, and G be a Gröbner basis of the ideal $I + (1 - x_{n+1}f)$ of $R[x_{n+1}]$. Then, show that $f \in \sqrt{I}$ if and only if G contains a constant.

Exercise 4.2.6. Let I be an ideal in $R = k[x_1, \ldots, x_n]$, k an algebraically closed field, and G be a Gröbner basis of I. Then, show that $V(I) = \emptyset$ if and only if G contains a constant.

Syzygy pairs. So far we have seen that Gröbner basis provides a useful tool. Our next step is to show how Gröbner basis can be characterized and computed.

Definition 4.2.7. *Let f, g be nonzero polynomials in $R = k[x_1, \ldots, x_n]$. Set $\mathrm{LT}(f) = cx^\alpha$, $\mathrm{LT}(g) = dx^\beta$ and $\mathrm{LCM}(x^\alpha, x^\beta) = x^\gamma$ ($\gamma_i = \max(\alpha_i, \beta_i)$ for all i). Then, we define the* syzygy pair

$$S(f, g) := \frac{x^\gamma}{\mathrm{LT}(f)} f - \frac{x^\gamma}{\mathrm{LT}(g)} g.$$

Theorem 4.2.8. *Let $G = \{g_1, \ldots, g_m\}$ be a set of polynomials in $k[x_1, \ldots, x_n]$ and set $I := (g_1, \ldots, g_m)$. Then, G is a Gröbner basis of I if and only if $\overline{S(g_i, g_j)}^G = 0$ for all $i \neq j$.*

Proof. The syzygy pairs belong to I by assumption, so if G is a Gröbner basis of I, then all the syzygy pairs reduce to zero, which proves $\Rightarrow$. The proof of the other direction is rather technical, so we only provide the main lines and refer to [CLO07, Chapter 2, Theorem 6].

Given $f \in I$, one has to show that $\mathrm{LT}(f) \in (\mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_m))$. Since $f \in I$, we have $f = \sum_{i=1}^m a_i g_i$ and from here we have two cases:

- *Case 1:* $\mathrm{LT}(f) = \mathrm{LT}(a_k g_k)$ for some integer k, and then we are done.
- *Case 2:* There are some cancellations in the leading terms of the $a_i g_i$'s. Each cancellation corresponds to a syzygy pair, and since syzygy pairs reduce to zero, one can replace them by a new combination of the g_i's. Proceeding this way, the total degree of the polynomial $\sum_{i=1}^m a_i g_i$ decreases, which allows to prove the claimed result.

□

Example 4.2.9. Consider the ideal $I = (y - z^2, z - x^3)$ in $R = k[x, y, z]$ with the lexicographic order $y > z > x$. We have

$$S(y - x^2, z - x^3) = \frac{yz}{y}(y - x^2) - \frac{yz}{z}(z - x^3) = -zx^2 + yx^3,$$

and the division algorithm gives

$$-zx^2 + yx^3 = x^3(y - x^2) + (-x^2)(z - x^3) + 0.$$

Therefore, we deduce that $G = \{y - x^2, z - x^3\}$ is a Gröbner basis of I for the lexicographic order $y > z > x$.

Buchberger's algorithm. Using syzygy pairs, the Buchberger's algorithm allow us to compute a Gröbner basis of a given ideal in a polynomial ring.

Algorithm 3 Buchberger's algorithm

```
$I := (f_1, \dots, f_r)$ an ideal. Set $G := \{f_1, \dots, f_r\}$.
repeat
    $G' := G$
    for each pair $p, q \in G'$ do
        $S := \overline{S(p,q)}^{G'}$
        if $S \neq 0$ then
            $G := G \cup \{S\}$
        end if
    end for
until $G' = G$
```

Proof of Buchberger's algorithm. First, we notice that $I = (G)$ at any step because syzygy pairs belong to I by construction. Also, when the algorithm stops, G is a Gröbner basis by Theorem 4.2.8. It remains to show that the algorithm stops.

At each step, $G' \subseteq G$, hence $(\mathrm{LT}(G')) \subseteq (\mathrm{LT}(G))$. If $G' \subsetneq G$ then $(\mathrm{LT}(G')) \subsetneq (\mathrm{LT}(G))$ because a syzygy pair added is a polynomial r such that $\mathrm{LT}(r) \notin (\mathrm{LT}(G'))$. Thus, this gives an increasing chain of ideals that must stop at some point by Noetherianity. Therefore $(\mathrm{LT}(G')) = (\mathrm{LT}(G))$ at some point, hence $G' = G$. □

Example 4.2.10. We compute the Gröbner basis of the ideal $I = (f_1, f_2)$, $f_1 = x^2 - y^2$, $f_2 = xy - 1$, with respect to the lexicographic order $x > y$.

$$S(f_1, f_2) = y(x^2 - y^2) - x(xy - 1) = x - y^3 =: f_3$$

is not reduced by $\mathrm{LT}(f_1), \mathrm{LT}(f_2)$.

$$S(f_1, f_3) = 1(x^2 - y^2) - x(x - y^3) = xy^3 - y^2$$

reduces to zero with respect to f_1, f_2, f_3.

$$S(f_2, f_3) = 1(xy - 1) - y(x - y^3) = y^4 - 1 =: f_4$$

does not reduce. In the final pass, all syzygy pairs reduce to zero and we get the Gröbner basis

$$G := \{f_1 = x^2 - y^2, f_2 = xy - 1, f_3 = x - y^3, f_4 = y^4 - 1\}.$$

Observe that the polynomials f_1 and f_2 seem to be superfluous in this basis (because of leading terms; remember the definition of Gröbner bases).

Exercise 4.2.11. Let $f_1 = xy - x$, $f_2 = x^2 - y$ in $\mathbb{Q}[x, y]$ with the grlex ordering and $x > y$. Build a Gröbner basis of the ideal $I = (f_1, f_2)$.

Minimality and reduction. As observed in Example 4.2.10, Buchberger's algorithm may produce Gröbner basis that have some redundant elements.

Lemma 4.2.12. *Let G be a Gröbner basis of an ideal I and let $p \in G$ be such that $\mathrm{LT}(p) \in (\mathrm{LT}(G \setminus \{p\}))$. Then, $G \setminus \{p\}$ is also a Gröbner basis of I.*

Proof. By definition, $(\mathrm{LT}(G)) = (\mathrm{LT}(I))$ and by assumption, $(\mathrm{LT}(G\setminus\{p\})) = (\mathrm{LT}(G))$. □

Definition 4.2.13. *A Gröbner basis G of an ideal I is* minimal *if*

i) $\mathrm{LC}(p) = 1$ *for all* $p \in G$,
ii) for all $p \in G$, $\mathrm{LT}(p) \notin (\mathrm{LT}(G \setminus \{p\}))$.

Definition 4.2.14. *A Gröbner basis G of an ideal I is* reduced *if*

i) $\mathrm{LC}(p) = 1$ *for all* $p \in G$,
ii) for all $p \in G$, *no term in* p *belongs to* $(\mathrm{LT}(G \setminus \{p\}))$.

Proposition 4.2.15. *Any ideal in $k[x_1, \ldots, x_n]$ has a unique reduced Gröbner basis.*

Proof. See [CLO07, Chapter 2, §7, Proposition 6]. □

4.3 Monomial Ideals and Applications

We have seen that the Hilbert polynomial of a finitely generated R-module M can be computed by means of a finite free resolution. In what follows we provide a more efficient method in the case of a quotient ring $M = R/I$.

Lemma 4.3.1 (Macaulay). *Let I be an homogeneous ideal in the graded polynomial ring $R = k[x_1, \ldots, x_n]$, then*

$$\mathrm{HF}_{R/I}(t) = \mathrm{HF}_{R/(\mathrm{LT}(I))}(t)$$

for all $t \geq 0$.

Proof. Pick an integer j and let $f_1, \ldots, f_l$ be homogeneous polynomials of degree j which form of basis of I_j, i.e. $I_j = \langle f_1, \ldots, f_l \rangle_k$. Without loos of generality, one can assume $\mathrm{LM}(f_1) > \mathrm{LM}(f_2) > \cdots > \mathrm{LM}(f_l)$. It follows that $\langle \mathrm{LM}(f_1), \ldots, \mathrm{LM}(f_l) \rangle_k$ is of dimension l. Thus, to prove the claimed result it is enough to show that

$$\mathrm{LT}(I)_j = \langle \mathrm{LM}(f_1), \ldots, \mathrm{LM}(f_l) \rangle_k.$$

Suppose this is not true, i.e. there exists $f \in \mathrm{LT}(I)_j$ but $f \notin \langle \mathrm{LM}(f_1), \ldots, \mathrm{LM}(f_l)\rangle_k$. We choose f such that $m := \mathrm{LM}(f)$ is minimal with respect the monomial ordering. Necessarily, there exists $g \in I_j$ such that $\mathrm{LM}(g) = m$. Moreover, since $g \in \langle f_1, \ldots, f_l\rangle_k$, $g = \sum_{i=1}^{l} \alpha_i f_i$, $\alpha_i \in k$, we deduce that $\mathrm{LM}(g) = \mathrm{LM}(f_i)$ for some i. We deduce that there exists $\gamma \in k$ such that $f - \gamma\mathrm{LM}(f_i) \notin \langle \mathrm{LM}(f_1), \ldots, \mathrm{LM}(f_l)\rangle_k$ and $\mathrm{LM}(f - \gamma\mathrm{LM}(f_i)) < m = \mathrm{LM}(f)$, which gives a contradiction. □

As a consequence on the above result, to compute the Hilbert function of R/I one can compute first a Gröbner basis $G := \{g_1, \ldots, g_m\}$ of I and then consider the monomial ideal $(\mathrm{LT}(I)) = (\mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_m))$. Indeed, monomial ideals are nice for computations:

Lemma 4.3.2. *Let $I = (x^{\alpha_1}, x^{\alpha_2}, \ldots, x^{\alpha_l})$ be a monomial ideal and x^α be a monomial. Then,*

i) $x^\alpha \in I$ if and only if $x^{\alpha_i} | x^\alpha$ for some i.

ii) $f \in I$ if and only if f is a linear combination of monomials that belong to I.

Proof. One direction of *i)* is obvious. For the other one, observe that if $x^\alpha = \sum_j f_j x_j^\alpha$ then each term on the right side is a multiple of x^{α_j} for some j. The proof of *ii)* is similar: if $f = \sum_j f_j x_j^\alpha$, then all terms in the right side are multiples of the monomials x^{α_j}, hence belongs to I. □

As another illustration of the nice behavior of monomial ideals with respect to computations, we show that Hilbert polynomials of a monomial ideals can be computed inductively. To be more precise, let I be a monomial ideal (in a graded ring) and let $x^\alpha \notin I$; set $d := |\alpha|$. From the exact sequence

$$0 \to R(-d)/(I : x^\alpha) \xrightarrow{\times x^\alpha} R/I \to R/(I, x^\alpha) \to 0$$

we deduce that

$$\mathrm{HP}_{R/(I,x^\alpha)}(i) = \mathrm{HP}_{R/I}(i) - \mathrm{HP}_{R/(I:x^\alpha)}(i - d).$$

On the left side, one has the Hilbert polynomial of a monomial ideal with l generators whereas one the right side one has Hilbert polynomials of monomial ideals with $l - 1$ generators. Indeed, $(I : x^\alpha)$ is easily described from I as follows.

Lemma 4.3.3. *Let $I = (x^{\alpha_1}, x^{\alpha_2}, \ldots, x^{\alpha_l})$ a monomial ideal and x^α a monomial. Then,*

$$(I : x^\alpha) = \left(\frac{x^{\alpha_1}}{\mathrm{GCD}(x^{\alpha_1}, x^\alpha)}, \ldots, \frac{x^{\alpha_l}}{\mathrm{GCD}(x^{\alpha_l}, x^\alpha)}\right).$$

Proof. The inclusion $\supseteq$ is clear. For the other one: if $x^\alpha g \in I$, $g = \sum_\gamma c_\gamma x^\gamma$, then $x^\alpha x^\gamma \in I$ for all γ by Lemma 4.3.2. It follows that x^{α_i} divides $x^\alpha x^\gamma$ for some i, and hence that $\frac{x^{\alpha_i}}{\mathrm{GCD}(x^{\alpha_i}, x^\alpha)}$ divides $\frac{x^\alpha}{\mathrm{GCD}(x^{\alpha_i}, x^\alpha)} x^\gamma$. As $\frac{x^{\alpha_i}}{\mathrm{GCD}(x^{\alpha_i}, x^\alpha)}$ and $\frac{x^\alpha}{\mathrm{GCD}(x^{\alpha_i}, x^\alpha)}$ are coprime, this concludes the proof. □

4.4 Computing Syzygies of an Ideal

Suppose given an ideal $I = (f_1, \dots, f_m)$ in $R = k[x_1, \dots, x_n]$. For computing the first step of a finite free resolution of I it is necessary to compute the syzygies of I, that is to say

$$\mathrm{Syz}(I) := \{(h_1, \dots, h_m) \text{ such that } \sum_i h_i f_i = 0\} \subset R^m.$$

Indeed, by definition one has the exact sequence

$$0 \to \mathrm{Syz}(I) \to R^m \xrightarrow{(f_1,\dots,f_m)} R \to R/I \to 0.$$

Actually, what we need is a set of generators of the syzygy module $\mathrm{Syz}(I) \subset R^m$. It turns out that such a set can be extracted from syzygy pairs. Let us be more precise.

Let $G := \{g_1, \dots, g_s\}$ be a Gröbner basis of $I = (g_1, \dots, g_s)$. We recall that we have introduced syzygy pairs

$$S(g_i, g_j) = \frac{x^\gamma}{\mathrm{LT}(g_i)} g_i - \frac{x^\gamma}{\mathrm{LT}(g_j)} g_j,$$

where $x^\gamma = \mathrm{LCM}(\mathrm{LM}(g_i), \mathrm{LM}(g_j))$. Since G is a Gröbner basis, these syzygy pairs reduce to zero. Therefore, applying the division algorithm we get

$$S(g_i, g_j) = \sum_{l=1}^{s} a_{i,j,l} g_l$$

where $a_{i,j,l}$ are polynomials. Now, for all i, j define

$$a_{i,j} := a_{i,j,1} e_1 + \cdots + a_{i,j,s} e_s \in R^s = \oplus R e_i$$

and

$$s_{i,j} := \frac{x^\gamma}{\mathrm{LT}(g_i)} e_i - \frac{x^\gamma}{\mathrm{LT}(g_j)} e_j - a_{i,j}.$$

Theorem 4.4.1. *The set $\{s_{i,j},\ 1 \le i, j \le s\}$ form a set of generators of $\mathrm{Syz}(I)$ as a R-module.*

Proof. See [CLO07, Chapter 2, §9]. □

Therefore, a slight modification of Buchberger's algorithm gives the first step of a finite free resolution of a quotient ring R/I:

$$R^N \xrightarrow{\begin{pmatrix} \vdots & s_{i,j} & \vdots \end{pmatrix}} R^s \xrightarrow{(g_1,\dots,g_s)} R \to R/I \to 0.$$

In order to continue this process we need to compute sygyzies of a submodule of R^s, not only a submodule of R (i.e. an ideal). In the next section we briefly address this question.

4.5 Computing Finite Free Resolutions

First, we will quickly overview the theory of Gröbner basis for submodules M of R^m, where $R = k[x_1, \ldots, x_n]$. The typical questions we would like to solve are:

- Given $M \subset R^m$ and $f \in R^m$, decide if $f \in M$.
- Given $\{f_1, \ldots, f_s\} \subset R^m$, find a set of generators for $\mathrm{Syz}(f_1, \ldots, f_s) \subset R^s$.

The extension of Gröbner basis from ideals to modules requires that we (1) define monomial orders, (2) construct a division algorithm, (3) devise an algorithm similar to the Buchberger's algorithm.

Monomial Orders. A monomial $m \in R^s = \oplus_{i=1}^s Re_i$ is an element of the form $x^\alpha e_i$. Every $f \in R^s$ can be written as $\sum_{i=1}^l c_i m_i$, $c_i \in k$.

Example 4.5.1. In $R = k[x, y, z]$,

$$f = \begin{pmatrix} 5x \\ 2y \\ x+3z \end{pmatrix} = 5xe_1 + 2ye_2 + xe_3 + 3ze_3 \ \in R^3.$$

Let $m_1 = x^\alpha e_i$ and $m_2 = x^\beta e_j$. We say that m_2 divides m_1 if and only if $i = j$ and x^β divides x^α. In this case, we set $m_1/m_2 = x^\alpha/x^\beta \in R$. In the same vain, we set

$$\mathrm{GCD}(m_1, m_2) = \begin{cases} 0 & \text{if } i \neq j \\ \mathrm{GCD}(x^\alpha, x^\beta)e_i & \text{if } i = j \end{cases}$$

and

$$\mathrm{LCM}(m_1, m_2) = \begin{cases} 0 & \text{if } i \neq j \\ \mathrm{LCM}(x^\alpha, x^\beta)e_i & \text{if } i = j. \end{cases}$$

Definition 4.5.2. *An ordering relation $>$ on the monomials of R^s is a monomial ordering if*

- *$>$ is a total order,*
- *$m_1, m_2 \in R^s$, if $m > n$ then $x^\alpha m_1 > x^\alpha m_2$ for all x^α,*
- *$>$ is a well-ordering.*

This definition is very similar to the one in the case $s = 1$. Actually, on can get monomial orders on R^s by extending monomial orders on R as follows. First, choose an order on the entries in a column, say $e_1 > e_2 > \cdots > e_s$. Then, one has the two following monomial orders on R^s:

- TOP: $x^\alpha e_i > x^\beta e_j$ if $x^\alpha > x^\beta$, or if $x^\alpha = x^\beta$ and $i < j$.
- POT: $x^\alpha e_i > x^\beta e_j$ if $i < j$, or if $i = j$ and $x^\alpha > x^\beta$.

Now, given $f \in R^s$ we can write $f = \sum_{i=1}^l c_i m_i$ with $m_1 > m_2 > \cdots$ and $c_1 \neq 0$. Thus, we define $\mathrm{LT}(f) := c_1 m_1$ and $\mathrm{LM}(f) := m_1$.

Division Algorithm. With above definition of monomial orders in R^s, we obtain directly a division algorithm.

Let $\{f_1, \ldots, f_s\}$ be a s-tuple of elements in R^m. Then, any $f \in R^m$ can be written as

$$f = a_1 f_1 + \cdots + a_s f_s + r, \ a_i \in R, \ r \in R^m$$

where

- $\mathrm{LM}(f) \geq \mathrm{LM}(a_i f_i)$ for all i,
- $r = 0$ or r is a k-linear combination of monomials, none of which is divisible by any $\mathrm{LM}(f_i)$.

(The proof is the same as in the case $m = 1$).

Gröbner Basis. Let M be a submodule of R^m and $>$ a monomial order. Let $\langle \mathrm{LT}(M) \rangle$ be the monomial submodule generated by the leading terms of all $f \in M$.

Definition 4.5.3. *A finite set $\{g_1, \ldots, g_s\} \subset M \subset R^m$ is called. a Gröbner basis for M if $\langle \mathrm{LT}(M) \rangle = \langle \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_s) \rangle$.*

Proposition 4.5.4. *Let $G := \{g_1, \ldots, g_s\}$ be a Gröbner basis for a submodule $M \subset R^m$ and let $f \in R^m$. Then,*

i) G generates M as an R-module.

ii) $f \in M$ if and only if the remainder of the division by G is 0,

Finally, we mention that Gröbner bases do exist for all submodules of R^m and that minimal and reduced Gröbner basis can also be defined; see [CLO98, Chapter 5, §2] for the details.

Syzygy Pairs. Choosing a monomial order on R^m, given $f, g \in R^m$ we define their syzygy pair by

$$S(f, g) = \frac{m}{\mathrm{LT}(f)} f - \frac{m}{\mathrm{LT}(g)} g$$

where $m = \mathrm{LCM}(\mathrm{LT}(f), \mathrm{LT}(g))$.

Example 4.5.5. Let $R = k[x, y]$ equipped with the POT extension of the lexicographic order. If

$$f = \begin{pmatrix} xy - x \\ x^3 + y \end{pmatrix} \text{ and } f = \begin{pmatrix} x^2 + 2y^2 \\ x^2 - y^2 \end{pmatrix},$$

Then, $\mathrm{LT}(f) = xye_1$, $\mathrm{LT}(g) = x^2 e_1$, $\mathrm{LCM}(\mathrm{LT}(f), \mathrm{LT}(g)) = x^2 y e_1$ and

$$S(f, g) = xf - yg = \begin{pmatrix} -x^2 - 2y^3 \\ x^4 - x^2 y + xy + y^3 \end{pmatrix}.$$

Buchberger's Criterion.

Theorem 4.5.6. *A set $G := \{g_1, \ldots, g_s\} \subset R^m$ is a Gröbner basis for the module M it generates if and only if all syzygy pairs reduce to zero.*

The proof of the above theorem is essentially the same as in the case of Gröbner bases of ideals. Moreover, one can derive the same Buchberger algorithm to build a Gröbner basis.

Computing Finite Free Resolutions. We proceed as in Section 4.4 to compute syzygies of a submodule $M \subset R^m$. Let $G := \{g_1, \dots, g_s\}$ be a Gröbner basis of M. The syzygy pairs reduce to zero and hence

$$S(g_i, g_j) = \sum_k a_{i,j,k} g_k,$$

which, as explained in Section 4.4, yields a syzygy that we denote by $s_{i,j}$.

Theorem 4.5.7. *Let $G := \{g_1, \dots, g_s\}$ be a Gröbner basis. The set of syzygies $s_{i,j}$'s form a Gröbner basis for the syzygy module* $\operatorname{Syz}(g_1, \dots, g_s)$.

Proof. See [CLO98, Chapter 5]. □

From there, one can compute a finite free resolution iteratively (up to change of basis matrices if one wants to keep a specific list of generators instead of a Gröbner basis; see [CLO98, Chapter 5] for more details).

Chapter 5

Projection and Elimination

5.1 Elimination Ideal

An important application of Gröbner basis theory is solving systems of polynomial equations. The main idea is to project down to a lower dimensional space, solve and then lift the solutions. In some sense, this is very similar to Gaussian elimination for systems of linear equations.

Definition 5.1.1. *Let $I \subset R = k[x_1, \ldots, x_n]$ be an ideal. The m^{th} elimination ideal of I is the ideal*

$$_mI := I \cap k[x_{m+1}, \ldots, x_n].$$

Geometrically, consider the projection map

$$\begin{aligned} \pi_m : \mathbb{A}^n &\to \mathbb{A}^{n-m} \\ (a_1, \ldots, a_n) &\mapsto (a_{m+1}, \ldots, a_n). \end{aligned}$$

It turns out that the projection of a variety is not necessarily a variety (give an example!), but taking algebraic closure one has the following result.

Theorem 5.1.2. *Let $I \subset R = k[x_1, \ldots, x_n]$ be an ideal and assume that k is an algebraically closed field, then*

$$\overline{\pi_m(V(I))} = V({}_mI).$$

Moreover, if I is radical or prime then the elimination ideal $_mI$ has the same property.

Proof. We begin with the inclusion $\subseteq$. Let $(a_{m+1}, \ldots, a_n) \in \pi_m(V(I))$. By construction, there exists $(a_1, \ldots, a_n) \in V(I)$ such that $\pi_m(a_1, \ldots, a_n) = (a_{m+1}, \ldots, a_n)$. Let $f \in {}_mI$. Since $f \in I$, $f(a_1, \ldots, a_n) = 0$. But $f \in k[x_{m+1}, \ldots, x_n]$, so $f(a_{m+1}, \ldots, a_n) = 0$, i.e. f vanishes on $\pi_m(V(I))$.

Next we show that $I(\pi_m(V(I))) \subseteq I(V({}_mI))$, which concludes the proof by passing to varieties. Let $g \in I(\pi_m(V(I))) \subset k[x_{m+1}, \ldots, x_n]$. Seen as a polynomial in R, g vanishes on $V(I)$, so $g^p \in I$ for some integer p. But we also have $g^p \in k[x_{m+1}, \ldots, x_n]$, so $g^p \in {}_mI$, i.e. $g \in \sqrt{{}_mI} = I(V({}_mI))$. □

The above theorem shows that the algebraic operation of elimination corresponds to the geometric operation of projection. These operations are fundamental in many applications. Here is an example.

Example 5.1.3 (Matrix completion)**.** Consider the variety V of symmetric matrices $M = (x_{i,j})$ of size 5×5 and rank ≤ 2. It is an irreducible variety: the ideal of 3-minors of M is a prime ideal minimally generated by 50 homogeneous polynomials in the $x_{i,j}$'s. The elimination of the diagonal entries $x_{i,i}$, $i = 1, \ldots, 5$, leads to a principal ideal:

$$\begin{aligned} J = (&x_{12}x_{13}x_{24}x_{35}x_{45} - x_{12}x_{13}x_{25}x_{34}x_{45} - x_{12}x_{14}x_{23}x_{35}x_{45} + x_{12}x_{14}x_{25}x_{34}x_{35} + \\ &x_{12}x_{15}x_{23}x_{34}x_{45} - x_{12}x_{15}x_{24}x_{34}x_{35} + x_{13}x_{14}x_{23}x_{25}x_{45} - x_{13}x_{14}x_{24}x_{25}x_{35} - \\ &x_{13}x_{15}x_{23}x_{24}x_{45} + x_{13}x_{15}x_{24}x_{25}x_{34} + x_{14}x_{15}x_{23}x_{24}x_{35} - x_{14}x_{15}x_{23}x_{25}x_{34}). \end{aligned}$$

Therefore, given the ten entries which are not on the diagonal, a 5×5 symmetric matrix of rank ≤ 2 can be completed providing the above polynomial constraint is satisfied. Matrix completion appears in many applied fields, including algebraic statistics; see [MS20, Exercise 4.3] for more details.

How the generator of the principal ideal in the above example can be computed? It turns out that Gröbner basis can be used to compute elimination ideals.

Theorem 5.1.4. *Let $I \subset R = k[x_1, \ldots, x_n]$ be an ideal and let $G := \{g_1, \ldots, g_s\}$ be a Gröbner basis for I with respect to the lexicographic order $x_1 > x_2 > \cdots > x_n$. Then,*

$$_mG := G \cap k[x_{m+1}, \ldots, x_n]$$

is a Gröbner basis for $_mI$. In addition, if G is a reduced Gröbner basis for I, then $_mG$ is also a reduced Gröbner basis for $_mI$.

Proof. It is clear that $(_mG) \subseteq {}_mI$.

Pick $f \in {}_mI$. By the division algorithm we can write $f = \sum_{i=1}^{s} h_i g_i$. Since the monomial order is the lex order, then all the g_i's that appear in this equality must be in $_mI$. We deduce that $_mI = (_mG)$. It remains to show that syzygy pairs reduce to zero, but this is automatic as G is a Gröbner basis. □

Exercise 5.1.5. We would like to compute the extrema of the real-valued function

$$f(x, y, z) = x^3 + 2xyz - z^2$$

restricted to the unit sphere, i.e. under the constraint $h(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$. The Lagrange multiplier method suggests to form the polynomial system corresponding to the partial derivatives of the polynomial $f + \lambda h$. Explain how you could put this system in a triangular structure, ready for solving, and provide a bound for the number of extrema.

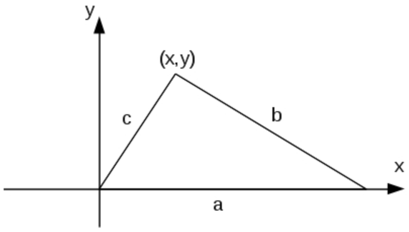


Figure 5.1: Héron formula

Exercise 5.1.6. Recover the Héron formula that allows to compute the area s of a planar triangle from the lengths a, b, c of its edges, namely

$$s^2 = \frac{1}{16}(a+b+c)(a+b-c)(a-b+c)(-a+b+c).$$

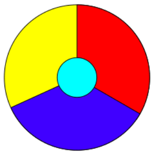
1. Using the notation in Figure 5.1, show that we have the equations:
$$b^2 = (a-x)^2 + y^2,\ c^2 = x^2 + y^2,\ 2s = ay.$$
2. Deduce the expected formula by polynomial elimination techniques.

Exercise 5.1.7.
The famous “Four Color Theorem” shows that only four colors are needed to color planar map so that no bordering regions have the same color. Typical examples are a colored world map, a colored map of the states of the USA, or a colored map of the French regions. In this exercise, we will provide a method to determine if three colors are sufficient for a particular map.

1. Could you provide a simple planar map to illustrate that three colors are not always enough to color it so that no bordering regions have the same color?
2. The three colors are represented by a complex cubic root of the unit and each region is represented by a variable x_i. Justify that for each region we have the polynomial equation
$$x_i^3 - 1 = 0.$$
3. Let x_j and x_k be two neighboring regions. As neighboring regions cannot have the same color, show that x_j and x_k must satisfy a polynomial equation of degree 2.
(Hint: use that $x_j^3 - x_k^3 = 0$).
4. Deduce from the previous questions that there exists a polynomial system such that a map with n regions can be colored with three colors if and only if there exists at least one solution to this polynomial system.
5. Given a particular map, explain how you would use a computer algebra system to determine if it can be colored with three colors.

Solution.

1. It is easy tp design small maps that cannot be colored with three colors. An example (from wikipedia):



2. The equation $x_i^3 = 1$ has three distinct complex roots $\{1, j, j^2\}$ that can be bijectively associated to three colors.
3. If x_j and x_k are two neighboring regions then the variables x_j and x_k are not allowed to take the same value. Since

$$0 = x_j^3 - x_k^3 = (x_j - x_k)(x_j^2 + x_j x_k + x_k^2)$$

   we deduce that the polynomial $x_j^2 + x_j x_k + x_k^2$ vanishes if and only if $x_j \neq x_k$, always under the assumption that $x_j^3 = 1$ and $x_k^3 = 1$.
4. For all i we set $f_i = x_i^3 - 1$ and for all couple (j, k) we set $g_{j,k} = x_j^2 + x_j x_k + x_k^2$. These are polynomials in $\mathbb{C}[x_1, \ldots, x_n]$. Consider the algebraic affine variety V defined by all the f_i's and the $g_{j,k}$'s such that x_j and x_k are neighboring regions. We deduce that the map can be colored with three colors, so that no bordering regions have the same color, if and only if $V \neq \emptyset$.
5. To conclude, by Hilbert Nullstellensatz we have to decide if $1 \in I$, where I is the ideal generated by the equations defining V. This can be done by computing a Gröbner basis of I, a task that can/must be done with a computer algebra system.

5.2 First Applications of Elimination

In this section, we provide some first applications of elimination. We begin with some basic operations on ideals that we encountered previously, then we discuss the computation of the image of a polynomial map.

Intersection of ideals. Let I, J be two ideals of $R = k[x_1, \ldots, x_n]$. Introduce a new indeterminate t and consider the ideal $L = t \times I + (1 - t) \times J$ of $R[t]$ (where the products are formed by multiplying generators by t and $(1 - t)$ respectively). Then, $I \cap J$ can be computed as an elimination ideal: $I \cap J = L \cap R$.

Indeed, if $f \in I \cap J$, then $f = tf + (1 - t)f$ in $R[t]$, so $f \in L \cap R$. Conversely, if $f \in L \cap R$, then there exists $g \in I$ and $h \in J$ such that $f = tg + (1 - t)h$. As $f \in R$, one can evaluate this equality at $t = 0$ and $t = 1$ and deduce that necessarily $f = g = h$, and hence that $f \in I \cap J$.

Ideal Quotient. Let I, J be two ideals of $R = k[x_1, \ldots, x_n]$. Let us write $J = (f_1, \ldots, f_m)$ and let $f \in k[x_1, \ldots, x_n]$. Then, it is immediate to verify that

- $(I : J) = \cap_{i=1}^m (I : (f_i))$,

- $(I : (f)) = (I \cap (f))f^{-1}$, which means the ideal of the elements in $I \cap (f)$, divided by f.

It follows that the computation of quotient ideals relies on the computation of intersection of ideals, an operation that can performed by means of Gröbner basis as we have just noticed above.

Saturation. Let I, J be two ideals of $R = k[x_1, \ldots, x_n]$, then the saturation of I by J is defined as the ideal $(I : J^\infty) = \cup_{i \in \mathbb{N}}(I : J^i)$. Observe that since the ideals $(I : J^i)$ form an ascending chain of ideals then there exists an integer s such that $(I : J^\infty) = (I : J^s)$, as R is a Noetherian ring.

Now, to compute $(I : J^\infty)$ we can proceed as follows. First, if we write $J = (f_1, \ldots, f_m)$, then $(I : J^\infty) = \cap_{i=1}^n (I : (f_i)^\infty)$. Then, suppose f is one of the f_i's and consider the ideal $I_t := I + (tf - 1) \subseteq R[t]$, then $(I : (f)^\infty) = I_t \cap R$.

Indeed, let $g \in I_t \cap R$. Then $g = g_1 p + g_2(tf - 1)$, with $g_1, g_2 \in R[t]$ and $p \in I$. Substituting t by f^{-1} and clearing denominators, we deduce that $f^s g \in I$ for some integer s, i.e. $g \in (I : (f)^\infty)$. Conversely, assume $f^s g \in I$ for some integer s. As $tf - 1 \in I_t$, $1 = tf + q$ with $q \in I_t$ and hence $1 = (tf)^s + q'$ with $q' \in I_t$. It follows that $g = t^s g f^s + g q' \in I_t$.

Image of a polynomial map. Computing the image of a polynomial map is called *implicitization*. It is a special instance of elimination as it can be done by forming the graph of the map and then projecting onto the image coordinates. To be more precise, consider a map of the form

$$\begin{aligned} \phi : \mathbb{A}^m &\to \mathbb{A}^n \\ x = (x_1, \ldots, x_m) &\mapsto (f_1(x), \ldots, f_n(x)) \end{aligned} \tag{5.2.1}$$

when $f_1, \ldots, f_n$ are polynomials in $R = k[x_1, \ldots, x_m]$, and k is an algebraically closed field. We write $\operatorname{Im}(\phi) \subset \mathbb{A}^n$ as the set-theoretic image of ϕ. In general, this set is not a variety.

Example 5.2.1. Assume $m = 2, n = 3$ and $f_1 = x_1$, $f_2 = x_1 x_2$ and $f_3 = x_1 x_2^2$. The Zarisky closure of the image is the surface of equation $y_1 y_3 - y_2^2 = 0$. The point $(0, 0, 1)$ is in this surface, but not in $\operatorname{Im}(\phi)$ (notice that for all $z \neq 0$, $\phi(z^2, 1/z) = (z^2, z, 1) \in \operatorname{Im}(\phi)$).

The closed image of the map ϕ is the Zarisky closure of $\operatorname{Im}(\phi)$, which is denoted by $\overline{\operatorname{Im}(\phi)}$. It is a subvariety in $\mathbb{A}^n$.

Corollary 5.2.2. *Given the map ϕ in (5.2.1), let I be the ideal in the polynomial ring $k[x_1, \ldots, x_m, y_1, \ldots, y_n]$ in $n+m$ variables which is generated by the polynomials $f_i(x_1, \ldots, x_m) - y_i$ for $i = 1, \ldots, n$. Then, the closed image of ϕ is the variety defined by the elimination ideal $J = I \cap k[y_1, \ldots, y_n]$; in other words, $\overline{\operatorname{Im}(\phi)} = V(J)$.*

More generally, if $X = V(I_X) \subset \mathbb{A}^m$ is a variety, then its closed image $\overline{\phi(X)}$ via ϕ is the variety defined by the elimination ideal $J = (I + I_X) \cap k[y_1, \ldots, y_n]$.

Proof. The graph of ϕ restricted to X is closed in $\mathbb{A}^m \times \mathbb{A}^n$ and $I + I_X$ is the ideal that defines it. The image of X is the projection of the graph onto $\mathbb{A}^n$. Therefore, the claim follows from Theorem 5.1.2. □

Exercise 5.2.3. Consider the twisted cubic curve in $\mathbb{R}^3$; it can be obtained as the image of the parameterization

$$\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}^3 \\ t & \mapsto & (t, t^2, t^3). \end{array}$$

1. Using a new parameter u, compute parameterizations of the tangent line to the twisted cubic.
2. Provide a parameterization of the surface obtained as the union of all the tangent line to the twisted cubic.
3. Compute the smallest algebraic set that contains this tangent surface.

5.3 The Sylvester Resultant

The most basic setting in elimination arises when one needs to eliminate n variables from $n+1$ equations. In this case, one expect the result to be a single equation in the coefficients of the system. Such equations, that can be seen as generalization of the determinant of a linear system of $n+1$ equations in n variables, are called *resultants*. These objects dated back to the 19th century and since them their theory have been widely developed.

Example 5.3.1 (Hyperdeterminant)**.** A tensor of format $2 \times 2 \times 2$ has 8 entries, so it can be viewed as an element in a 8-dimensional linear space, with basis indexed by three integers $0 \leq i, j, k \leq 1$ (the vertices of a cube). To such a tensor one can associate an affine trilinear form (each vertex i, j, k of the cube corresponds to the monomial $x_1^i x_2^j x_3^k$):

$$f = x_1\, x_2\, x_3\, y_{111} + x_1\, x_2\, y_{110} + x_1\, x_3\, y_{101} + x_1\, y_{100} + x_2\, x_3\, y_{011} + x_2\, y_{010} + x_3\, y_{001} + y_{000}.$$

Given specific values to the tensor coefficients, the polynomial f defines a surface in $\mathbb{A}^3$. It turns out that this surface has a singular point if and only if there is a point in $\mathbb{A}^3$ where f and its three partial derivatives vanish simultaneously. Thus, let I be the ideal generated by f, $\partial f/\partial x_1$, $\partial f/\partial x_2$ and $\partial f/\partial x_3$. To obtain the condition on the coefficients of the tensor so that the surface fails to be smooth we have to compute the elimination ideal $I \cap k[y_{i,j,k}]$; it turns out that this ideal is principal, generated by a quartic polynomial. It is called the *hyperdeterminant* of a $2 \times 2 \times 2$-tensor. See [MS20, Example 4.10] for more details and references.

In what follows we will discuss the first case of resultant theory, i.e. the case of two univariate polynomials. We will come back to the multivariate case in the next chapter.

Definition and main properties. Let A be a commutative ring and consider the two univariate polynomials

$$\left\{\begin{array}{lcl} f(x) & := & a_0x^m + a_1x^{m-1} + \cdots + a_m \\ g(x) & := & b_0x^n + b_1x^{n-1} + \cdots + b_n \end{array}\right. \tag{5.3.1}$$

of degree m and n in $A[x]$. The Sylvester matrix of f and g is defined as

$$\mathrm{Sylv}_{m,n}(f,g) := \begin{pmatrix} a_m & 0 & \cdots & 0 & b_n & 0 & 0 \\ a_{m-1} & a_m & & \vdots & b_{n-1} & \ddots & 0 \\ \vdots & & \ddots & 0 & \vdots & & b_n \\ a_0 & & & a_m & b_1 & & b_{n-1} \\ 0 & a_0 & & a_{m-1} & b_0 & & \vdots \\ \vdots & & \ddots & \vdots & 0 & \ddots & b_1 \\ 0 & \cdots & 0 & a_0 & 0 & 0 & b_0 \end{pmatrix}.$$

It is a square matrix of size $(m+n)$

Definition 5.3.2. *We define the* resultant *of f and g in degree (m,n) as the determinant of the Sylvester matrix* $\mathrm{Sylv}_{m,n}(f,g)$. *It is denoted by* $\mathrm{Res}_{m,n}(f,g)$.

Remark 5.3.3. We emphasize the degrees (m,n) in the definition of the resultant because we are considering an affine setting, so they are important. For instance, if $\deg(f) = m$ and $n \geq \deg(g)$ then

$$\mathrm{Res}_{m,n}(f,g) = a_0^{n-\deg(g)}\mathrm{Res}_{m,n-\deg(g)}(f,g).$$

Example 5.3.4. If $f := ax^2 + bx + c$ and $g = \partial_x f = 2ax + b$ then

$$\mathrm{Res}_{2,1}(f,g) = \begin{vmatrix} c & b & 0 \\ b & 2a & b \\ a & 0 & 2a \end{vmatrix} = a(b^2 - 4ac).$$

Do you recognize a classical quantity attached to $f(x)$?

Exercise 5.3.5. If $f = a_0x^m + \cdots + a_m$ and $g = x - b$ then show that

$$\mathrm{Res}_{m,1}(f,g) = \begin{vmatrix} a_m & -b & 0 & \cdots & 0 \\ a_{m-1} & 1 & -b & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & 1 & -b \\ a_0 & 0 & \cdots & 0 & 1 \end{vmatrix} = f(b).$$

Exercise 5.3.6. Show that $\mathrm{Res}_{m,n}(f,g) = (-1)^{mn}\mathrm{Res}_{n,m}(g,f)$.

Observe that by definition, we have the equality

$$\mathrm{Sylv}_{m,n}(f,g)^T \begin{pmatrix} 1 \\ x \\ \vdots \\ x^{m+n-2} \\ x^{m+n-1} \end{pmatrix} = \begin{pmatrix} f \\ xf \\ \vdots \\ x^{n-1}f \\ g \\ xg \\ \vdots \\ x^{m-1}g \end{pmatrix} \tag{5.3.2}$$

in $A[x]$, where $(-)^T$ stands for the transpose matrix. In particular, applying Cramer's rules we deduce that $\mathrm{Res}_{m,n}(f,g) = pf + qg$ with $p \in A[x]_{<n}$ and $q \in A[x]_{<m}$, where $A[x]_{<d}$ denotes the set of polynomials of degree $< d$.

Another interpretation of the Sylvester matrix is as follows. The polynomials f and g define a map of free $A[x]$-modules

$$A[x] \oplus A[x] \to A[x] : (u,v) \mapsto uf + vg$$

that induces another map of free A-modules by restriction,

$$\phi : A[x]_{<n} \times A[x]_{<m} \to A[x]_{<m+n} : (u,v) \mapsto uf + vg.$$

The Sylvester matrix of f and g is nothing but the matrix of ϕ in canonical bases. In particular, if A is a domain then ϕ is injective if and only if $\mathrm{Res}_{m,n}(f,g) \neq 0$ (this is a well-know property of linear algebra over a field, but it also holds over a domain - hint: consider the fraction field and clear denominators).

Proposition 5.3.7. *Assume that A is a domain and let $K = \mathrm{Frac}(A)$ be its fraction field. Let f and g be two polynomials in $A[x]$ defined by (5.3.1) such that $a_0 \neq 0$. Then, $\mathrm{Res}_{m,n}(f,g) \neq 0$ if and only if $f(x)$ and $g(x)$ are relatively prime polynomials in $K[x]$. In particular, $\mathrm{Res}_{m,n}(f,g) \neq 0$ if and only if f and g have no common root in the algebraic closure of K.*

Proof. Assume f and g are relatively prime and let (u,v) be an element in the kernel of ϕ. From $uf + vg = 0$ we deduce that v divides f, which is of degree m by assumption. As $v \in A[x]_{<m}$, then $v = 0$ and hence $u = 0$. Now, if f and g have a common factor h of positive degree in $K[x]$. Then, $f = h\tilde{f}$ and $g = h\tilde{g}$. It follows that, up to multiplication by an element in A, $(-\tilde{g}, \tilde{f})$ is a nonzero in element in the kernel ϕ. □

Corollary 5.3.8. *Assume A is a domain and let f and g be two polynomials in $A[x]$ defined by (5.3.1). Then, $\mathrm{Res}_{m,n}(f,g) = 0$ if and only if f and g have a common root in the algebraic closure of $K = \mathrm{Frac}(A)$ or $a_0 = b_0 = 0$.*

Exercise 5.3.9. Explain how $\mathrm{Res}_{m,n}(f,g)$ can be computed by means of Gröbner basis.

Some formal properties of the resultant. The resultant has many nice properties and in what follows we give some of them. As a first observation, it is important to notice that the resultant is *universal*, in the sense that it commutes with the specialization of the coefficients of the polynomials f and g. This is inherited from the definition as a determinant, since the determinant itself has this property (the computation of a determinant commutes with the specialization of its entries).

Let A be a commutative ring and f, g be two polynomials defined by (5.3.1).

Lemma 5.3.10 (Homogeneity)**.** *For any $a \in A$, $\mathrm{Res}_{m,n}(af,g) = a^n \mathrm{Res}_{m,n}(f,g)$ and $\mathrm{Res}_{m,n}(f,ag) = a^m \mathrm{Res}_{m,n}(f,g)$.*

Proof. This is immediate from the definition of the resultant as the determinant of the Sylvester matrix. □

Assuming that A is the universal ring of coefficients of the polynomials f and g, i.e. that $A := \mathbb{Z}[a_0, \ldots, a_n, b_0, \ldots, b_m]$, the above lemma means that $\mathrm{Res}_{m,n}(f,g)$ is a homogeneous polynomial of degree n in the variables $a_0, \ldots, a_m$ and of degree m in the variables $b_0, \ldots, b_n$.

Proposition 5.3.11 (Poisson's formula)**.** *Suppose that a_0 is an invertible element in A and consider the multiplication map by g in the quotient ring $A[x]/(f)$:*

$$\psi : A[x]/(f) \to A[x]/(f) : \overline{u} \mapsto \overline{ug}.$$

Then, the determinant of the matrix ψ is equal to $a_0^{-n}\mathrm{Res}_{m,n}(f,g)$.

Proof. First, recall that since a_0 is invertible, $A[x]/(f)$ is a free A-module with basis $\{\overline{x}^{m-1}, \ldots, \overline{1}\}$ by Euclidian division: any polynomial $u(x) \in A[x]$ can be uniquely written as $u(x) = q(x)f(x) + r(x)$ with $\deg(r(x)) < m$, and we have $\overline{u} = r(\overline{x})$.

Now, consider the two A-module morphisms

$$\phi : A[x]_{<n} \times A[x]_{<m} \to A[x]_{<m+n} : (u,v) \mapsto uf + vg$$

and

$$\theta : A[x]_{<m+n} \to A[x]_{<n} \times A[x]_{<m} : p \mapsto (q,r)$$

where (q,r) correspond to the quotient and remainder of the Euclician division of p by f, respectively, i.e. $p = qf + r$. In canonical bases, the matrices M_ϕ, M_θ and $M_{\theta\circ\phi}$ of the maps ϕ, θ and $\theta \circ \phi$ respectively, satisfy to

$$\det(M_\phi)\det(M_\theta) = \det(M_{\theta\circ\phi}). \tag{5.3.3}$$

Since $M_\phi = \mathrm{Sylv}_{m,n}(f,g)$, we deduce $\det(M_\phi) = \mathrm{Res}_{m,n}(f,g)$. Moreover, the matrices M_θ et $M_{\theta\circ\phi}$ are of the form

$$M_\theta = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & * & * & * \\ 0 & \ddots & 0 & * & * & * \\ 0 & 0 & 1 & * & * & * \\ \hline & & & a_0^{-1} & * & * \\ & 0 & & 0 & \ddots & * \\ & & & 0 & 0 & a_0^{-1} \end{array}\right) \quad \text{and} \quad M_{\theta\circ\phi} = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & * & * & * \\ 0 & \ddots & 0 & * & * & * \\ 0 & 0 & 1 & * & * & * \\ \hline 0 & \dots & 0 & & & \\ \vdots & 0 & \vdots & & M_\psi & \\ 0 & \dots & 0 & & & \end{array}\right).$$

Therefore, (5.3.3) yields the claimed formula: $a_0^{-n}\mathrm{Res}_{m,n}(f,g) = \det(M_\psi)$. □

Proposition 5.3.12 (Multiplicativity)**.** *Let* $f(x) = a_0x^n + \cdots + a_n \in A[x]$ *and suppose given two polynomials* $g_1(x)$ *and* $g_2(x)$ *in* $A[x]$ *such that* $\deg(g_1) \leq n_1$ *and* $\deg(g_2) \leq n_2$*. Then,*

$$\mathrm{Res}_{m,n_1+n_2}(f, g_1g_2) = \mathrm{Res}_{m,n_1}(f, g_1)\mathrm{Res}_{m,n_2}(f, g_2).$$

Proof. By specialization of the coefficients of f, g_1 and g_2, it is enough to prove this property in the universal setting, i.e. assuming that

$$A := \mathbb{Z}[\mathrm{coeff}(f), \mathrm{coeff}(g_1), \mathrm{coeff}(g_2)].$$

In the ring $A_{a_0}[x]$ (localization ring where a_0 becomes an invertible element; recall that there is a canonical map $A \to A_{a_0} : a \mapsto a/1$ which is here injective as A has no torsion), the following diagram is commutative:

$$\begin{tikzcd} A_{a_0}[x]/(f) \arrow[rr, "\times g_1g_2"] \arrow[dr, "\times g_1"] & & A_{a_0}[x]/(f) \\ & A_{a_0}[x]/(f) \arrow[ur, "\times g_2"] & \end{tikzcd}$$

Applying Proposition 5.3.11, we deduce that

$$a_0^{-n_1-n_2}\mathrm{Res}_{m,n_1+n_2}(f, g_1g_2) = a_0^{-n_1}\mathrm{Res}_{m,n_1}(f, g_1)a_0^{-n_2}\mathrm{Res}_{m,n_2}(f, g_2).$$

The element a_0 is not a zero divisor in A, so the previous equality becomes an equality in A after simplification by a_0, which concludes the proof. □

Lemma 5.3.13 (Elementary transformations)**.** *If* $n \geq m$ *(resp.* $m \geq n$*), then for any polynomial* $h \in A[x]_{\leq n-m}$ *(resp.* $h \in A[x]_{\leq m-n}$*), the following equality holds in* A*:*

$$\mathrm{Res}_{m,n}(f, g + hf) = \mathrm{Res}_{m,n}(f,g) \quad (\text{resp. } \mathrm{Res}_{m,n}(f + hg, g) = \mathrm{Res}_{m,n}(f,g)).$$

Proof. Exercise (use the Sylvester matrix). □

Proposition 5.3.14 (Expressions in the roots)**.** *Assume that f and g are split in A, that is*

$$f(x) := a_0 \prod_{i=1}^{m}(x-\alpha_i) \quad \textit{and} \quad g(x) := b_0 \prod_{i=1}^{n}(x-\beta_i).$$

Then, the following equalities hold in A:

$$\mathrm{Res}_{m,n}(f,g) = (-1)^{mn} a_0^n b_0^m \prod_{\substack{1\le i\le m\\1\le j\le n}} (\alpha_i - \beta_j) = (-1)^{mn} a_0^n \prod_{i=1}^{m} g(\alpha_i) = b_0^m \prod_{i=1}^{n} f(\beta_i).$$

Proof. The first and third formulas are obtained as follows:

$$\begin{aligned}
\mathrm{Res}_{m,n}(f,g) &= \mathrm{Res}_{m,n}(f, b_0 \prod_{i=1}^{n}(X-\beta_i)) \\
&= b_0^m \mathrm{Res}_{m,n}(f, \prod_{i=1}^{n}(X-\beta_i)) && \text{by Lemma 5.3.10} \\
&= b_0^m \prod_{i=1}^{n} \mathrm{Res}_{m,n}(f, X-\beta_i) && \text{by Proposition 5.3.12} \\
&= b_0^m \prod_{i=1}^{n} f(\beta_i) && \text{par l'exemple 5.3.5} \\
&= (-1)^{mn} a_0^n b_0^m \prod_{j=1}^{n}\prod_{i=1}^{m}(\alpha_i-\beta_j).
\end{aligned}$$

A similar computation, by swapping the roles of f and g, yields the last formula. □

Proposition 5.3.15 (Twisted homogeneity)**.** *Assume that $A = \mathbb{Z}[a_0, \dots, a_m, b_0, \dots, b_n]$ and use the following weights:*

$$\begin{cases} \deg(p) = 0 & \textit{for all } p \in \mathbb{Z}, \\ \deg(a_i) = i \quad (\textit{resp. } m-i) & \textit{for all } i = 0, \dots, m, \\ \deg(b_j) = j \quad (\textit{resp. } n-j) & \textit{for all } j = 0, \dots, n. \end{cases}$$

Then, $\mathrm{Res}_{m,n}(f,g) \in A$ is homogeneous of degree mn.

Proof. It is a consequence of Proposition 5.3.14: $\mathrm{Res}_{m,n}(f,g)$ is homogeneous of degree mn in the roots of $f(x)$ and $g(x)$ (after choosing a suitable ring A), and the coefficients a_i and b_j are themselves homogeneous of degree i and j, respectively, with respect to those roots. □

It is worth mentioning that the homogeneity and twisted homogeneity properties of the resultant imply that

$$\mathrm{Res}_{m,n}(f,g) = \sum_{\substack{i_0+i_1+\cdots+i_m=n\\ j_0+j_1+\cdots+j_n=m\\ i_1+2i_2+\cdots+mi_m+j_1+2j_2+\cdots+nj_n=mn}} c_{i_0,i_1,\dots,i_m,j_0,\dots,j_n} a_0^{i_0} a_1^{i_1} \dots a_m^{i_m} b_0^{j_0} b_1^{j_1} \dots b_n^{j_n}$$

where $c_{i_0,i_1,\ldots,i_m,j_0,\ldots,j_n} \in \mathbb{Z}$ for all multi-indices $(i_0, i_1, \ldots, i_m, j_0, \ldots, j_n) \in \mathbb{N}^{m+n+2}$ (observe that the twisted homogeneity condition $mi_0 + (m-1)i_1 + \cdots + i_{m-1} + nj_0 + (n-1)j_1 + \cdots + j_{n-1} = mn$ is already contained in the three conditions already appearing in the above sum).

Cokernel of the Sylvester matrix. The fact that the resultant is defined as the determinant of the Sylvester matrix yields interesting properties, as this matrix carries more informations than simply the existence of roots.

Exercise 5.3.16. Assume that $A = K$ is a field and that $(a_0, b_0) \neq (0, 0)$. Show that

$$\operatorname{corank} \operatorname{Sylv}_{m,n}(f, g) = \deg \operatorname{GCD}(f, g).$$

Actually, one can recover the roots (in general numerical approximations of the roots) form the cokernel of the Sylvester matrix. Let us be more precise. Let $A = K$ be a field and assume that $(a_0, b_0) \neq (0, 0)$ and that $\operatorname{GCD}(f, g)$ is equal to $\prod_{i=1}^{r}(x - \alpha_i)^{m_i}$, $\alpha_i \neq \alpha_j$, in some extension $\bar{K}$ of K.

Lemma 5.3.17. *A basis of the cokernel of* $\operatorname{Sylv}_{m,n}(f, g)$ *is given by the columns of the block matrix*

$$V := \begin{pmatrix} V_{m+n-1}(\alpha_1; m_1) & V_{m+n-1}(\alpha_2; m_2) & \cdots & V_{m+n-1}(\alpha_r; m_r) \end{pmatrix} \tag{5.3.4}$$

where $V_d(\alpha; k)$ *is the generalized Vandermonde matrix*

$$V_d(\alpha; k) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \alpha & 1 & \cdots & 0 \\ \alpha^2 & 2\alpha & \cdots & 0 \\ \alpha^3 & 3\alpha^2 & \ddots & \vdots \\ \vdots & \vdots & \cdots & \frac{(d-1)!}{(d-k)!}\alpha^{d-k} \\ \alpha^d & d\alpha^{d-1} & \cdots & \frac{d!}{(d-k+1)!}\alpha^{d-k+1} \end{pmatrix}.$$

Proof. First, we prove that the columns of a block $V_{m+n-1}(\alpha; m)$, where α is a root of multiplicity at least m for f and g, belongs to the cokernel of the Sylvester matrix. Specializing the matrix equality (5.3.2) at α shows that the first column of $V_{m+n-1}(\alpha; m)$ belongs to the cokernel. Now, by computing the derivatives on both sides of (5.3.2), and then specializing at α, we see also that the second column of $V_{m+n-1}(\alpha; m)$ belongs to the cokernel. Following the same strategy, we can continue this way until we reach the multiplicity m of the root α.

To conclude, it remains to show that the $\delta := \sum_{i=1}^{r} m_i$ columns we got are linearly independent, because the dimension of the cokernel of the Sylvester matrix is precisely equal to this number by Exercise 5.3.16. For that purpose, consider the top square $\delta \times \delta$ submatrix of V; this is a generalized Vandermonde matrix and it is known that its determinant is equal to $\prod_{1 \leq i < j \leq r}(\alpha_j - \alpha_i)^{m_j m_i}$, which is nonzero since $\alpha_i \neq \alpha_j$ for all $i \neq j$. □

Now, let Δ_0 be the top square block of V defined by (5.3.4), of (maximal) size $\sum_{i=1}^r m_i = \deg \mathrm{GCD}(f,g)$ and define Δ_1 similarly with a shift down by one row.

Lemma 5.3.18. *With the above notation,* $\det(\Delta_1 - x\Delta_0)$ *is equal to* $\mathrm{GCD}(f,g)$ *up to a non-zero multiplicative constant. In particular, the generalized eigenvalues of the pencil* (Δ_1, Δ_0) *are* $\alpha_1, \ldots, \alpha_r$ *with multiplicity* $m_1, \ldots, m_r$*, respectively.*

Proof. Set $\delta := \sum_{i=1}^r m_i$ and consider the top submatrix of V, defined by (5.3.4), composed of the top $\delta+1$ rows and all the columns. It is a matrix of size $(\delta+1)\times\delta$. Now, denote by M the matrix obtained by adding a new column on the right which is equal to the column vector $(1, t, t^2, \cdots, t^\delta)^T$, where t is a new indeterminate. Consider the matrix

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -t & 1 & 0 & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & -t & 1 & 0 \\ 0 & \cdots & 0 & -t & 1 \end{pmatrix}$$

of size $(\delta+1)\times(\delta+1)$. Then, the product PM yields the matrix

$$\left(\begin{array}{ccc|c} 1 & \cdots & 1 & 1 \\ \hline & & & 0 \\ & \Delta_1 - t\Delta_0 & & 0 \\ & & & 0 \end{array}\right)$$

and hence we deduce that $\det(M) = \det(\Delta_1 - t\Delta_0)$ (notice that $\det(P) = 1$). But the matrix M is a generalized Vandermonde matrix where t can be seen as a new distinct root, so it follows that

$$\det(M) = \prod_{i=1}^{r}(t-\alpha_i)^{m_i} \prod_{i<j}(\alpha_j - \alpha_i)^{m_j m_i}.$$

From here, the claimed result follows as $\prod_{i<j}(\alpha_j-\alpha_i)^{m_j m_i} \neq 0$. □

Exercise 5.3.19. Consider the two univariate polynomials

$$f(x) = x^5 - 12x^3 + 13x^2 + 12x - 12, \ g(x) = -x^5 + 5x^4 - 5x^3 - 7x^2 + 8x + 4.$$

Decide if they have any common root, and if yes compute them by means of Lemma 5.3.18 (you might want to use a computer for these computations).

Resultant of homogeneous polynomials. Finally, we introduce the homogeneous setting which allows us to simplify the notation. Consider the two homogeneous polynomials in the graded polynomial ring $A[x,y]$ obtained by homogenization of (5.3.1):

$$\left\{ \begin{array}{lcl} F(x,y) & := & a_0x^m + a_1x^{m-1}y + \cdots + a_my^m \\ G(x,y) & := & b_0x^n + b_1x^{n-1}y + \cdots + b_ny^n \end{array} \right. \tag{5.3.5}$$

We define the Sylvester matrix and resultant of F and G by

$$\mathrm{Sylv}(F,G) := \mathrm{Sylv}_{m,n}(F(x,1),G(x,1)),$$

$$\mathrm{Res}(F,G) := \mathrm{Res}_{m,n}(F(x,1),G(x,1)) = \det(\mathrm{Sylv}_{m,n}(F(x,1),G(x,1))).$$

Exercise 5.3.20. With the above notation, assuming that A is a domain and denoting $K = \mathrm{Frac}(A)$, show that:

i) $\mathrm{Res}(F,G) = 0$ if and only if F and G has a common root in the projective line $\mathbb{P}^1$ over the algebraic closure of K.

ii) if $A = K$ then $\mathrm{corank}\,\mathrm{Sylv}(F,G) = \deg \mathrm{GCD}(F,G)$.

iii) If $A = K$ and $\mathrm{GCD}(F,G)$ is equal to $y^{m_\infty}\prod_{i=1}^{r}(x-\alpha_i y)^{m_i}$, $\alpha_i \neq \alpha_j$, in some extension $\bar{K}$ of K, up to a nonzero multiplicative constant, then a basis of the cokernel of $\mathrm{Sylv}(F,G)$ is given by the columns of the matrix

$$\begin{pmatrix} V_{m+n-1}(\infty;m_\infty) & V_{m+n-1}(\alpha_1;m_1) & V_{m+n-1}(\alpha_2;m_2) & \cdots & V_{m+n-1}(\alpha_r;m_r)\end{pmatrix}$$

with

$$V_{m+n-1}(\infty;m_\infty) = \left(\frac{0}{\mathrm{Id}_{m_\infty}}\right),$$

where the top block is a null matrix and the bottom block is the identity matrix of size $m_\infty \times m_\infty$.

For the sake of completeness, we close this paragraph with the invariance property of the resultant under the action of SL_2; we will prove it in a more general setting when we will introduce the Macaulay resultant.

Proposition 5.3.21 (Invariance)**.** *Let $ax+by$ and $cx+dy$ in $A[x,y]$ be two linear forms, and let $F(x,y)$ and $G(x,y)$ be two homogeneous polynomials of degree m and n, respectively. Then, the following equality holds in A:*

$$\mathrm{Res}(F(ax+by,cx+dy),G(ax+by,cx+dy)) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}^{mn}\mathrm{Res}(F,G).$$

Proof. [CLO98, Theorem 3.5] and references therein. See also next chapter for a proof in a more general framework. □

Exercise 5.3.22 (Discriminant)**.** Let A be a commutative ring and suppose given

$$f(x) = a_0x^m + a_1x^{m-1} + \cdots + a_m \in A[x]$$

with $m \geq 2$.

1. Assuming A is the universal ring of coefficients $\mathbb{Z}[a_0,\ldots,a_m]$, show that there exist an element in A, denoted $\mathrm{Disc}_m(f)$ such that

$$\mathrm{Res}_{m,m-1}(f,\partial f/\partial x) = a_0\mathrm{Disc}_m(f).$$

It is called the *discriminant* of f. If A is an arbitrary commutative ring, then the discriminant of f is defined by specialization from the universal setting.

2. Give the degree of the discriminant with respect to the coefficients of f.
3. Let $F(x,y)$ be the homogenization of f with respect to the variable y; it has degree m. Prove that

$$\mathrm{Res}(\partial F/\partial x, \partial F/\partial y) = m^{m-2}\mathrm{Disc}(F).$$

4. Prove the following polarization formula: for any couple of homogeneous polynomials $F(x,y)$ and $G(x,y)$ of degree d and e respectively,

$$\mathrm{Disc}(FG) = (-1)^{de}\mathrm{Disc}(F)\mathrm{Disc}(G)\mathrm{Res}(F,G)^2.$$

5.4 Bézout Theorem for Plane Curves

We now provide a quick proof of Bézout Theorem (see Theorem 2.1.1) which uses the Sylvester resultant as a key ingredient. The main idea is to project the intersection points of the two curves on one of the coordinate axes. Although we will not discuss this with details, we mention that this computational approach, which was essentially the original one used by Bézout, also leads to algorithms for computing the intersection points.

Let k be an algebraically closed field and let f, g be two homogeneous polynomials in $k[x,y,z]$ of degree d, e, respectively, with no common factor. Bézout theorem claims that the two projective curves $V(f)$ and $V(g)$ in $\mathbb{P}^2_k$ meet in d, e points, counted with multiplicities.

Considering the polynomials f and g as univariate polynomials in y with coefficients in $A = k[x,z]$, we write

$$\left\{ \begin{array}{rcl} f(x,y,z) & = & a_0(x,z)y^m + a_1(x,z)y^{m-1} + \cdots + a_{m-1}(x,z)y + a_m(x,z) \\ g(x,y,z) & = & b_0(x,z)y^n + b_1(x,z)y^{n-1} + \cdots + b_{n-1}(x,z)y + b_n(x,z) \end{array} \right. \tag{5.4.1}$$

where $a_i(x,z)$ and $b_j(x,z)$ are homogeneous polynomials in A, with $a_0 \neq 0$ and $b_0 \neq 0$. By a linear change of coordinates, one can assume that the point $\infty_y = (0:1:0)$ does not belong to $V(f) \cup V(g) \subset \mathbb{P}^2$, which implies that a_0 and b_0 are actually nonzero constant in k. It also follows that $m = d$, $n = e$, $a_i(x,z)$ are homogeneous polynomials in A of degree i and $b_j(x,z)$ are homogeneous polynomials in A of degree j.

Now, since f ang g have no common factor (in $k[x,y,z]$), their intersection consists of finitely many points, say $\{p_1, \ldots, p_r\}$ (this is a slice of the curve $V(f)$ by the curve $V(g)$). In particular, $\mathrm{Res}_{m,n}(f,g)$ is a nonzero polynomial in $k[x,z]$ (otherwise, for any $(x:z) \in \mathbb{P}^1$ the polynomials f ang g would have a common root in y, which is not possible). Set $p_i = (x_i : y_i : z_i)$ for all $i = 1, \ldots, r$. Then, by property of the resultant, we can write

$$\mathrm{Res}_{m,n}(f,g) = c\prod_{i=1}^{r}(z_i x - x_i z)^{m_i}$$

where the m_i's are integers and c is a nonzero constant in k. Moreover, by Proposition 5.3.15, $\sum_i m_i = mn = de$. We notice that by a linear change of coordinates, one can assume that $(x_i : 0 : z_i) \neq (x_j : 0 : z_j)$ for all $i \neq j$. In other words, one can assume that any two intersection points of $V(f)$ and $V(g)$ will not project on the same point after elimination of the variable y. Thus, defining m_i as the intersection multiplicity of the point $p_i \in V(f) \cap V(g)$, the theorem is proved.

Intersection multiplicity. Let $p \in V(f) \cap V(g)$. By de-homogenizing and applying a linear change of coordinates, one can assume that $p = (0,0) \in \mathbb{A}^2$. Providing that $V(f) \cap V(g) \cap V(x) = \{p\}$, we have defined the intersection multiplicity of p, denoted $m_p(V(f), V(g))$, as the valuation with respect to x of the resultant of f and g with respect to y:

$$m_p(V(f), V(g)) := \operatorname{val}_x \operatorname{Res}_y(f, g).$$

To clarify the notation, if $\operatorname{Res}_y(f,g) = x^m \prod (x - x_i)^{\xi_i}$, where $x_i \neq 0$ for all i, then $\operatorname{val}_x \operatorname{Res}_y(f,g) = m$.

One can show that this valuation agrees with the dimension, as a k-vector space, of the ring $k[x,y]/(f,g)$ localized at the point p (the more common definition of the intersection multiplicity). This shows in particular that this valuation does not depend on the choice of coordinates. Also, if $V(f) \cap V(g) \cap V(x) = \{p_1, \ldots, p_s\}$ then

$$\operatorname{val}_x \operatorname{Res}_y(f,g) = \sum_{i=1}^{s} m_{p_i}(V(f), V(g)).$$

Computation of intersection points. By combining the strategy of the proof of Bézout theorem and the methods exploiting the cokernel of the Sylvester matrix to compute the common roots of two polynomials, one can devise algorithms for computing the intersection points between two algebraic plane curves. We will not go further in this direction and we encourage the reader to do some experiments.

Exercise 5.4.1 (Singular points)**.** Assume that k is an algebraically closed field of characteristic zero. Suppose given a non-constant polynomial $f(x,y) \in k[x,y]$ and denote by $\mathcal{C}$ the algebraic curve defined by f in $\mathbb{A}^2$.

Let P be a point on $\mathcal{C}$; up to change of coordinates one can assume that $P = (0,0)$. Show that the following are equivalent.

i) $\partial f/\partial x$ and $\partial f/\partial x$ does not vanish simultaneously at $P = (0,0)$.

ii) $f(x,y)$ is of order 1 et P.

iii) In the pencil of lines passing through P, all, except finitely many, have an intersection multiplicity at P with $\mathcal{C}$ equal to 1.

A point P that satisfies to the above properties is said to be a regular point on $\mathcal{C}$, otherwise P is said to be a *singular point*. Prove that if $f(x,y)$ is square-free, then the set of singular points of $\mathcal{C}$ is finite.

5.5 Implicitization of Plane Curves

In this section we focus on the image of a plane curve parameterization. A situation we already encountered is the computation of the image of a polynomial map

$$\begin{array}{rcl}\phi : \mathbb{A}^1 & \to & \mathbb{A}^2 \\ t & \mapsto & (f_1(t), f_2(t)),\end{array}$$

where f_1 and f_2 are univariate polynomials in the variable t. From Corollary 5.2.2 and the properties of the Sylvester resultant, we expect the resultant of $x_1 - f_1(t)$ and $x_2 - f_2(t)$ to give an implicit equation of the image of ϕ.

Example 5.5.1. Let $f_1(t) = t$ and $f_2(t) = t^2$. Then, $\mathrm{Res}_{1,2}(x_1 - t, x_2 - t^2) = x_2 - x_1^2$ and the elimination ideal $(x_1 - t, x_2 - t^2) \cap k[x_1, x_2]$ is also generated by $x_2 - x_1^2$. Now, let $f_1(t) = t^2$ and $f_2(t) = t^4$. Obviously, the set-theoretic image of the map is not changed by substituting t with t^2, so the implicit equation of the image should not changed. Indeed, the elimination ideal $(x_1 - t^2, x_2 - t^4) \cap k[x_1, x_2]$ is generated by $x_2 - x_1^2$. However, $\mathrm{Res}_{2,4}(x_1 - t^2, x_2 - t^4) = (x_2 - x_1^2)^2$.

Degree of ϕ. This example illustrates how the resultant captures what is called the *degree* of the map ϕ (more precisely of ϕ co-restricted to its image). Assuming that k is algebraically closed of characteristic zero, this degree is the number of pre-images of a general point in the image of ϕ. In the previous example, this degree is 2 after the substitution of t by t^2 because any point on the parabola $x_2 - x_1^2$ will have two pre-images. Why is this number well defined, that is why the number of pre-images of a general point in the image of ϕ is constant? This can be seen with the Sylvester resultant. Consider the Sylvester matrix

$$M(x_1, x_2) := \mathrm{Sylv}_{\deg(f_1),\deg(f_2)}(x_1 - f_1(t), x_2 - f_2(t)),$$

which depends on the implicit coordinates x_1, x_2. The determinant of $M(f_1(t), f_2(t))$ is obviously equal to zero, but this matrix has a positive corank over the fraction field $k(t)$. This corank is the corank of $M(f_1(t), f_2(t))$ where t is specialized to a general value (which is described by the vanishing of some minors of M). In addition, Exercise 5.3.16 shows that this corank is precisely the number of pre-images of such a general point on the image. See [Har92, Lecture 7] for more details on the degree of maps.

The projective setting. In order to explain better the observation in Example 5.5.1, we introduce the projective version of our problem, that will moreover allow to deal with maps defined by rational functions not only polynomial functions, which enlarge the class of curves we consider.

Exercise 5.5.2. Is the plane circle can be seen as the image of a polynomial map? If not, what parameterization can you suggest? How could you compute an implicit equation?

Consider a map

$$\begin{array}{rcl} \phi : \mathbb{P}^1 & \to & \mathbb{P}^2 \\ (s:t) & \mapsto & (f_0(s,t) : f_1(s,t) : f_2(s,t)) \end{array}$$

where f_0, f_1 ad f_2 are homogeneous polynomials of the same degree $d \geq 1$ in the ring $R = k[s,t]$, k being assumed to be an algebraically closed field of characteristic zero (for simplicity). Without loss of generality, one can assume that f_0, f_1, f_2 have no common factor in R, which implies that these polynomials have no common root in P^1 and that ϕ is defined at any point of P^1 (this is called a regular map).

The degree formula. Suppose that the image of ϕ is a plane curve $\mathcal{C}$. If $C(x_0, x_1, x_2)$ is a reduced homogeneous polynomial defining $\mathcal{C}$, then the degree of the curve is equal to the degree of the polynomial C (see Section 2.4); we will denote it by $\deg(\mathcal{C})$. From what we proved in Section 3.3, it is equal to the number of intersection points between $\mathcal{C}$ and a general line in P^2.

Now, we denote by $\deg(\phi)$ that degree of the map ϕ; as explained above, this is the number of pre-images of a general point on $\mathcal{C}$ via ϕ. Then, we have the following equality:

$$d = \deg(\phi)\deg(\mathcal{C}). \tag{5.5.1}$$

In particular, if ϕ is generically injective, i.e. a general point on $\mathcal{C}$ has a single pre-image via ϕ, then $\mathcal{C}$ is a curve of degree d. Why (5.5.1) holds? Consider a general line $\mathcal{L}$ in P^2, say of equation $\alpha_0 x_0 + \alpha_1 x_1 + \alpha_2 x_2 = 0$, where α_i's are constants in k. Then, the curve $\mathcal{C}$ and the line $\mathcal{L}$ intersect in $\deg(\mathcal{C})$ points. These points are in correspondence via ϕ with the roots of the equation $\alpha_0 f_0 + \alpha_1 f_1 + \alpha_2 f_2 = 0$ which defines d points in $\mathbb{P}^1$ (because it is a polynomial equation of degree d in two homogeneous variables). As the line is general, each intersection point of $\mathcal{L}$ and $\mathcal{C}$ gives $\deg(\phi)$ roots of $\sum_{i=0}^{2} \alpha_i f_i(s,t) = 0$, and (5.5.1) follows.

A first implicitization formula. A classical approach to determine a defining polynomial of the curve $\mathcal{C}$ is to consider the restriction $\tilde{\mathcal{C}}$ of $\mathcal{C}$ to the affine chart $\mathbb{A}^2$ of $\mathbb{P}^2$, which corresponds to points such that $x_0 \neq 0$. Thus, we assume that $f_0 \neq 0$ otherwise the image of ϕ is contained in the line at infinity $x_0 = 0$. The curve $\tilde{\mathcal{C}}$ is parameterized by

$$\begin{array}{rrcl} \tilde{\phi} : & \mathbb{P}^1 \setminus V(f_0) & \to & \mathbb{A}^2 \\ & (s:t) & \mapsto & \left(\frac{f_1(s,t)}{f_0(s,t)}, \frac{f_2(s,t)}{f_0(s,t)}\right). \end{array}$$

Thus, the point $(1 : x_1 : x_2)$ belongs to $\tilde{\mathcal{C}}$ if and only if there exists $(s_0 : t_0) \in \mathbb{P}^1 \setminus V(f_0)$ such that

$$(x_1, x_2) = \left(\frac{f_1(s_0,t_0)}{f_0(s_0,t_0)}, \frac{f_2(s_0,t_0)}{f_0(s_0,t_0)}\right).$$

In other words, the graph of $\tilde{\phi}$ in $\mathbb{P}^1 \times \mathbb{A}^2$ is defined by the two polynomial equations

$$f_1(s,t) - x_1 f_0(s,t) = 0, \quad f_2(s,t) - x_2 f_0(s,t) = 0$$

(recall that f_0, f_1, f_2 cannot vanish simultaneously by our assumption). Therefore, it is expected that the elimination of the homogeneous variables (s,t) from this polynomial system of two equations yields a defining polynomial of $\tilde{\mathcal{C}}$.

Lemma 5.5.3. *With the above notation,*

$$\mathrm{Res}(f_1(s,t) - x_1 f_0(s,t), f_2(s,t) - x_2 f_0(s,t)) = \tilde{C}(x_1,x_2)^{\deg(\phi)}$$

where $\tilde{C}(x_1,x_2) = C(1,x_1,x_2)$ *is a defining polynomial of* $\tilde{\mathcal{C}}$*, which is of degree* $d/\deg(\phi)$

Proof. First, we notice that that $\mathcal{C}$ is an irreducible plane curve in $\mathbb{P}^2$, as the image of an irreducible variety. In more algebraic terms, $k[x_0,x_1,X_2]/(C)$ is a domain because there is a canonical injective map to $k[s,t]$ (this map sends x_i to $f_i(s,t)$, it is the algebraic counterpart of ϕ). It follows that $\tilde{C}(x_1,x_2) = C(1,x_1,x_2)$ is irreducible.

Now, we claim that the resultant and $\tilde{C}$ define the same algebraic varieties in $\mathbb{A}^2$. This follows form the property of the Sylvester resultant (recall that f_0, f_1 and f_2 cannot vanish simultaneously). Therefore, there exists an integer p such that

$$\mathrm{Res}(f_1(s,t) - x_1 f_0(s,t), f_2(s,t) - x_2 f_0(s,t)) = \tilde{C}(x_1,x_2)^p.$$

Applying the degree formula (5.5.1), it is enough to prove that the above resultant is a polynomial of degree d in x_1, x_2 to conclude the proof. This property can be seen from the definition of the resultant as the determinant of the Sylvester matrix and the multi-linearity of the determinant. Indeed, denote by $S_i^0, \ldots, S_i^{d-1}$ the d columns corresponding to the polynomial f_i in a Sylvester block matrix built in degree d. Then,

$$\begin{aligned}&\mathrm{Res}(f_1(s,t) - x_1 f_0(s,t), f_2(s,t) - x_2 f_0(s,t)) = \\ &\qquad \det\langle S_1^0 - x_1 S_0^0, \ldots, S_1^{d-1} - x_1 S_0^{d-1}, S_2^0 - x_2 S_0^0, \ldots, S_2^{d-1} - x_2 S_0^{d-1}\rangle.\end{aligned}$$

Now, by multi-linearity of the determinant, it is clear that the coefficient of any monomial in x_1, x_2 of degree $> d$ will vanish, so the resultant is of degree at most d. To see that it is exactly d, we notice that the coefficient of x_1^d, respectively x_2^d, in the expansion of the resultant is $\mathrm{Res}(f_0, f_2)$, respectively $\mathrm{Res}(f_1, -f_0)$. If one of these two resultants are nonzero, then we are done, otherwise one needs to argue via some more technical arguments relying on changes of coordinates, which we will not discuss here. □

Exercise 5.5.4. Consider the following parameterization of a circle

$$\begin{array}{ccc}\mathbb{P}^1 & \to & \mathbb{P}^2 \\ (s:t) & \mapsto & (s^2+t^2 : s^2-t^2 : 2st).\end{array}$$

Applying Lemma 5.5.3, compute an implicit equation of the circle.

We notice that the Sylvester matrix of $f_1(s,t) - x_1 f_0(s,t)$ and $f_2(s,t) - x_2 f_0(s,t)$ is of size $2d \times 2d$ where as it determinant is a degree d polynomial. This gap in the degrees can be explained by coming back to $\mathbb{P}^2$. Indeed, by homogenizing equations with respect to the variable x_0, we obtain

$$\operatorname{Res}(x_0 f_1(s,t) - x_1 f_0(s,t), x_0 f_2(s,t) - x_2 f_0(s,t)) = x_0^d \, C(x_0, x_1, x_2)^{\deg(\phi)} \qquad (5.5.2)$$

This equality holds because the determinant of the corresponding Sylvester matrix is a homogeneous polynomial of degree $2d$ and it must vanishes when $x_0 = 0$, which implies that x_0^d is a factor in (5.5.2). Thus, the resultant (5.5.2) yields a curve of degree $2d$ in $\mathbb{P}^2$ which is the union of our curve $\mathcal{C}$ and the line at infinity, i.e. the line of equation $x_0 = 0$, with multiplicity d (the number of roots of $f_0(s,t)$). The reason why this line at infinity appears in this resultant is because we chose two equations among the three equations that are needed to fully express the homogeneous constraint

$$(x_0 : x_1 : x_2) = \phi(s:t) = (f_0(s,t) : f_1(s,t) : f_2(s,t)), \qquad (5.5.3)$$

namely the three equations

$$x_0 f_1(s,t) - x_1 f_0(s,t) = 0, \;\; x_0 f_2(s,t) - x_2 f_0(s,t) = 0, \;\; x_1 f_2(s,t) - x_2 f_1(s,t) = 0. \qquad (5.5.4)$$

These equations are the 2-minors of the matrix

$$\begin{pmatrix} f_0(s,t) & f_1(s,t) & f_1(s,t) \\ x_0 & x_1 & x_2 \end{pmatrix}$$

and they clearly define the graph of ϕ in $\mathbb{P}^1 \times \mathbb{P}^2$. It turns out that the third equation $x_1 f_2(s,t) - x_2 f_1(s,t) = 0$ is redundant if $x_0 \neq 0$, but it is not if $x_0 = 0$. In order to fix this problem, we will refine our approach by considering other equations in the defining ideal of the graph of ϕ.

Syzygies of curve parameterizations. A syzygy of the polynomials f_0, f_1, f_2 is a triple of polynomials g_0, g_1, g_2 in $R = k[s,t]$ such that $\sum g_i f_i = 0$. It can be identified with the polynomial $\sum_{i=0}^2 x_i g_i(s,t) \in R[x_0, x_1, x_2]$ which is a linear form in x_0, x_1, x_2 with coefficients in R. Thus, the equations (5.5.4) are syzygies of f_0, f_1, f_2, that are actually Koszul syzygies (see Section 3.4); we denote by I_K the ideal generated by the three Koszul syzygies (5.5.4). We already noticed that the algebraic variety defined by the ideal I_K is the graph of ϕ in $\mathbb{P}^1 \times \mathbb{P}^2$.

Denote by I_S the ideal of $R[x_0, x_1, x_2]$ generated by all the syzygies of f_0, f_1, f_2. Clearly, $I_K \subset I_S$ so that $V(I_S)$, which denotes the algebraic variety defined by the ideal I_S, is contained in the graph of ϕ. Now, by definition any syzygy of f_0, f_1, f_2 vanishes on the graph of ϕ so we deduce that $V(I_S) = V(I_K)$ and hence that $V(I_S)$ is also the graph of ϕ.

Remark 5.5.5. Actually, one can be a little more precise: for any syzygy $\sum_{i=0}^2 x_i g_i$ we have the equality

$$f_0(x_0 g_0 + x_1 g_1 + x_2 g_2) = g_1(x_1 f_0 - x_0 f_1) + g_2(x_2 f_0 - x_0 f_2),$$

as well as two similar equalities replacing f_0 by f_1 and f_2 on the left-hand side. It follows that the ideals I_S and I_K are equal after localization by f_i for all $i = 0, 1, 2$. Since $V(f_0, f_1, f_2) = \emptyset$ by assumption, this implies that I_S and I_K have the same saturation with respect to the homogeneous ideal (s,t), i.e. $I_S : (s,t)^\infty = I_K : (s,t)^\infty$.

As $I_K \subset I_S$, one can expect to find some non-obvious syzygies in I_S that would help to get ride of the extraneous factor x_0^d appearing in (5.5.2). Actually, the situation is particularly nice because of the following property.

Theorem 5.5.6 (Hilbert-Burch Theorem)**.** *The ideal I admits a finite free resolution of the form*

$$0 \to \oplus_{i=1}^2 R(-d-\mu_i) \xrightarrow{\psi} R^3(-d) \xrightarrow{(f_0 \ f_1 \ f_2)} R \to R/I \to 0$$

where $\mu_1 \le \mu_2$ are non-negative integers such that $\mu_1 + \mu_2 = d$.

Proof. From the Hilbert Syzygy Theorem (see Theorem 3.2.3), the finite free resolution of I is of the form

$$0 \to \oplus_{i=1}^n R(-a_i) \to R(-d)^3 \to R \to R/I \to 0. \tag{5.5.5}$$

Since the f_i's have no common root, the Hilbert polynomial of R/I is equal to zero. Therefore,

$$\mathrm{HP}_R(\ell) - 3\mathrm{HP}_{R(-d)}(\ell) + \sum_{i=1}^n \mathrm{HP}_{R(-a_i)}(\ell) = 0.$$

It follows that $n = 2$ and that $3d - a_1 - a_2 = 0$. Setting $a_1 = d + \mu_1$ and $a_2 = d + \mu_2$, this latter condition gives $d = \mu_1 + \mu_2$, as claimed. □

Remark 5.5.7. In Theorem 5.5.6, it can also be proved that the ideal generated by the 2-minors of a matrix of ψ is equal to I, up to multiplication by a non-zero constant in k. We refer to [Eis95, §20.4]; see also [CLO98, Theorem 4.17].

The graded R-module of syzygies of ϕ,

$$\mathrm{Syz}(\phi) = \{(g_0, g_1, g_2) \in R^3 : g_0 f_0 + g_1 f_1 + g_2 f_2 = 0\},$$

is hence a free module generated in degree μ_1 and μ_2. Let $p = (p_0, p_1, p_2)$, $q = (q_0, q_1, q_2)$ be a basis of this module with $\deg p = \mu_1$ and $\deg q = \mu_2$; they form the two columns of a matrix of ψ. Using the identification of syzygies of I with linear forms in x_0, x_1, x_2, we define

$$L_1(s,t;x_0,x_1,x_2) = x_0 p_0(s,t) + x_1 p_1(s,t) + x_2 p_2(s,t),$$

$$L_2(s,t;x_0,x_1,x_2) = x_0 q_0(s,t) + x_1 q_1(s,t) + x_2 q_2(s,t),$$

so that $I_S = (L_1, L_2) \subset R[x_0, x_1, x_2]$. It follows that the graph of ϕ is actually a complete intersection defined by L_1 and L_2. Consequently, the implicitization formula (5.5.2) can be refined by taking the resultant of L_1 and L_2 with respect to the homogeneous variables s, t; we have

$$\mathrm{Res}(L_1, L_2) = C(x_0, x_1, x_2)^{\deg(\phi)} \tag{5.5.6}$$

where $C(x_0, x_1, x_2)$ is an implicit equation of $\mathcal{C}$.

Exercise 5.5.8. Take again the parameterization of a plane circle given in Exercise 5.5.4 and apply (5.5.6).

We mention that the two polynomials L_1 and L_2 have been first introduced by the geometric modeling community to solve the implicitization problem for plane rational curves. They are called *moving lines following the parameterization* ϕ because of the following geometric interpretation: For any parameter value $(s : t) \in \mathbb{P}^1$, each polynomials L_1 and L_2 define a line in $\mathbb{P}^2$. When the parameter $(s : t)$ varies, these two lines move as well, hence the terminology of moving lines. In addition, for all parameter values $(s : t) \in \mathbb{P}^1$, the lines L_1 and L_2 are linearly independent and they both go through the point $\phi(s : t) \in \mathbb{P}^2$, which explains the terminology *moving lines following the parameterization* ϕ. Observe that the two Koszul syzygies we considered earlier, namely

$$x_0 f_1(s,t) - x_1 f_0(s,t) = 0, \quad x_0 f_2(s,t) - x_2 f_0(s,t) = 0,$$

both satisfy the second property but not the first one. Indeed, they define the same line for all parameters $(s : t) \in \mathbb{P}^1$ such that $f_0(s,t) = 0$ (there are d of them, counting multiplicities, which explains the factor x_0^d in (5.5.2)).

Pre-images and eigenvalues. Given a point on the curve $\mathcal{C}$, an important problem in practice is to determine its corresponding pre-image(s), i.e. its corresponding parameter values, via the parameterization ϕ.

As the graph of ϕ is defined by the ideal $I_S = (L_1, L_2)$, we deduce that for any point $P \in \mathbb{P}^2$,

$$\mathrm{GCD}(L_1(s,t;P), L_2(s,t;P)) = \prod_{i=1}^{r_P} (\beta_i s - \alpha_i t)^{m_i} \tag{5.5.7}$$

where the product is taken over all distinct pairs $(\alpha_i : \beta_i) \in \mathbb{P}^1$ such that $\phi(\alpha_i : \beta_i) = P$, i.e. all pre-images of P via ϕ. The multiplicity m_i is called the multiplicity of the branch curve at $\phi(\alpha_i : \beta_i)$, and the multiplicity of the point P on $\mathcal{C}$ is defined as $\sum_{i=1}^{r_P} m_i = m_P(\mathcal{C})$. We notice that (5.5.7) is a constant if and only if $P \notin \mathcal{C}$.

Remark 5.5.9. Observe that from the above discussions, we see some connections between the degrees μ_1 and μ_2 of minimal syzygies and singular points on $\mathcal{C}$. For instance, if $P \in \mathcal{C}$ is a point of multiplicity $m \geq 2$, then $m \leq \mu_1$ or $m = \mu_2$. If $\mu_1 < m$, then the equality $m = \mu_2$ implies $L_1(s,t;P) = 0$.

Since the resultant of L_1 and L_2 can be computed as the determinant of the corresponding Sylvester matrix, Equation (5.5.7) can be turned into linear algebra computations, as explained in Section 5.3. Compared to computations with polynomial equations, this point of view allows to rely on well established methods from linear algebra, but especially it allows to deal with approximate data, which is of capital importance for applications in the field of geometric modeling.

From now on, we will denote by $\mathbb{M}$ the Sylvester matrix (with respect to the variables s, t) of the two polynomials $L_1 = \sum_{i=0}^{2} x_i p_i$ and $L_2 = \sum_{i=0}^{2} \sum x_i q_i$ defined in the previous section, i.e.

$$\mathbb{M}(x_0, x_1, x_2) := \mathrm{Sylv}(L_1, L_2) \tag{5.5.8}$$

It is a $d \times d$-matrix whose entries are linear forms in $k[x_0, x_1, x_2]$. Given any point $P \in \mathbb{P}^2$ we denote by $\mathbb{M}(P)$ the evaluation of $\mathbb{M}$ at P. Then, for any point P in $\mathbb{P}^2$,

$$\mathrm{corank}(\mathbb{M}(P)) = m_P(\mathcal{C}).$$

Moreover, the pre-images of P via ϕ can be extracted form the cokernel of $\mathbb{M}(P)$, as explained in Section 5.3. We illustrate it with the following example.

Example 5.5.10. Consider the curve parameterization given by

$$f_0 = s^3, \;\; f_1 = st^2 - s^2t = st(t-s), \;\;\; f_2 = 2t^3 - 7st^2 + 5s^2t = t(s-t)(5s-2t).$$

Computations show that $\mu_1 = 1$, $\mu_2 = 2$ and

$$\mathbb{M}(x_0, x_1, x_2) = \begin{pmatrix} 5x_1 + x_2 & 0 & x_1 \\ -2x_1 & 5x_1 + x_2 & x_0 \\ 0 & -2x_1 & -x_0 \end{pmatrix}.$$

The rank of $\mathbb{M}(1, 0, 0)$ is equal to 1, so this point has 2 pre-images. The computation of the corresponding cokernel yields a vector space generated by the two vectors $(1, 0, 0)$ and $(0, 1, 1)$, with respect to the monomial basis (s^2, st, t^2). We observe that $\phi(0 : 1) = (0 : 0 : 1)$ so the point at infinity $(0 : 1)$ does not belong to the fiber of the point $(1 : 0 : 0)$. Thus, we recover its two pre-images by solving the eigenvalue problem $\det(\Delta_1 - t\Delta_0) = 0$ where

$$\Delta_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \;\; \Delta_1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

We obtain $(s : t) = (1 : 0)$ and $(s : t) = (1 : 1)$, as expected.

Inverse maps. When a point P has a single point pre-image, i.e. $m_P(\mathcal{C}) = 1$, then the above process to compute its pre-image yields a column vector and the eigenvalue computation reduces to a ratio. This fact can be used to compute an inverse of ϕ when it is a generically injective ($\deg(\phi) = 1$).

Let $\mathbb{T}$ be a submatrix of $\mathbb{M}$ which is obtained by removing one column of $\mathbb{M}$ and which is chosen such that $\mathrm{rank}\, \mathbb{T}(P) = d - 1$ for a general point P on $\mathcal{C}$. We notice that such a matrix $\mathbb{T}$ exists if and only if ϕ is birational onto $\mathcal{C}$. It follows that the column vector of signed $(d-1)$-minors of $\mathbb{T}$

$$\Big(\det(\mathbb{T}_0), -\det(\mathbb{T}_1), \ldots, (-1)^d \det(\mathbb{T}_d)\Big),$$

where $\mathbb{T}_i$ is the minor of $\mathbb{T}$ obtained by removing the row number $i+1$, is a basis for the cokernel of $\mathbb{M}_{d-1}$ after evaluation at a general point on $\mathcal{C}$. Consequently, the maps

$$\begin{array}{ccc} \mathbb{P}^2 & \dashrightarrow & \mathbb{P}^1 \\ (x_0 : x_1 : x_2) & \mapsto & (\det(\mathbb{T}_i) : -\det(\mathbb{T}_{i+1})) \end{array}$$

for all $i = 0, \ldots, d-1$, give the inverse of ϕ when restricted to $\mathcal{C}$.

Example 5.5.11. Consider the following parameterization of a circle:

$$f_0 = s^2 + t^2,\ f_1 = 2st,\ f_2 = s^2 - t^2.$$

Then, the computation of the matrix $\mathbb{M}$ gives

$$\mathbb{M} = \begin{pmatrix} x_1 & -x_0 + x_2 \\ -x_0 - x_2 & x_1 \end{pmatrix}$$

where the columns are indexed with the monomial basis $\{s, t\}$ (from top to bottom). We deduce two inversion formulas for ϕ from the two columns of $\mathbb{M}$, namely

$$\mathbb{P}^2 \dashrightarrow \mathbb{P}^1 : (x_0 : x_1 : x_2) \mapsto (-x_0 - x_2 : -x_1),$$

$$\mathbb{P}^2 \dashrightarrow \mathbb{P}^1 : (x_0 : x_1 : x_2) \mapsto (x_1 : x_0 - x_2).$$

They both coincide after restriction to $\mathcal{C}$; here is the `Macaulay2` code:

To conclude this chapter, we mention that from a computational point of view, the computation of a basis (p_0, p_1, p_2) and (q_0, q_1, q_2) of the syzygy module $\mathrm{Syz}(\phi)$ of I is not needed to build the matrix $\mathbb{M}$. Indeed, any matrix whose columns form a basis of the k-vector space $\mathrm{Syz}(\phi)_{d-1}$ of syzygies of I of degree $d-1$ can be used in the place of the Sylvester matrix $\mathbb{M}$ (which actually correspond to a specific choice of basis for $\mathrm{Syz}(\phi)_{d-1}$). The computation of a basis of $\mathrm{Syz}(\phi)_{d-1}$ amounts to solve a linear system, which can also be done approximately via Singular Value Decomposition.

Chapter 6

Resultants over a Projective Space

The goal of this section is to introduce resultants of homogeneous multivariate polynomials, more precisely of n homogeneous polynomials in n variables. It is a generalization of the Sylvester resultant of 2 homogeneous polynomials in 2 variables (see Section 5.3). We begin with some preliminaries about the elimination of variables in a list of an arbitrary number of homogeneous multivariate polynomials.

6.1 The Elimination Theorem

We suppose given r homogeneous polynomials in n variables $x_1, \ldots, x_n$:

$$f_i(x_1, \ldots, x_n) = \sum_{|\alpha|=d_i} u_{i,\alpha} x^\alpha, \quad i = 1, \ldots, r.$$

We use the same notation as in Chapter 4 for monomials, that is $\alpha = (\alpha_1, \ldots, \alpha_n)$ denotes a multi-index and x^α is the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ which is of degree $|\alpha| = \sum_{i=1}^n \alpha_i$. We assume that $n \geq 1$, $r \geq 1$ and $d_i \geq 1$ for all $i = 1, \ldots, n$.

We define the ring

$$A := \mathbb{Z}[u_{i,\alpha} \ : \ i = 1, \ldots, r, |\alpha| = d_i]$$

which is called the *universal ring of coefficients*. We also define the polynomial ring $C = A[x_1, \ldots, x_n]$ that we see as a graded ring by setting $\deg(x_i) = 1$ and $\deg(u_{i\alpha}) = 0$. Thus, $f_i \in C_{d_i}$ for all i. Finally, we set $I = (f_1, \ldots, f_r) \subset C$ and $\mathfrak{m} = (x_1, \ldots, x_n) \subset C$; they are both homogeneous ideals of C.

Specialization. Let k be a field and suppose that a ring map $\rho : A \to k$ is given, sending $u_{i,\alpha}$ to $c_{i,\alpha} \in k$. The map ρ is called a specialization map as it corresponds to specializing the coefficients of the ‘generic” polynomials $f_1, \ldots, f_r$ to the field k. More precisely, the polynomial $f_i \in A[x_1, \ldots, x_n]$ is specialized to the polynomial

$$\rho(f_i)(x_1, \ldots, x_n) = \sum_{|\alpha|=d_i} c_{i,\alpha} x^\alpha \in k[x_1, \ldots, x_n]$$

(observe that we are abusing notation here, but it is clear that ρ can be canonically extended to a ring map from $A[x_1, \dots, x_n]$ to $k[x_1, \dots, x_n]$ by leaving invariant the variables $x_1, \dots, x_n$).

Now, a natural question is to ask whether there exists a necessary and sufficient condition on the coefficients $u_{i,\alpha}$ such that the polynomials $\rho(f_1), \dots, \rho(f_r)$ in $k[x_1, \dots, x_n]$ have a common root in $\mathbb{P}^{n-1}_{\bar{k}}$, where $\bar{k}$ is the algebraic closure of the field k.

Theorem 6.1.1 (Elimination Theorem)**.** *Assuming the above notation, there exist polynomials $p_1, \dots, p_\ell \in A$ such that $\rho(p_1) = \cdots = \rho(p_\ell) = 0$ if and only if the polynomials $\rho(f_1), \dots, \rho(f_r)$ have a common root in $\mathbb{P}^{n-1}_{\bar{k}}$.*

Moreover, the ideal

$$\mathfrak{A}(I) := (I : \mathfrak{m}^\infty) \cap A = \{a \in A \ : \ \forall i \ \exists m \ : \ x_i^m a \in I\}$$

is such that $\rho(\mathfrak{A}(I)) = 0$ if and only if the polynomials $\rho(f_1), \dots, \rho(f_r)$ have a common root in $\mathbb{P}^{n-1}_{\bar{k}}$. It is called the resultant ideal *or* elimination ideal *associated to I with respect to $\mathfrak{m}$.*

Proof. We refer to [Eis95, Chapter 14] for a proof and more general versions. □

In comparison to Theorem 5.1.2, this theorem shows that the elimination of homogeneous variables, i.e. the projection of a projective variety, is closed (there is no need to take algebraic closure). However, in this projective context, it is necessary to saturate the ideal with respect to the homogeneous variables that are to be eliminated (we notice that a homogeneous ideal $J \subset k[x_1, \dots, x_n]$ and the ideals $\mathfrak{m}^p J$, $p \in \mathbb{N}$, define same projective varieties in $\mathbb{P}^{n-1}_k$).

Example 6.1.2. Assuming $r = n = 2$, we know that the Sylvester resultant $\mathrm{Res}(f_1, f_2)$ (which eliminates the variables x_1, x_2) is such that $\rho(\mathrm{Res}(f_1, f_2)) = 0$ if and only if $\rho(f_1)$ and $\rho(f_2)$ have a common root in $\mathbb{P}^1_{\bar{k}}$. Actually, in this case one can show that $\mathfrak{A}(I) = (\mathrm{Res}(f_1, f_2))$ (we will prove this later).

A quick geometric overview of the case $r = n$. In what follows we provide some geometric intuition (without proofs) about Theorem 6.1.1. We assume that $r = n$ in order to focus on what we will do in the next sections. We also assume that k is an algebraically closed field for simplicity.

First, we notice that homogeneous polynomials of degree d form an affine space by identifying $\sum_{|\alpha|=d} u_\alpha x^\alpha$ with the point $(u_\alpha)_{|\alpha|=d} \in \mathbb{A}^{N(d)}$, where $N(d) = \binom{n+d-1}{n-1}$.

Now, consider the incidence variety

$$W := V(f_1, \dots, f_r) \subset \mathbb{P}^{n-1} \times \prod_{i=1}^{n} \mathbb{A}^{N(d_i)}.$$

There are two canonical projections, namely π_1 onto the first factor $\mathbb{P}^{n-1}$ and π_2 on the second factor $\prod_{i=1}^n \mathbb{A}^{N(d_i)}$.

The projection $\pi_1 : W \to \mathbb{P}^{n-1}$ is surjective and its fibers are linear spaces of codimension n (show this!). We deduce that W is an irreducible variety of dimension $(\sum_{i=1}^n N(d_i)) - 1$ (W has actually a structure of fiber bundle).

Now, set $\nabla := \pi_2(W)$; this is an irreducible variety (it is a variety by the Elimination Theorem, and irreducible as the projection of an irreducible variety). As the general fiber over ∇ is a finite set of points (show this!), we deduce that ∇ has the same dimension as W. Therefore, ∇ is an irreducible hypersurface in $\prod_{i=1}^n \mathbb{A}^{N(d_i)}$. Such a hypersurface is defined by a single equation, which is called the resultant of $f_1, \ldots, f_n$ with respect to $x_1, \ldots, x_n$.

In what follows, we will provide a more algebraic treatment in order to provide a better definition of resultants (being an equation of a hypersurface is not satisfactory in many regards, as multiplicity or multiplicative constant).

6.2 Inertia Forms and Saturation

We take again the notation of Section 6.1, but now we let k be an arbitrary commutative ring (with unit) and we set $A := k[u_{i,\alpha}]$ (ring of all coefficients over k).

Definition 6.2.1. *The ideal* $(I : \mathfrak{m}^\infty) \subset C$ *is called the ideal of* inertia forms*. An element in* $(I : \mathfrak{m}^\infty)$ *is called an inertia form.*

We notice that $(I : \mathfrak{m}^\infty)$ is a graded ideal in $C = A[x_1, \ldots, x_n]$. Moreover, we have

$$\mathfrak{A}(I) = (I : \mathfrak{m}^\infty)_0 = (I : \mathfrak{m}^\infty) \cap A.$$

Proposition 6.2.2. $f \in (I : \mathfrak{m}^\infty)$ *if and only if there exist* $i \in \{1, \ldots, n\}$ *and* $m \in \mathbb{N}$ *such that* $x_i^m f \in I$.

Proof. Pick $i \in \{1, \ldots, n\}$ and for any $j = 1, \ldots, r$ set

$$f_j(x_1, \ldots, x_n) = \varepsilon_{i,j} x_i^{d_j} + \sum_{\substack{|\alpha| = d_j \\ u_{j,\alpha} \neq \varepsilon_{i,j}}} u_{j,\alpha} x^\alpha.$$

Thus, in the extended ring $C[x_i^{-1}]$ we have

$$f_j = x_i^{d_j} \left(\varepsilon_{i,j} + \sum_{\substack{|\alpha| = d_j \\ u_{j,\alpha} \neq \varepsilon_{i,j}}} u_{j,\alpha} \frac{x^\alpha}{x_i^{d_j}} \right) \in C[x_i^{-1}].$$

Setting $B = C/I$, we get an isomorphism of graded k-algebras

$$\begin{array}{rcl} B_{x_i} & \xrightarrow{\sim} & k[u_{j,\alpha} \ : \ u_{j,\alpha} \neq \varepsilon_{i,j}, j = 1, \ldots, r][x_1, \ldots, x_n][x_i^{-1}] \\ \varepsilon_{i,j} & \mapsto & - \displaystyle\sum_{\substack{|\alpha| = d_j \\ u_{j,\alpha} \neq \varepsilon_{i,j}}} u_{j,\alpha} \frac{x^\alpha}{x_i^{d_j}} \end{array}$$

(observe that taking quotient by f_j amounts to impose $\varepsilon_{i,j} = -\sum u_{j,\alpha} x^\alpha / x_i^{d_j}$, so the above claim follows from classical property of univariate polynomials, seeing here f_j as a polynomial in $\varepsilon_{i,j}$).

It follows that x_j is not a zero divisor in B_{x_i} for any couple of integers (i, j). So we have commutative diagrams

$$\begin{tikzcd}
C \arrow[r] \arrow[d] & B_{x_i} \arrow[d, hook] \\
B_{x_j} \arrow[r, hook] & B_{x_i x_j}
\end{tikzcd}$$

where maps are the canonical localization maps. Therefore, if $x_i^m f \in I$, i.e. $f = 0$ in B_{x_i}, then $f = 0$ in $B_{x_i x_j}$ and hence in B_{x_j}, which means that $x_j^{m'} f \in I$. □

Corollary 6.2.3. *If k is a domain then $(I : \mathfrak{m}^\infty)$, and hence $\mathfrak{A}(I)$, are prime ideals.*

Proof. In the proof of Proposition 6.2.2, we proved that $(I : \mathfrak{m}^\infty) = \operatorname{Ker}(C \to B_{x_n})$ and that B_{x_n} is a polynomial ring over k, hence is a domain. It follows that $C/(I : \mathfrak{m}^\infty)$ is a domain, and hence that $(I : \mathfrak{m}^\infty)$ is a prime ideal. The same conclusion follows for $\mathfrak{A}(I) = i^{-1}(I : \mathfrak{m}^\infty)$ where $i : A \hookrightarrow A[x_1, \ldots, x_n]$ is the canonical inclusion. □

Geometrically, following the discussion after Theorem 6.1.1, the homogeneous ideal $(I : \mathfrak{m}^\infty)$ defines the incidence variety W and $\mathfrak{A}(I)$ its image via π_2. We have just proved that they are both irreducible objects.

Theorem 6.2.4. *If $r < n$ then $(I : \mathfrak{m}^\infty) = I$.*

This theorem implies that the projection π_2 of the incidence variety W is surjective, since $\mathfrak{A}(I) = I \cap A = 0$ ($d_i \geq 1$ for all i). To prove this theorem, we will rely on the properties of regular sequences given in Exercise 3.5.3.

Lemma 6.2.5. *If $r \leq n$ then the sequence $\{f_1, \ldots, f_r\}$ is a regular sequence in C.*

Proof. In order to emphasize the coefficient of $x_i^{d_i}$ in f_i for all i, we set $f_i = \varepsilon_i x_i^{d_i} + \cdots$. We first consider the sequence of elements

$$S := \{\text{all } u_{1,\alpha} \text{ except } \varepsilon_1, \text{all } u_{2,\alpha} \text{ except } \varepsilon_2, \ldots, \text{all } u_{r,\alpha} \text{ except } \varepsilon_r\}.$$

It is obviously a regular sequence and in the quotient ring $C/(S)$, the class of f_i is equal to the class of $\varepsilon_i x_i^{d_i}$ for all i. Now, we add to S the elements

$$\varepsilon_1 - x_1, \varepsilon_2 - x_2, \ldots, \varepsilon_r - x_r.$$

The new sequence S' we obtain this way is still regular and moreover, in the quotient ring $C/(S') \simeq k[x_1, \ldots, x_n]$, the class of f_i is equal to the class of $x_i^{d_i+1}$ for all i. Finally, the sequence $x_1^{d_1+1}, \ldots, x_r^{d_r+1}$ is regular in $k[x_1, \ldots, x_n]$ (show this!), which implies that the sequence $S'' = S' \cup \{f_1, \ldots, f_r\}$ is a regular sequence in C.

To conclude, we observe that the elements in S'' are homogeneous elements so the sequence S'' remains regular after any permutation of its elements. In particular, the sequence $\{f_1, \dots, f_r\} \cup S'$ is regular. From the definition of regular sequences, we deduce that $\{f_1, \dots, f_r\}$ is a regular sequence in C. □

Proof of Theorem 6.2.4. Given $f \in (I : \mathfrak{m}^\infty)$, we have to show that $f \in I$. By assumption, there exists $s \in \mathbb{N}$ such that $x_n^s f \in I$ (use Proposition 6.2.2). If $s = 0$ then we are done. If not, it is enough to show that if $x_n f \in I$ then $f \in I$, because then, one can conclude by iterations ($x_n^s f = x_n(x_n^{s-1} f)$).

So, let $f \in (I : \mathfrak{m}^\infty)$ be such that $x_n f \in I$, i.e.

$$x_n f = h_1 f_1 + h_2 f_2 + \dots + h_r f_r \in C = A[x_1, \dots, x_n].$$

Specializing x_n to 0 in the above equality we get

$$0 = \bar{h}_1 \bar{f}_1 + \dots + \bar{h}_r \bar{f}_r \in A[x_1, \dots, x_{n-1}],$$

where we use the notation $\bar{p} := p(x_n = 0)$ for any polynomial $p \in C$. Since $r < n$, Lemma 6.2.5 implies that $\{\bar{f}_1, \dots, \bar{f}_r\}$ is a regular sequence. Therefore, applying Exercise 3.5.3, there exists a skew-symmetric matrix $M = (\lambda_{i,j})$ such that

$$\begin{pmatrix} \bar{h}_1 \\ \vdots \\ \bar{h}_r \end{pmatrix} = M \begin{pmatrix} \bar{f}_1 \\ \vdots \\ \bar{f}_r \end{pmatrix},$$

where $\lambda_{i,i} = 0$ for all i and $\lambda_{i,j} = -\lambda_{j,i}$ for all couple (i, j). Now, we define polynomials $g_1, \dots, g_r$ by setting

$$\begin{pmatrix} g_1 \\ \vdots \\ g_r \end{pmatrix} := M \begin{pmatrix} f_1 \\ \vdots \\ f_r \end{pmatrix},$$

i.e. $g_i = \sum_{j=1}^r \lambda_{i,j} f_j$ for all $i = 1, \dots, r$. As M is skew-symmetric, $\sum_{i=1}^r g_i f_i = 0$. Moreover, $\bar{g_i} = \bar{h_i}$ for all i, by definition of the g_i's. So, for all i there exists a polynomial l_i such that $-g_i + h_i = x_n l_i$ in $A[x_1, \dots, x_n]$. It follows that

$$\begin{aligned} x_n f &= (g_1 + x_n l_1) f_1 + (g_2 + x_n l_2) f_2 + \dots + (g_r + x_n l_r) f_r \\ &= (\sum_{i=1}^r g_i f_i) + x_n(\sum_{i=1}^r l_i f_i) = x_n(\sum_{i=1}^r l_i f_i). \end{aligned}$$

As x_n is not a zero divisor in C, we deduce that $f = \sum_{i=1}^r l_i f_i \in I$. □

6.3 Definition of Resultants

From now on we will focus on the case $r = n$. Observe that by Theorem 6.2.4, this is the first interesting case as the saturation of I is equal to I if $r < n$. Moreover,

$I \neq (I : \mathfrak{m}^\infty)$ if $r \geq n$. Indeed, consider the Jacobian determinant

$$\mathrm{Jac}(f_1, \dots, f_n) := \det \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}. \tag{6.3.1}$$

Then, $x_i \mathrm{Jac}(f_1, \dots, f_n) \in I$ for all i (multiply the first column of the above matrix by x_1, add to it a combination of the others in order to get the polynomials $d_i f_i$ is the first column, and use the Euler formulas $d_i f_i = \sum_{j=1}^n x_j \partial f_j / \partial_{x_j}$). So, $\mathrm{Jac}(f_1, \dots, f_n)$ belongs to $(I : \mathfrak{m}^\infty)$. One can also show that $\mathrm{Jac}(f_1, \dots, f_n)$ does not belong to I, although this is less obvious (see e.g. [Jou97]).

The goal of this section is to prove the following result.

Theorem 6.3.1. *Assume $r = n$ and k is a UFD, then $\mathfrak{A}(I)$ is a prime and principal ideal in A. It has a unique generator, which is denoted by $\mathrm{Res}(f_1, \dots, f_n)$, such that $\mathrm{Res}(x_1^{d_1}, \dots, x_n^{d_n}) = 1$.*

Before going further, let us explain the notation $\mathrm{Res}(f_1, \dots, f_n)$, and in the same time define the resultant of any sequence of homogeneous polynomials having the same degrees.

Let $g_1, \dots, g_n$ be n homogeneous polynomials in $S[x_1, \dots, x_n]$, where S is a commutative ring, such that g_i is of degree $d_i = \deg(f_i)$ for all integer i. More precisely, for all i

$$g_i(x_1, \dots, x_n) = \sum_{|\alpha| = d_i} c_{i,\alpha} x^\alpha \in S[x_1, \dots, x_n].$$

Consider the ring morphism (specialization map)

$$\begin{array}{rcl} \rho : \mathbb{Z}[u_{i,\alpha} : i = 1, \dots, n, |\alpha| = d_i] & \to & S \\ u_{i,\alpha} & \mapsto & c_{i,\alpha} \end{array}$$

(notice that $\rho(1) = 1$ necessarily).

Definition 6.3.2. *With the above notation, the resultant of the polynomials $g_1, \dots, g_n$ is the element in S, denoted $\mathrm{Res}(g_1, \dots, g_n)$, which is defined as*

$$\mathrm{Res}(g_1, \dots, g_n) := \rho(\mathrm{Res}(f_1, \dots, f_n)) \in S.$$

Resultants are hence defined by specialization from the generic case, i.e. the case where the coefficients of the polynomials are seen as variables over the integers. Thus, resultants can be seen as operators, similarly to determinants.

Finally, we notice that the notation $\mathrm{Res}(g_1, \dots, g_n)$ is compatible with the notation used in Theorem 6.3.1, considering the specialization map ρ as the identity map.

Macaulay determinants. A first step towards the proof of Theorem 6.3.1 is to show that $\mathfrak{A}(I) \neq 0$. For that purpose, we introduce the Macaulay determinants.

Set $\delta := \sum_{i=1}^{n}(d_i - 1)$ (recall that $d_i := \deg(f_i)$ for all $i = 1, \ldots, n$) and let $\mathrm{Mon}(t)$ be the set of all monomials of degree t, i.e.

$$\mathrm{Mon}(t) := \{x^\alpha \text{ such that } |\alpha| = t\}.$$

For all $t \geq \delta + 1$, any monomial $x^\alpha \in \mathrm{Mon}(t)$ is divisible by $x_i^{d_i}$ for some i (because $x_1^{d_1-1} \ldots x_n^{d_n-1}$ is of degree δ), so we can set the following definition: for any monomial $x^\alpha \in \mathrm{Mon}(t)$, $t \geq \delta + 1$, we define its index as

$$i(\alpha) := \min\{i \text{ such that } d_i \geq \alpha_i\}.$$

We notice that $i(\alpha)$ depends on the order of the variables $x_1, \ldots, x_n$.

Definition 6.3.3. *We define the* Macaulay matrix $\mathbb{M}(f_1, \ldots, f_n; t) = (m_{\alpha,\beta})_{\alpha,\beta}$ *by*

$$\begin{array}{rcl} \mathrm{Mon}(t) \times \mathrm{Mon}(t) & \to & A = k[u_{i,\alpha} \ : \ i = 1, \ldots, n, |\alpha| = d_i] \\ (\alpha, \beta) & \mapsto & m_{\alpha,\beta} \end{array}$$

where

$$\frac{x^\beta}{x_{i(\beta)}^{d_{i(\beta)}}} f_{i(\beta)} = \sum_{|\alpha|=t} m_{\alpha,\beta} x^\alpha$$

for all $x^\beta \in \mathrm{Mon}(t)$ (assume an order is chosen for $\mathrm{Mon}(t)$). In other words, the columns of $\mathbb{M}(f_1, \ldots, f_n; t)$ are built from the coefficients of the homogeneous polynomials $\frac{x^\beta}{x_{i(\beta)}^{d_{i(\beta)}}} f_{i(\beta)}$ with respect to $\mathrm{Mon}(t)$.

Example 6.3.4. If $r = n = 2$, the Macaulay matrix $\mathbb{M}(f_1, f_2; \delta + 1)$ in the canonical basis of $\mathrm{Mon}(\delta + 1)$ is nothing but the Sylvester matrix of f_1 and f_2 (check this!).

Proposition 6.3.5. *The determinant $D(f_1, \ldots, f_n; t) := \det(\mathbb{M}(f_1, \ldots, f_n; t))$, $t \geq \delta + 1$, belongs to the ideal $\mathfrak{A}(I)$. Moreover, $D(f_1, \ldots, f_n; t)$ is nonzero and is homogeneous with respect to the coefficients of each polynomial f_i.*

Proof. For the first assertion, suppose that the order chosen for $\mathrm{Mon}(t)$ is the lex order $x_1 > \ldots > x_n$. Then, multiply the first row of $\mathbb{M}$ by x_1^t and then add to it a combination of the other rows in order get multiples of the f_i's in the first row. It follows that $x_1^t D(f_1, \ldots, f_n; t) \in (I : \mathfrak{m}^\infty)$ (apply Proposition 6.2.2 or repeat this process for all monomial $x^\alpha \in \mathrm{Mon}(t)$).

To prove that $D(f_1, \ldots, f_n; t)$ is nonzero, we specialize each f_i to $x_i^{d_i}$. The matrix $\mathbb{M}(f_1, \ldots, f_n; t)$ specializes to the matrix $\mathbb{M}(x_1^{d_1}, \ldots, x_n^{d_n}; t)$ which is nothing but the identity matrix (assuming that the same basis is chosen for the row and columns of $\mathbb{M}$), whose determinant is nonzero. We deduce that $D(f_1, \ldots, f_n; t)$ is nonzero (if it is identically 0 then any of its specialization must be zero as well).

Finally, the claimed homogeneity property is a straightforward consequence of the definition of $\mathbb{M}(f_1, \ldots, f_n; t)$ and the expansion of determinants with respect to columns. □

Proposition 6.3.6. *For any $t \geq \delta + 1$, the degree of $D(f_1, \ldots, f_n; t)$ with respect to the coefficients of f_n is equal to $d_1 d_2 \ldots d_{n-1}$.*

Proof. By definition of the Macaulay determinant $D(f_1, \ldots, f_n; t)$, its degree with respect to the coefficients of f_j is simply the number of monomials x_α such that $|\alpha| = t$ and $i(\alpha) = j$. Now, $i(\alpha) = n$ if and only if $0 \leq \alpha_i \leq d_i - 1$ for all $i = 1, \ldots, n-1$. □

Corollary 6.3.7. *Let $t \geq \delta + 1$. For all $i = 1, \ldots, n$, there exists a nonzero Macaulay determinant $D_i(t) \in \mathfrak{A}(I)$ which is homogeneous with respect to the coefficients of f_i of degree $d_1 d_2 \ldots d_n / d_i$.*

Proof. $D_n(t)$ is the determinant $D(f_1, \ldots, f_n; t)$ as defined in Proposition 6.3.5. To obtain the others, simply permute the order of the polynomials f_i's and variables x_i's in the previous constructions. □

Back to the proof of Theorem 6.3.1. We are almost ready to prove the defining property of resultants. We need a last result on inertia forms. We maintain the notation of the previous sections.

Proposition 6.3.8. *Let $f \in (I : \mathfrak{m}^\infty)$, then either $f \in I$ (in which case f is called a trivial inertia form) or f depends on each coefficient of each polynomial f_i, $i = 1, \ldots, n$.*

Proof. Let u be one of the coefficients $u_{i,\alpha}$ of the polynomial f_i, for some i; we have $f_i = ux^\alpha + g_i$.

Suppose that there exits $f \in (I : \mathfrak{m}^\infty)$ such that f is independent on u. By assumption, there exists m such that (apply Proposition 6.2.2)

$$x_n^m f = h_1 f_1 + h_2 f_2 + \cdots + h_n f_n, \;\; h_i \in C.$$

Consider the morphism of k-algebras

$$\begin{array}{rcl} \varphi : A[x_1, \ldots, x_n] & \to & A[x_1, \ldots, x_n]_{x_1 x_2 \ldots x_n} \\ u & \mapsto & -g_i / x^\alpha \\ u_{j,\beta} & \mapsto & u_{j,\beta}, \;\; (j, \beta) \neq (i, \alpha) \\ x_i & \mapsto & x_i. \end{array}$$

Since f is independent on u, one has $\varphi(x_n^m f) = x_n^m f$. Therefore, since $\varphi(f_i) = 0$ we get

$$x_n^m f = \varphi(h_1) f_1 + \cdots + \varphi(h_{i-1}) f_{i-1} + \varphi(h_{i+1}) f_{i+1} + \cdots + \varphi(h_n) f_n.$$

But $x_1 x_2 \ldots x_n$ is not a zero divisor in $C = A[x_1, \ldots, x_n]$, so there exists a monomial x^β such that

$$x^\beta x_n^m f = l_1 f_1 + l_2 f_2 + \cdots + l_{i-1} f_{i-1} + l_{i+1} f_{i+1} + \cdots + l_n f_n$$

in C. Therefore $f \in ((f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_n) : \mathfrak{m}^\infty)$ and hence , applying Theorem 6.2.4 we deduce that $f \in (f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_n)$. □

Proof of Theorem 6.3.1. We select a coefficient u of one of the polynomial f_i, $i = 1, \dots, n$ ($u = u_{i,\alpha}$ for some i and α). Let A' be the ring of coefficients without u, i.e. $A = A'[u]$. We notice that, similarly to A, A' is a UFD as k is assumed to be a UFD. We know that

- $\mathfrak{A}(I) \neq 0$ by Proposition 6.3.5,
- for all nonzero $a \in \mathfrak{A}(I)$, a is of positive degree as a polynomial in the variable u, a property that we summarize by $\deg_u(a) \geq 1$. This follows from Proposition 6.3.8 (observe that $\mathfrak{A}(I) \cap I = 0$ as $d_i \geq 1$ for all i).

Therefore, the integer s defined as

$$s := \inf_{a \in \mathfrak{A}(I), a \neq 0} \deg_u(a)$$

is a positive integer: $s \geq 1$.

Our first step is to show that there exists a prime element $R \in \mathfrak{A}(I)$ such that $\deg_u(R) = s$. Let a be a nonzero element in $\mathfrak{A}(I)$ such that $\deg_u(a) = s$. One can decompose a as a product $a = q_1 q_2 \dots q_t$ where the q_i's are primes in A (which is a UFD). But $\mathfrak{A}(I)$ is a prime ideal by Corollary 6.2.3, so $q_i \in \mathfrak{A}(I)$ for some i. Moreover $\deg_u(q_i) \leq \deg_u(a) = s$, so from the definition of s we deduce that $\deg_u(q_i) = s$. Thus, we set $R := q_i$ and the expected property is proved.

Our second step is to prove that R is a generator of $\mathfrak{A}(I)$. For any $b \in \mathfrak{A}(I)$, the pseudo-euclidian division by R as a polynomial in u, i.e. in $A'[u]$, yields the equality

$$\lambda b = qR + v$$

where $\lambda \in A'$ and $v \in A'[u]$ is such that $v = 0$ or $\deg_u(v) < s$. We notice that $v = \lambda b - qR \in \mathfrak{A}(I)$. Therefore, if $v \neq 0$ then there is a contradiction with the definition of s. It follows that $v = 0$ and hence that $\lambda b = qR$ in $A'[u]$. Now, as R is irreducible, it must divides λ or b. But λ does not depend on u so we deduce that R divides b, proving the claimed property.

To conclude, we proved that $\mathfrak{A}(I)$ is generated by R, which is defined up to an invertible element in A', hence in k. This invertible element is set by the condition $\mathrm{Res}(x_1^{d_1}, \dots, x_n^{d_n}) = 1$. □

6.4 Formal Properties of Resultants

In this section, we give some properties of the resultants. Those properties are important ingredients to understand the deep geometric meaning of resultants, but also to compute with them efficiently, as an operator. As we will see, a key property to develop the formalism of resultants is their stability under specialization, which follows by Definition 6.3.2.

Resultant of linear forms. Let R be a commutative ring and $l_1, \dots, l_n$ be n linear forms in $R[x_1, \dots, x_n]$:

$$l_i(x_1, \dots, x_n) = \sum_{j=1}^{n} u_{i,j} x_j, \ i = 1, \dots, n.$$

Then

$$\operatorname{Res}(l_1, \ldots, l_n) = \det(u_{i,j})_{1 \le i,j \le n}.$$

Proof. By Definition 6.3.2, it is enough to prove this formula in the generic setting over the integers, i.e. assuming that $R = \mathbb{Z}[u_{i,j} \ : \ 1 \le i, j \le n]$.

By a trick we already used several times, we have $x_i \det(u_{i,j}) \in (l_1, \ldots, l_n)$, so $\det(u_{i,j}) \in \mathfrak{A}(I)$. It is also nonzero (e.g. one can specialize it to the identity matrix). Applying Theorem 6.3.1, $\mathfrak{A}(I)$ is generated by $\operatorname{Res}(l_1, \ldots, l_n)$, so $\operatorname{Res}(l_1, \ldots, l_n)$ divides $\det(u_{i,j})$. The determinant $\det(u_{i,j})$ is homogeneous of degree 1 in the coefficients of each l_i, $i = 1, \ldots, n$, by construction. But $\operatorname{Res}(l_1, \ldots, l_n)$ is homogeneous of degree at least 1 in the coefficients of each of the l_i's by Proposition 6.3.8, so we deduce that there exists a nonzero integer c such that $\det(u_{i,j}) = c\operatorname{Res}(l_1, \ldots, l_n)$ (we notice that we could have used Macaulay determinants, in particular Proposition 6.3.5, to reach this conclusion). By specializing each l_i to x_i, we get $c = 1$, which concludes the proof. □

Divisibility. Let R be a commutative ring and $f_1, \ldots, f_n$ and $g_1, \ldots, g_n$ be two sequences of homogeneous polynomials in $R[x_1, \ldots, x_n]$ such that $g_i \in (f_1, \ldots, f_n)$ for all $i = 1, \ldots, n$. Then

$$\operatorname{Res}(f_1, \ldots, f_n) \text{ divides } \operatorname{Res}(g_1, \ldots, g_n) \text{ in } R.$$

Proof. By assumption, $g_i = \sum_j h_{i,j} f_j$ for all i. Thanks to the specialization property of resultants, one can assume that the f_i's and the $h_{i,j}$'s are generic homogeneous polynomials, and that R is their universal ring of coefficients over the integers.

Now, there exists an integer m such that $x_n^m \operatorname{Res}(g_1, \ldots, g_n) \in (g_1, \ldots, g_n)$, because $\operatorname{Res}(g_1, \ldots, g_n)$ is an inertia forms (a property that is stable under specialization). It follows that $x_n^m \operatorname{Res}(g_1, \ldots, g_n) \in (f_1, \ldots, f_n)$ and hence that $\operatorname{Res}(g_1, \ldots, g_n)$ is an inertia form of $(f_1, \ldots, f_n)$. As the f_i's are generic polynomials, Theorem 6.3.1 implies that $\operatorname{Res}(f_1, \ldots, f_n)$ divides $\operatorname{Res}(g_1, \ldots, g_n)$, which concludes the proof. □

Multi-degree of resultants. Assuming we are in the generic setting, $\operatorname{Res}(f_1, \ldots, f_n)$ is an homogeneous polynomial in the coefficients of f_i of degree $d_1 d_2 \ldots d_n / d_i$ for all $i = 1, \ldots, n$, where d_i denotes the degree of the homogeneous polynomial f_i.

Proof. Theorem 6.3.1 and the existence of Macaulay determinants, more precisely Corollary 6.3.7, show that

$$\deg_{f_i} \operatorname{Res}(f_1, \ldots, f_n) \le \frac{d_1 d_2 \ldots, d_n}{d_i}, \ i = 1, \ldots, n$$

(the notation $\deg_{f_i}(-)$ means the degree with respect to the coefficients of f_i).

Now, consider the specialization that sends each f_i to a product of generic linear forms $l_{i,j}$:

$$\rho : f_i \mapsto g_i := \prod_{j=1}^{d_i} l_{i,j}, \ i = 1, \ldots, n.$$

By the divisibility property, $\mathrm{Res}(l_{1,j_1},\dots,l_{n,j_n})$ divides $\mathrm{Res}(g_1,\dots,g_n)$ for all $j_1,\dots,j_n$, so we deduce that

$$\prod_{1\leq j_1\leq d_1,\dots,1\leq j_n\leq d_n} \mathrm{Res}(l_{1,j_1},\dots,l_{n,j_n}) \text{ divides } \mathrm{Res}(g_1,\dots,g_n). \tag{6.4.1}$$

Moreover, $\deg_{l_{i,j_k}}\mathrm{Res}(l_{1,j_1},\dots,l_{n,j_n})=1$ for all pairs i,j_k. We have $d_1d_2\dots d_n$ terms in the product (6.4.1). In addition, the coefficients of f_i are specialized to homogeneous polynomials of degree d_i in the coefficients of the l_{i,j_k}'s via ρ. Therefore, we conclude that

$$\deg_{f_i}\mathrm{Res}(f_1,\dots,f_n)\geq \frac{d_1d_2\dots,d_n}{d_i},\ i=1,\dots,n.$$

□

Exercise 6.4.1. In the case $n=2$, show that $\mathrm{Res}(f_1,f_2)$ is equal to the Sylvester resultant defined in Section 5.3 (hint: show first that the determinant of the Sylvester matrix is an inertia form of the expected degree with respect to the coefficients of f_1 and of f_2, and then conclude with a well chosen specialization).

Multiplicativity property. Let R be a commutative ring and let $f_1,\dots,f_{i-1},f'_i$, $f''_i,f_{i+1},\dots,f_n$ be $n+1$ homogeneous polynomials in $R[x_1,\dots,x_n]$ of positive degree. Then,

$$\begin{aligned}&\mathrm{Res}(f_1,\dots,f_{i-1},f'_if''_i,f_{i+1},\dots,f_n)=\\ &\qquad \mathrm{Res}(f_1,\dots,f_{i-1},f'_i,f_{i+1},\dots,f_n)\mathrm{Res}(f_1,\dots,f_{i-1},f''_i,f_{i+1},\dots,f_n).\end{aligned}$$

Proof. By the specialization property of resultants, it is enough to prove the claimed formula assuming that R is the universal ring of coefficients of the polynomials $f_1,\dots,f_{i-1},f'_i,\ f''_i,f_{i+1},\dots,f_n$.

Set $f_i:=f'_if''_i$. By the divisibility property, both $\mathcal{R}':=\mathrm{Res}(f_1,\dots,f'_i,\dots,f_n)$ and $\mathcal{R}'':=\mathrm{Res}(f_1,\dots,f''_i,\dots,f_n)$ divides $\mathcal{R}:=\mathrm{Res}(f_1,\dots,f_i,\dots,f_n)$. As $\mathcal{R}'$ and $\mathcal{R}''$ are irreducible (because their input polynomials are generic polynomials) and coprime (as they do not depends on all the same coefficients), we deduce that $\mathcal{R}'\mathcal{R}''$ divides $\mathcal{R}$. Now, set $d'_i:=\deg(f'_i)$ and $d''_i:=\deg(f''_i)$ and $d_i:=d'_i+d''_i=\deg(f_i)$. Computing the multi-degrees of resultants, it is clear that for all $j\neq i$ we have that $\deg_{f_j}(\mathcal{R}'\mathcal{R}'')=\deg_{f_j}(\mathcal{R})$. Moreover,

$$\deg_{f'_i}(\mathcal{R}')=\deg_{f''_i}(\mathcal{R}'')=d_1\dots d_{i-1}d_{i+1}\dots,d_n,$$

so $\mathcal{R}''\mathcal{R}''$ is homogeneous of degree $2d_1\dots d_{i-1}d_{i+1}\dots,d_n$ with respect to the coefficients of f'_i and f''_i. But $\deg_{f_i}(\mathcal{R})=d_1\dots d_{i-1}d_{i+1}\dots,d_n$, and since $f_i=f'_if''_i$, the coefficients of f_i are degree 2 polynomials in the coefficients of f'_i and f''_i. All this shows that $\mathcal{R}$ and $\mathcal{R}'\mathcal{R}''$ have the same degrees with respect to the coefficients of $f_1,\dots,f_{i-1},f'_i,\ f''_i,f_{i+1},\dots,f_n$. Therefore, there exists a nonzero integer c such that $\mathcal{R}=c\mathcal{R}'\mathcal{R}''$. By specializing each f_j to $x_j^{d_j}$, f'_i to $x_i^{d'_i}$ and f''_i to $x_i^{d''_i}$, we conclude that $c=1$. □

Permutation of variables. Let σ a permutation of the group of n elements and let $f_1, \dots, f_n$ be homogeneous polynomials of positive degree in $R[x_1, \dots, x_n]$, with R a commutative ring. Then,

$$\operatorname{Res}(f_{\sigma(1)}, f_{\sigma(2)}, \dots, f_{\sigma(n)}) = \varepsilon(\sigma)^{d_1 d_2 \dots d_n} \operatorname{Res}(f_1, \dots, f_n),$$

where $\varepsilon(\sigma)$ denotes the signature of the permutation σ.

Proof. First, observe that since resultants of linear forms are simply determinants of coefficient matrices (first property above), we have

$$\operatorname{Res}(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma).$$

Now, as always it is enough to prove the claimed equality in the generic setting. By the divisibility property, both resultants $\operatorname{Res}(f_{\sigma(1)}, f_{\sigma(2)}, \dots, f_{\sigma(n)})$ and $\operatorname{Res}(f_1, \dots, f_n)$ divides each others and hence there exists an inverstible element in $\mathbb{Z}$, i.e. $c = \pm 1$, such that

$$\operatorname{Res}(f_{\sigma(1)}, f_{\sigma(2)}, \dots, f_{\sigma(n)}) = c \operatorname{Res}(f_1, \dots, f_n).$$

To determine c, we specialize each f_i to $x_i^{d_i}$ for all i. On the one hand, we get

$$\operatorname{Res}(x_{\sigma(1)}^{d_{\sigma(1)}}, x_{\sigma(2)}^{d_{\sigma(2)}}, \dots, x_{\sigma(n)}^{d_{\sigma(n)}}) = c \operatorname{Res}(x_1^{d_1}, \dots, x_n^{d_n}) = c,$$

the last equality following from the normalization of the resultant. On the other hand, by the multiplicativity property we have

$$\operatorname{Res}(x_{\sigma(1)}^{d_{\sigma(1)}}, x_{\sigma(2)}^{d_{\sigma(2)}}, \dots, x_{\sigma(n)}^{d_{\sigma(n)}}) = \operatorname{Res}(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)})^{d_1 d_2 \dots d_n} = \varepsilon(\sigma)^{d_1 d_2 \dots d_n}.$$

□

Invariance under elementary transformations. Let R be a commutative ring and $f_1, \dots, f_n$ be homogeneous polynomials of positive degree in $R[x_1, \dots, x_n]$. Then,

$$\operatorname{Res}(f_1, \dots, f_{i-1}, f_i + \sum_{j \neq i} h_{i,j} f_j, f_{i+1}, \dots, f_n) = \operatorname{Res}(f_1, \dots, f_{i-1}, f_i, f_{i+1}, \dots, f_n)$$

for any $i \in \{1, \dots, n\}$ and for any collection of homogeneous polynomials $h_{i,j}$ such that $f_i + \sum_{j \neq i} h_{i,j} f_j$ is homogeneous of the same degree as f_i.

Proof. As always, one can assume that we are in the generic setting (coefficients of the f_i's and $h_{i,j}$'s). Now, by the divisibility property, these two resultants divide each others, so they are equal up to an invertible multiplicative constant in $\mathbb{Z}$. This constant is proved to be equal to one be specializing all $h_{i,j}$ to 0. □

Reduction by one variable. Let R be a commutative ring and let $f_1, \dots, f_{n-1}$ be homogeneous polynomials of positive degree in $R[x_1, \dots, x_n]$, $n \geq 2$. We set $\bar{f}_i(x_1, \dots, x_{n-1}) := f_i(x_1, \dots, x_{n-1}, 0) \in R[x_1, \dots, x_{n-1}]$. Then,

$$\operatorname{Res}(f_1, \dots, f_{n-1}, x_n) = \operatorname{Res}(\bar{f}_1, \dots, \bar{f}_{n-1}) \in R.$$

Proof. One can assume that we are in the generic setting, R being the universal ring of coefficients of the f_i's. As resultants are inertia forms, there exists an integer m such that $x_1^m \mathrm{Res}(f_1, \ldots, f_{n-1}, x_n) \in (f_1, \ldots, f_{n-1}, x_n)$ (the property of being an inertia form is stable under specialization). Specializing x_n to 0 we deduce that $x_1^m \mathrm{Res}(f_1, \ldots, f_{n-1}, x_n) \in (\bar{f}_1, \ldots, \bar{f}_{n-1})$. As the $\bar{f}_i$'s are generic polynomials, it follows that $\mathrm{Res}(\bar{f}_1, \ldots, \bar{f}_{n-1})$ divides $\mathrm{Res}(f_1, \ldots, f_{n-1}, x_n)$. From here, we conclude by comparing the multi-degrees of these two resultants and by considering the specialization sending each f_i to $x_i^{d_i}$. □

The base change formula. Let R be a commutative ring and $f_1, \ldots, f_n$ be n homogeneous polynomials in $R[x_1, \ldots, x_n]$ of positive degrees $d_1, \ldots, d_n \geq 1$ respectively. Moreover, suppose given n homogeneous polynomials $g := (g_1, \ldots, g_n)$ in $R[x_1, \ldots, x_n]$ of the same degree $d \geq 1$. Then,

$$\mathrm{Res}(f_1 \circ g, \ldots, f_n \circ g) = \mathrm{Res}(g_1, \ldots, g_n)^{d_1 d_2 \ldots d_n} \mathrm{Res}(f_1, \ldots, f_n)^{d^{n-1}}.$$

Exercise-Proof.

1. Justify that it is enough to prove the above formula over a universal ring of coefficients. Describe this ring.
2. Show that there exists an integer m such that for all $i = 1, \ldots, n$

$$g_i^m \mathrm{Res}(f_1, \ldots, f_n) \in (f_1 \circ g, \ldots, f_n \circ g).$$

3. Deduce that

$$\mathrm{Res}(f_1 \circ g, \ldots, f_n \circ g) = \varepsilon \mathrm{Res}(g_1, \ldots, g_n)^{\lambda} \mathrm{Res}(f_1, \ldots, f_n)^{\mu}$$

with λ, μ positive integers and $\varepsilon = \pm 1$.

4. Conclude the proof with the help of the specialization $f_j \mapsto u_j x_j^{d_j}$, $g_j \mapsto v_j x_j^d$ for all j.

Solution.

1. By definition, the resultant is a universal object: it is first defined in the universal setting and then defined over any commutative by specialization (there is always a ring map from $\mathbb{Z}$ to any commutative ring). In our setting, the universal ring A is the polynomial ring built from the coefficients of the f_i's and g_j's over the ring integers.
2. The resultant is an inertia form, so there exists N such that for all $i = 1, \ldots, n$ we have

$$x_i^N \mathrm{Res}(f_1, \ldots, f_n) \in (f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n)).$$

Substituting x_i by $g_i(x_1, \ldots, x_n)$ in the above equality gives the claimed relation (notice that the resultant belongs to A and hence does not depend on the x_i's).

3. Using the relations obtained in the previous question and the divisibility property of the resultant, we get that

$$\mathrm{Res}(f_1 \circ g, \ldots, f_n \circ g) \text{ divides } \mathrm{Res}(g_1^N \mathrm{Res}(f_1, \ldots, f_n), \ldots, g_n^N \mathrm{Res}(f_1, \ldots, f_n))$$

   in A. But since $\mathrm{Res}(f_1, \ldots, f_n) \in A$, by homogeneity and multiplicativity of the resultant we have

$$\begin{aligned}\mathrm{Res}(g_1^N \mathrm{Res}(f_1, \ldots, f_n), \ldots, g_n^N \mathrm{Res}(f_1, \ldots, f_n)) = \\ \mathrm{Res}(f_1, \ldots, f_n)^{n(Nd)^{n-1}} \mathrm{Res}(g_1^N, \ldots, g_n^N) = \\ \mathrm{Res}(f_1, \ldots, f_n)^{n(Nd)^{n-1}} \mathrm{Res}(g_1, \ldots, g_n)^{N^n}.\end{aligned}$$

   Now, since we are in the universal setting, over A, the resultants $\mathrm{Res}(f_1, \ldots, f_n)$ and $\mathrm{Res}(g_1, \ldots, g_n)$ are both irreducible polynomials that are moreover coprime (they do not depend on the same variables). It follows that

$$\mathrm{Res}(f_1 \circ g, \ldots, f_n \circ g) = \varepsilon \mathrm{Res}(g_1, \ldots, g_n)^\lambda \mathrm{Res}(f_1, \ldots, f_n)^\mu \tag{6.4.2}$$

   for some non negative integers λ, μ and an invertible element ε in $\mathbb{Z}$.
4. Using this specialization, (6.4.2) yields the equality

$$\mathrm{Res}(u_1 v_1^{d_1} x_1^{dd_1}, \ldots, u_n v_n^{d_n} x_n^{dd_n}) = \varepsilon \mathrm{Res}(v_1 x_1^{d_1}, \ldots, v_n x_n^{d_n})^\lambda \mathrm{Res}(u_1 x_1^{d_1, \ldots, u_n x_n^{d_n}})^\mu.$$

   Applying the homogeneity and multiplicativity properties of the resultant we get

$$\prod_i \left(u_i v_i^{d_i}\right)^{d^{n-1} \frac{d_1 \ldots d_n}{d_i}} = \varepsilon \left(\prod_i v_i^{d^{n-1}}\right)^\lambda \left(\prod_i u_i^{\frac{d_1 \ldots d_n}{d_i}}\right)^\mu,$$

   so we deduce that $\varepsilon = 1$, $\mu = d^{n-1}$ and $\lambda = d_1 \ldots d_n$.
5. If the g_i's are linear forms $g_i = \sum_{i=1}^n a_{i,j} x_j$ then we know that

$$\mathrm{Res}(g_1, \ldots, g_n) = \det(a_{i,j})_{i,j=1,\ldots,n}.$$

   In this case, we get

$$\mathrm{Res}(f_1 \circ g, \ldots, f_n \circ g) = \det(a_{i,j})^{d_1 d_2 \ldots d_n} \mathrm{Res}(f_1, \ldots, f_n).$$

   The resultant is said to be invariant under a linear change of coordinates.

□

A consequence of the above property is the **invariance of the resultant under the action of the general linear group** GL(n). Indeed, assume that $d = 1$, i.e. that the g_i's are linear forms and let φ be their matrix of coefficients, then

$$\mathrm{Res}(f_1 \circ g, \ldots, f_n \circ g) = \det(\varphi)^{d_1 \ldots d_n} \mathrm{Res}(f_1, \ldots, f_n).$$

The Macaulay formula. We take again the notation we introduced to state Proposition 6.3.6 on Macaulay determinants. Recall that for any integer $t \geq \delta + 1$, where $\delta = \sum_{i=1}^{n}(d_i - 1)$, and for any $i \in \{1, \ldots, n\}$, we defined the Macaulay determinant $D(f_1, \ldots, f_n; t) = \det(\mathbb{M}(f_1, \ldots, f_n; t)) \in A = k[u_{i,\alpha}]$ (the ring of coefficients over the commutative ring k) and proved that its degree with respect to the coefficients of the polynomial f_n is equal to $d_1 d_2 \ldots d_{n-1}$. Observe that this degree is precisely the degree of the resultant of $f_1, \ldots, f_n$ with respect to the coefficients of f_n. Therefore, there exists a polynomial $H(f_1, \ldots, f_{n-1}, d_n; t)$, which is independent of the coefficients of f_n and homogeneous with respect to the coefficients of each polynomial $f_1, \ldots, f_{n-1}$, such that

$$D(f_1, \ldots, f_n; t) = \mathrm{Res}(f_1, \ldots, f_n) H(f_1, \ldots, f_{n-1}, d_n; t).$$

It turns out that the polynomial H can be obtained as the determinant of a certain submatrix of the Macaulay matrix $\mathbb{M}(f_1, \ldots, f_n; t)$. More precisely, for all $t \geq \delta + 1$ define the set of monomials

$$\mathrm{Dod}(t) := \{x^\alpha \text{ such that } \exists i \neq j \ : \ \alpha_i \geq d_i \text{ and } \alpha_j \geq d_j\} \subset \mathrm{Mon}(t)$$

and let $\mathbb{H}(f_1, \ldots, f_{n-1}, d_n; t)$ be the submatrix of $\mathbb{M}(f_1, \ldots, f_n; t)$ whose rows and columns are indexed by $\mathrm{Dod}(t)$.

Theorem 6.4.2 (Macaulay Formula)**.** *For all $t \geq \delta + 1$,*

$$\det(\mathbb{M}(f_1, \ldots, f_n; t)) = \mathrm{Res}(f_1, \ldots, f_n) \det(\mathbb{H}(f_1, \ldots, f_{n-1}, d_n; t)).$$

For the proof of this result, we refer to [Jou97, Proposition 3.9.4.4]; see also [CLO98, Chapter 3, §4] for more comments. It is important to notice that this formula holds over the universal ring of coefficients of the f_i's over the integers, and hence remains valid through any specialization. Thus, it provides a very useful formula to compute resultants.

Exercise 6.4.3. Write down explicitly the two matrices of the above Macaulay formula in the case $n = 3$, $d_1 = d_2 = 1$ and $d_2 = 2$.

6.5 Some Applications

Resultants has many, many applications. In this section, we briefly overview three of them of different flavors. The topics we cover are the implicitization of parameterized surfaces in $\mathbb{P}^3$, some old-fashioned and classical theorems in Euclidian plane geometry and the definition of the discriminant of a hypersurface. The application in plane geometry relies on Poisson's formula, a celebrated result that opens the door to numerous other applications for solving polynomial systems by means of resultants (we refer to [CLO98, Chapter 3,§4 and §6]) as a first reading on this topic).

6.5.1 Implicitization of parameterized surfaces in $\mathbb{P}^3$

Suppose given a parameterization of an algebraic surface $\mathcal{S}$ in $\mathbb{P}^3_k$, k being an algebraically closed field:

$$\begin{array}{rcl} \phi : \mathbb{P}^2 & \to & \mathbb{P}^3 \\ (s:t:u) & \mapsto & (f_0(s,t,u) : f_1(s,t,u) : f_2(s,t,u) : f_3(s,t,u)) \end{array}$$

where the f_i's are homogeneous polynomials in $k[s,t,u]$ of the same degree $d \geq 1$. For simplicity here, we also assume that $V(f_0, \ldots, f_3) \subset \mathbb{P}^2$ is empty, so that ϕ is well defined and $\mathrm{Im}(\phi) = \mathcal{S}$.

Irreducibility. The algebraic counterpart of ϕ is the map of k-algebras

$$\begin{array}{rcl} h : k[x_0, x_1, x_2, x_3] & \to & k[s,t,u] \\ x_i & \mapsto & f_i, \ i = 0, \ldots, 3 \end{array}$$

where $x_0, \ldots, x_3$ denote the coordinates in $\mathbb{P}^3$. The kernel of h contains all polynomials in $k[x_0, \ldots, x_3]$ that vanish on $\mathrm{Im}(\phi) = \mathcal{S}$. It is the defining ideal of $\mathcal{S}$, and it is a prime ideal as $k[s,t,u]$ is a domain. Therefore, the surface $\mathcal{S}$ is an irreducible surface.

Degree formula. Before trying to compute a defining equation of $\mathcal{S}$, it is better to have an idea of how big this equation could be. A first estimation is to guess the degree of $\mathcal{S}$.

The degree of $\mathcal{S}$ is the number of intersection points of $\mathcal{S}$ with a general line in $\mathbb{P}^3$. Such a line is the intersection of two hyperplanes, i.e. two linear forms in x_0, x_1, x_2, x_3: $\sum_{i=0}^3 a_i x_i$ and $\sum_{i=0}^3 b_i x_i$. Similarly to what we did for plane curves (see Section 5.5), to count the intersection points we pullback them via ϕ: we get the two equations $\sum_{i=0}^3 a_i f_i(s,t,u) = 0$ and $\sum_{i=0}^3 b_i f_i(s,t,u) = 0$ that correspond to two plane curves of degree d in $\mathbb{P}^2$. By Bézout theorem (see Theorem 2.1.1), they intersect in d^2 points. So, if the map ϕ is generically injective onto $\mathcal{S}$, then we deduce that $\deg(\mathcal{S}) = d^2$. If the map ϕ is not generically injective, we need to consider the degree of ϕ, which is defined as the number of pre-images of a general point on $\mathcal{S}$ (similarly to the case of plane curves again). Therefore, we get the formula

$$\deg(\phi)\deg(\mathcal{S}) = d^2,$$

assuming, as we did, that the f_i's have not common root in $\mathbb{P}^2$.

Before moving on, let us mention the more general case where the f_i's have common roots, which are called *base points* of the map ϕ. Without loss in generality, one can assume that the f_i's have only finitely many common points because otherwise they must share a common factor that can be easily removed. Now, to each common point $p \in V(f_0, \ldots, f_3)$ one can attached a multiplicity, called the Hilbert-Samuel multiplicity, that we denote by e_p. Then, we have the following degree formula:

$$\deg(\phi)\deg(\mathcal{S}) = d^2 - \sum_{p \in V(f_0, \ldots, f_3)} e_p.$$

This formula shows in particular that base points are unavoidable to parameterize (with $\mathbb{P}^2$) a lot of surfaces in $\mathbb{P}^3$, typically surfaces whose degree is not a square number.

Implicitization formula. Similarly to the result we obtained in the case of plane curves (see Section 5.5), it is no surprise that the resultant yields an implicit equation of $\mathcal{S}$ in the absence of base points.

Proposition 6.5.1. *Assume that $V(f_0, \ldots, f_3) = \emptyset$ then*

$$\mathrm{Res}(f_1 - x_1 f_0, f_2 - x_2 f_0, f_3 - x_3 f_0) = H(1, x_1, x_2, x_3)^{\deg(\phi)}$$

where $H(x_0, x_1, x_2, x_3)$ is a defining equation of the surface $\mathcal{S}$, which is of degree $d^2/\deg(\phi)$.

Proof. The fact that the resultant (which is taken with respect to the variables s, t, u) vanishes if and only if H vanishes for all points such that $x_0 \neq 0$ is easy to see. It remains to adjust the powers on these equations; we admit this result. □

To compute the resultant in the above proposition, the Macaulay formula can be used. However, in this setting (3 homogeneous polynomials of the same degree in three variables) there exists a more compact formula yielding the resultant as the determinant of a square matrix. Below, we provide this formula as an exercise.

Exercise 6.5.2. Suppose given an integer $d \geq 2$ and 3 generic homogeneous polyomials of degree d in the variables $x = (x_1, x_2, x_3)$:

$$f_1 = \sum_{|\alpha|=d} u_{1,\alpha} x^\alpha \quad , \quad f_2 = \sum_{|\alpha|=d} u_{2,\alpha} x^\alpha \quad , \quad f_3 = \sum_{|\alpha|=d} u_{3,\alpha} x^\alpha.$$

1. Let i, j, k be three non-negative integers such that $i + j + k = d - 1$. Show that there exist polynomials p_i, q_i, r_i such that

$$\begin{aligned} f_1 &= x_1^{i+1} p_1 + x_2^{j+1} q_1 + x_3^{k+1} r_1 \\ f_2 &= x_1^{i+1} p_2 + x_2^{j+1} q_2 + x_3^{k+1} r_2 \\ f_3 &= x_1^{i+1} p_3 + x_2^{j+1} q_3 + x_3^{k+1} r_3. \end{aligned} \tag{6.5.1}$$

2. Suppose given a decomposition (6.5.1) for all $(i, j, k) \in \mathbb{N}^3$ such that $i + j + k = d - 1$ and set

$$\Delta_{i,j,k} = \det \begin{pmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_3 \end{pmatrix}.$$

Show that $\Delta_{i,j,k}$ is an inertia form of (f_1, f_2, f_3) and give its degree.

3. Let M be the matrix whose columns are filled with the coefficients of the polynomials
$$X^\alpha f_i \text{ avec } i=1,2,3 \text{ et } |\alpha| = d-2 \quad , \quad \Delta_{i,j,k} \text{ avec } i+j+k=d-1,$$
in the canonical monomial bases. Show that M is a square matrix and that $\det(M)$ is a nonzero inertia form of (f_1, f_2, f_3).
4. Show that $\text{Res}(f_1, f_2, f_3) = \pm\det(M)$ and explain how this matrix can be used to implicitize a parameterized surface in $\mathbb{P}^3$ under some suitable assumptions.

6.5.2 Poisson's formula and classical geometry

Let k be an algebraically closed field and suppose given $n+1$ homogeneous polynomials
$$f_i \in k[x_1, \ldots, x_n]_{d_i}, \; i = 1, \ldots, n-1$$
and
$$f, g \in k[x_1, \ldots, x_n]_d$$
where $d \geq 1$ and $d_i \geq 1$ for all i.

Theorem 6.5.3 (Poisson's formula)**.** *With the above notation, if the algebraic set $V(f_1, \ldots, f_{n-1}, g) \subset \mathbb{P}_k^{n-1}$ is empty then*
$$\frac{\text{Res}(f_1, \ldots, f_{n-1}, f)}{\text{Res}(f_1, \ldots, f_{n-1}, g)} = \prod_{\xi \in V(f_1, \ldots, f_{n-1})} \left(\frac{f}{g}(\xi)\right)^{\mu(\xi)}$$
where $\mu(\xi)$ is the multiplicity of the point $\xi \in V(f_1, \ldots, f_{n-1})$.

We notice that the assumption $V(f_1, \ldots, f_{n-1}, g) = \emptyset$ implies that the resultant $\text{Res}(f_1, \ldots, f_{n-1}, g)$ is nonzero in k, and that the product in Poisson's formula is finite. For a proof of Poisson's formula, we refer to [Jou91].

One application of Poisson formula in classical Euclidian geometry is the following generalization of Menelau's Theorem.

Theorem 6.5.4. *Suppose given a (closed) polygon $(A_i)_{i=1,\ldots,n}$ and an algebraic plane curve $\mathcal{C}$, then*
$$\prod_{i=1}^{n} \left(\prod_{\alpha \in (A_i A_{i+1}) \cap \mathcal{C}} \frac{\overline{\alpha A_i}}{\overline{\alpha A_{i+1}}} \right) = 1$$
with the convention that $A_{n+1} = A_1$ and where the second product runs other all the intersection points α of the line $(A_i A_{i+1})$ and the curve $\mathcal{C}$ (see Figure 6.5.2).

Remark 6.5.5. The notation $\overline{\alpha A_i}$ in the above theorem stands for the signed distance between the two points α and A_i that belong to the line $(A_i A_{i+1})$. If $\vec{u}$ is a unitary vector of the line $(A_i A_{i+1})$ then $\overline{\alpha A_i} = \vec{\alpha A_i} \cdot \vec{u}$.

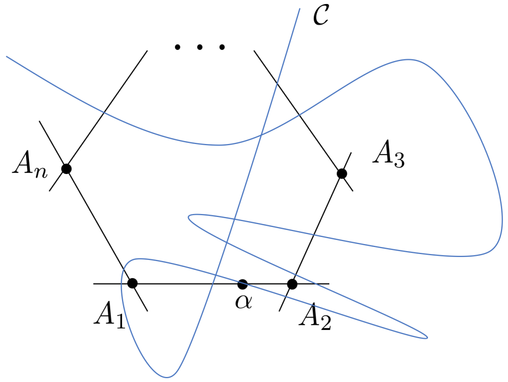


Figure 6.1: Intersection of a polygon and an algebraic curve in the plane.

The main ingredient to prove Theorem 6.5.4 is a generalized property/definition of the power of a point with respect to a plane algebraic curve in a given direction.

Power of a point with respect to a curve and a direction. Suppose given a point I, a curve $\mathcal{C} : f(x,y,z) = 0$ of degree $d \geq 1$ and a line $\mathcal{D}$ passing through the point I. We suppose that $\mathcal{C} \cap \mathcal{D} = \{P_1, \dots, P_r\}$ is at finite distance (i.e. no point P_i belongs to the line of equation $z = 0$ at infinity) and we set $I = (x_0, y_0) = (x_0 : y_0 : 1)$ and $\mathcal{D} : l = ax + by - (ax_0 + by_0)z = 0$. See Figure 6.2.

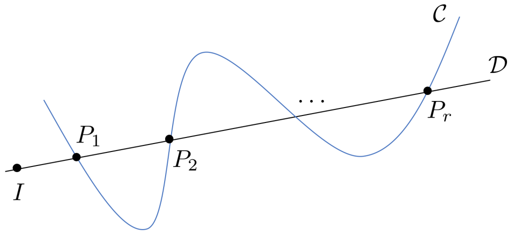


Figure 6.2: Intersection of a line and an algebraic curve.

Proposition 6.5.6. *With the above notation, we have*

$$\prod_{i=1}^{r} \overline{IP_i} = \frac{(a^2+b^2)^{\frac{d}{2}}}{f(-b,a,0)} f(x_0,y_0,1).$$

Proof. A unitary vector generating the line $\mathcal{D}$ is given by

$$\vec{u} = \begin{pmatrix} c \\ d \end{pmatrix} := \frac{1}{\sqrt{a^2+b^2}} \begin{pmatrix} -b \\ a \end{pmatrix}.$$

Thus, setting $P_i = (x_i, y_i)$, we get

$$\overline{IP_i} = \begin{pmatrix} x_i - x_0 \\ y_i - y_0 \end{pmatrix} \cdot \vec{u} = cx_i + dy_i - (cx_0 + dy_0) = l'(P_i)$$

where $l'(x,y,z) = cx + dy - (cx_0 + dy_0)z$ and $P_i = (x_i : y_i : 1)$. Therefore, Poisson's formula shows that

$$\prod_{i=1}^{r} \overline{IP_i} = \frac{\operatorname{Res}(f,l,l')}{\operatorname{Res}(f,l,z)}.$$

Applying the formalism of resultants, the denominator is easily computed as

$$\operatorname{Res}(f,l,z) = \operatorname{Res}(f(x,y,0), ax+by) = f(-b,a,0).$$

To compute the denominator, we apply again Poisson formula but by changing the order of the polynomials. More precisely, Poisson's formula yields

$$\frac{\operatorname{Res}(l,l',f)}{\operatorname{Res}(l,l',z^d)} = f(x_0,y_0,1)$$

because the intersection of l and l' is the point I. As

$$\operatorname{Res}(l,l',z^d) = \operatorname{Res}(l,l',z)^d = \operatorname{Res}(ax+by, cx+dy)^d = (ad-bc)^d = (a^2+b^2)^{\frac{d}{2}},$$

we deduce that $\operatorname{Res}(l,l',f) = (a^2+b^2)^{\frac{d}{2}} f(x_0,y_0,1)$. Since $\operatorname{Res}(l,l',f) = \operatorname{Res}(f,l,l')$, this concludes the proof. □

Remark 6.5.7. If $\mathcal{C}$ is a circle then $f(x,y,z) = \lambda(x^2+y^2) + zl(x,y,z)$ where λ is a nonzero constant and $l(x,y,z)$ is a linear form. Therefore, in this case we get

$$\prod_{i=1}^{2} \overline{IP_i} = \frac{1}{\lambda} f(x_0,y_0,1)$$

which is independent of the direction of the line $\mathcal{D}$, as expected (this property is known as the power of a point with respect to a circle).

What is remarkable in the formula given in Proposition 6.5.6 is that the product which is considered splits into two factors, one depending solely on the direction of the line $\mathcal{D}$ and the other one depending solely on the point I. This property implies the two following properties, as well as the proof of Theorem 6.5.4.

Corollary 6.5.8. *Let $\mathcal{C}$ be an algebraic curve, I, J be two points not on $\mathcal{C}$ and $\mathcal{D}_I$ and $\mathcal{D}_J$ be two parallel lines passing through I and J respectively. Then, denoting by P_i's the intersection points between $\mathcal{D}_I$ and $\mathcal{C}$ and by Q_j's the intersection points between $\mathcal{D}_J$ and $\mathcal{C}$, the ratio*

$$\frac{\prod_i \overline{IP_i}}{\prod_j \overline{JQ_j}}$$

is independent on the direction of both lines $\mathcal{D}_I$ and $\mathcal{D}_J$.

Corollary 6.5.9. *Let $\mathcal{C}$ be an algebraic curve, I be a point not on $\mathcal{C}$ and $\mathcal{D}_1$ and $\mathcal{D}_2$ be two lines passing through I. Then, denoting by P_i's the intersection points between $\mathcal{D}_1$ and $\mathcal{C}$ and by Q_j's the intersection points between $\mathcal{D}_2$ and $\mathcal{C}$, the ratio*

$$\frac{\prod_i \overline{IP_i}}{\prod_j \overline{IQ_j}}$$

is independent of the point I (it only depends on the directions of $\mathcal{D}_1$ and $\mathcal{D}_2$).

Proof of Theorem 6.5.4. Taking again the notation of Theorem 6.5.4, we have to compute the following product

$$\frac{\prod \overline{\alpha A_1}}{\prod \overline{\alpha A_2}} \cdot \frac{\prod \overline{\alpha A_2}}{\prod \overline{\alpha A_3}} \cdot \frac{\prod \overline{\alpha A_3}}{\prod \overline{\alpha A_4}} \cdots \frac{\prod \overline{\alpha A_n}}{\prod \overline{\alpha A_1}}. \tag{6.5.2}$$

But applying Proposition 6.5.6 we see immediately that this product is equal to 1. Indeed, replacing each product by the formula given in this proposition, we see that the terms depending on the direction of the line are killed in each ratio and the term depending on the point I are killed on diagonal (the two product depending on the same point A_i). □

6.5.3 Discriminant of a hypersurface

We consider a hypersurface $\mathcal{H}$ in $\mathbb{P}_k^{n-1}$, k an algebraically closed field, defined by the homogeneous polynomial $f(x_1, \ldots, x_n) \in k[x_1, \ldots, x_n]$ of degree $d \geq 1$. A point on $\mathcal{H}$ is called a regular point if $\mathcal{H}$ admits a tangent hyperplane, otherwise it is called a *singular point*. Thus, it is natural to ask whether a given hypersurface possesses singular points.

Let $p \in \mathcal{H}$. If p is a regular point then the tangent hyperplane of $\mathcal{H}$ at p is the hyperplane of equation

$$T_p\mathcal{H} : \partial_{x_1} f(p)x_1 + \partial_{x_2} f(p)x_2 + \cdots + \partial_{x_n} f(p)x_n = 0,$$

where the notation $\partial_{x_i} f$ denotes the partial derivative of f with respect to x_i. Therefore, singular points of $\mathcal{H}$ are points at which f and all its partial derivatives vanish simultaneously. Assuming that k is of characteristic zero, Euler formula shows that it is actually enough to consider the vanishing of all the partial derivatives. Hence, $\mathrm{Res}(\partial_{x_1} f, \ldots, \partial_{x_n} f)$ yields a necessary and sufficient condition on the coefficients of f for detecting the presence of singular points: the *discriminant* of the hypersurface $\mathcal{H}$. Here is a more precise definition.

Proposition 6.5.10. *Let $f(x_1, \ldots, x_n) = \sum_{|\alpha|=d} u_\alpha x^\alpha$ be the generic homogeneous polynomial of degree $d \geq 2$ and let $A = \mathbb{Z}[u_\alpha : |\alpha| = d]$ be its universal ring of coefficients. The discriminant of f, denoted by* $\mathrm{Disc}(f)$, *is the element in A defined by the equality*

$$\mathrm{Res}(\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f) = d^{\frac{(d-1)^n - (-1)^n}{d}} \mathrm{Disc}(f).$$

It is an irreducible and homogeneous polynomial in A, of degree $n(d-1)^{n-1}$.

We notice that the degree of the discriminant $\mathrm{Disc}(f)$ follows straightforwardly from the multi-degree property of resultants.

As we did for resultants, from Proposition 6.5.10 discriminants are defined by specialization: let $g = \sum_{|\alpha|=d} c_\alpha x^\alpha$ be a homogeneous polynomial in $R[x_1, \ldots, x_n]$, where R is a commutative ring, and let ρ be the specialization map from A to R sending each u_α to c_α. Then, we define the discriminant of g, denoted $\mathrm{Disc}(g)$, by

$$\mathrm{Disc}(g) := \rho(\mathrm{Disc}(f)) \in R.$$

The formalism of discriminants can be developed similarly to what we did for resultants, but it is much more delicate. We close this short introduction on discriminants with an example.

Example 6.5.11. Let $g := u_1x_1^d + u_2x_2^d + \cdots + u_nx_n^d$. Then, $\partial_{x_i} g = du_ix_i^{d-1}$ and hence, using the formalism of resultants, we deduce that

$$\mathrm{Res}(\partial_{x_1} g_1, \ldots, \partial_{x_n} g) = d^{n(d-1)^{n-1}} (u_1 \ldots u_n)^{(d-1)^{n-1}}.$$

It follows that

$$d^{\frac{(d-1)^n-(-1)^n}{d}} \mathrm{Disc}(g) = d^{n(d-1)^{n-1}} (u_1 \ldots u_n)^{(d-1)^{n-1}}.$$

In particular, this shows that a general hypersurface in $\mathbb{P}_k^{n-1}$, where k a field of characteristic $> d$, is smooth (i.e. has no singular point).

Bibliography

[CLO98] David A. Cox, John B. Little, and Donal O'Shea. *Using algebraic geometry.* Graduate texts in mathematics. Springer, New York, 1998.

[CLO07] David Cox, John Little, and Donald O'Shea. *Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra.* Springer, third edition, 2007.

[EH00] David Eisenbud and Joe Harris. *The geometry of schemes*, volume 197 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 2000.

[Eis95] David Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1995. With a view toward algebraic geometry.

[GS] Daniel R. Grayson and Michael E. Stillman. Macaulay2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/.

[Har77] Robin Hartshorne. *Algebraic geometry.* Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.

[Har92] Joe Harris. *Algebraic geometry*, volume 133 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1992. A first course.

[Jou91] Jean-Pierre Jouanolou. Le formalisme du résultant. *Adv. Math.*, 90(2):117–263, 1991.

[Jou97] Jean-Pierre Jouanolou. Formes d'inertie et résultant: un formulaire. *Adv. Math.*, 126(2):119–250, 1997.

[MS20] Mateusz Michalek and Bernd Sturmfels. *Invitation to Nonlinear Algebra*, volume 211 of *Graduate Studies in Mathematics.* AMS, 2020.

[Sch03] Hal Schenck. *Computational Algebraic Geometry.* Cambridge University Press, 2003.