Wolfram Decker and Frank-Olaf Schreyer

Varieties, Gröbner Bases, and Algebraic Curves

With Pictures by Oliver Labs

October 12, 2009

To Doris, Anne, and Matthias, with love
To Nora, Sarah, and Christine, with love

Contents

Part I Affine Algebraic Geometry

Part I

Affine Algebraic Geometry

Chapter 1

The Geometry-Algebra Dictionary

This chapter is an introduction to affine algebraic geometry. We will work over a field $\mathbb{k}$, writing $\mathbb{A}^n(\mathbb{k})$ for the affine n-space over $\mathbb{k}$ and $\mathbb{k}[x_1, \dots, x_n]$ for the ring of polynomials in n variables over $\mathbb{k}$. Our geometric objects of study will be (affine) algebraic sets, which are sets of solutions of polynomial equations in $\mathbb{A}^n(\mathbb{k})$. Although we will not require this in our definition, every algebraic set $A \subset \mathbb{A}^n(\mathbb{k})$ can be described using *finitely many* polynomials. The proof of this fact is a first example of how algebra enters the study of algebraic sets. Namely, writing A as the (common) vanishing locus $\mathrm{V}(I)$ of the elements in an ideal $I \subset \mathbb{k}[x_1, \dots, x_n]$, we will be able to apply Hilbert's basis theorem, a finiteness result for polynomial rings.

Relating to each algebraic set A the ideal $\mathrm{I}(A)$ of all polynomials vanishing on A, we will have defined maps

$$\{\text{algebraic subsets of } \mathbb{A}^n(\mathbb{k})\} \underset{\mathrm{V}}{\overset{\mathrm{I}}{\rightleftarrows}} \{\text{ideals of } \mathbb{k}[x_1, \dots, x_n]\}.$$

This correspondence is best understood over an algebraically closed field $\mathbb{k}$ where Hilbert's celebrated Nullstellensatz characterizes the ideals of type $\mathrm{I}(A)$. Based on the Nullstellensatz, we will develop a dictionary between geometric and algebraic statements. As part of the dictionary, we will study a number of natural geometric operations on algebraic sets together with their algebraic counterparts. Furthermore, we will give examples of how properties of algebraic sets can be expressed in terms of ideals $I \subset \mathbb{k}[x_1, \dots, x_n]$ or, in turn, of quotient rings $\mathbb{k}[x_1, \dots, x_n]/I$. The notion of modules will allow us to treat ideals and quotient rings on equal footing (modules other than ideals and quotient rings will arise naturally in subsequent chapters).

In the final section of this chapter, we will see that each algebraic set $A \subset \mathbb{A}^n(\mathbb{k})$ comes equipped with a ring of functions, the ring of polynomial functions, which is naturally isomorphic to the quotient ring $\mathbb{k}[x_1, \dots, x_n]/\mathrm{I}(A)$. We will use the polynomial functions to define the natural maps, or morphisms, between affine algebraic sets, and to relate these maps to ring homomorphisms on the algebraic side.

Throughout the chapter, in presenting explicit examples, we will occasionally use a piece of terminology whose meaning should be intuitively clear, but whose formal definition will be given later in the book.

1.1 Polynomials

In this section, we will fix our terminology for dealing with polynomials. If R is a ring, and $x_1, \dots, x_n$ is a collection of variables, $R[x_1, \dots, x_n]$ denotes the set of polynomials in $x_1, \dots, x_n$ with coefficients in R. To write the elements of $R[x_1, \dots, x_n]$, we use multiindices. First, a **monomial** in $x_1, \dots, x_n$ is a product $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, where $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$. A **term** in $R[x_1, \dots, x_n]$ is an element of R times a monomial. Finally, each polynomial $0 \neq f \in R[x_1, \dots, x_n]$ can be uniquely expressed as the sum of finitely many nonzero terms involving distinct monomials. These terms (monomials) are called the **terms** (**monomials**) **of** $\boldsymbol{f}$.

The **degree of** $\boldsymbol{x^\alpha}$ $= x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ is $|\alpha| = \alpha_1 + \cdots + \alpha_n$. The **degree of** $\boldsymbol{f}$, written $\deg f$, is the maximum degree of its monomials. The degree of the zero polynomial is $\deg 0 = -\infty$.

With the usual algebraic operations, the set $R[x_1, \dots, x_n]$ becomes a ring which contains R as the subring of constant polynomials, and which is characterized by the following **universal property**: Given any homomorphism ϕ from R to a ring S, and $s_1, \dots, s_n \in S$, there exists a unique homomorphism $\Phi : R[x_1, \dots, x_n] \to S$ extending ϕ, and such that $\Phi(x_i) = s_i$ for all i. In fact, Φ is the map $f \mapsto f(s_1, \dots, s_n)$, where the value $f(s_1, \dots, s_n)$ is obtained by substituting the s_i for the x_i in f and evaluating the corresponding expression in S. We refer to Φ as a **substitution homomorphism**, and write $R[s_1, \dots, s_n]$ for its image in S.

A polynomial in $R[x_1, \dots, x_n]$ is **homogeneous** (**of degree** $\boldsymbol{d}$), if all its monomials have degree d or if the polynomial is zero. We usually write

$$R[x_1, \dots, x_n]_d = \{f \in R[x_1, \dots, x_n] \mid f \text{ is homogenous of degree } d\}.$$

Subsets of polynomials such as $R[x_1, \dots, x_n]_{\leq d}$ and $R[x_1, \dots, x_n]_{<d}$ are defined similarly. Note that if $R = \Bbbk$ is a field, then $\Bbbk[x_0, \dots, x_n]_d$ is a $\Bbbk$-vector space of dimension $\binom{d+n}{n}$. Indeed, the monomials of degree d form a $\Bbbk$-basis.

Every nonzero polynomial $f \in R[x_1, \dots, x_n]$ can be uniquely written as a sum $f = f_0 + f_1 + f_2 + \dots$, where the f_i are homogenenous of degree i. The f_i are called the **homogeneous components** of f.

Given an extra variable x_0, the polynomial

$$f^h := x_0^{\deg(f)} f(x_1/x_0, \dots, x_n/x_0) \in R[x_0, x_1, \dots, x_n]$$

is homogeneous of degree $\deg(f)$, and is called the **homogenization** of f with respect to x_0. Conversely, the **dehomogenization** of a homogeneous polynomial $F \in R[x_0, x_1, \dots, x_n]$ with respect to x_0 is defined to be the polynomial $F(1, x_1, \dots, x_n) \in R[x_1, \dots, x_n]$. We have

$$f^h(1, x_1, \ldots, x_n) = f \text{ and } F = x_0^s \cdot F(1, x_1, \ldots, x_n)^h.$$

where s is the highest power of x_0 dividing F.

If $\boldsymbol{u} \subset \boldsymbol{x} = \{x_1, \ldots, x_n\}$ is a subset of variables, then $R[x_1, \ldots, x_n]$ is canonically isomorphic to $R[\boldsymbol{u}][\boldsymbol{x} \setminus \boldsymbol{u}]$. In particular,

$$R[x_1, \ldots, x_n] \cong R[x_1, \ldots, x_{n-1}][x_n]. \tag{1.1}$$

Explicitly, every polynomial in $R[x_1, \ldots, x_n]$ can be uniquely expressed as a polynomial in x_n with coefficients in $R[x_1, \ldots, x_{n-1}]$.

The isomorphism (1.1) is often used to prove a result on polynomials in several variables by induction on the number of variables. We briefly recall a typical example of how this works (for details, see, for instance, Dummit and Foote (2003), Sections 8.3 and 9.3): The polynomial ring $\mathbb{k}[x]$ in one variable x over a field $\mathbb{k}$ is an Euclidean domain and, hence, a principal ideal domain (PID for short). It is, then, also a unique factorization domain (UFD for short). In particular, if R is an integral domain, and $\mathrm{Q}(R)$ is its quotient field, then $\mathrm{Q}(R)[x]$ is a UFD. Using this and Gauss' lemma, one shows that if R is a UFD, then $R[x]$ is a UFD as well. Inductively, $R[x_1, \ldots, x_n]$ is a UFD.

We will return to some of this later in the book: Euclidean division with remainder will be a topic of Section 2.2, the definition of a PID will be recalled in Section 1.4 below, and quotient fields will be discussed in Section 2.6. As ususal, $\mathbb{k}(x_1, \ldots, x_n) = \mathrm{Q}(\mathbb{k}[x_1, \ldots, x_n])$ will denote the **field of rational functions** in $x_1, \ldots, x_n$ with coefficients in $\mathbb{k}$..

Partial derivatives of polynomials are defined for polynomials with coefficients in any ring R by formally writing the formula familiar from calculus:

Definition 1.1.1. If $f = \sum_\alpha c_\alpha x^\alpha \in R[x_1, \ldots, x_n]$ is a polynomial, its ***i*th formal partial derivative** is defined by the formula

$$\frac{\partial f}{\partial x_i} = \sum_\alpha c_\alpha \alpha_i x_1^{\alpha_1} \cdots x_i^{\alpha_i - 1} \cdots x_n^{\alpha_n} .$$

□

The usual rules of differentiation apply:

Exercise* 1.1.2. 1. Show that $\frac{\partial}{\partial x_i}$ is R-linear.

2. **(Product Rule)** Given $f, g \in R[x_1, \ldots, x_n]$, show that

$$\frac{\partial}{\partial x_i}(fg) = \frac{\partial f}{\partial x_i} g + f \frac{\partial g}{\partial x_i} .$$

3. **(Chain Rule)** Given $g \in R[y_1, \ldots, y_m]$ and $f_j \in R[x_1, \ldots, x_n]$, $j = 1, \ldots, m$, show that

$$\frac{\partial}{\partial x_i}(g(f_1, \ldots, f_m)) = \sum_{j=1}^m \frac{\partial g}{\partial y_j}(f_1, \ldots, f_m) \frac{\partial f_j}{\partial x_i}.$$

4. **(Euler's rule)** If $f \in R[x_1, \dots, x_n]$ is homogeneous of degree d, show that
$$d \cdot f = \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}.$$
□

A polynomial with coefficients in a field of characteristic zero is constant iff all its formal partial derivatives are zero. In characteristic $p > 0$, however, this is not true (for instance, $\frac{\partial x_i^p}{\partial x_i} = p x_i^{p-1} = 0$). Instead, we have:

Exercise* 1.1.3. Show that if $\mathbb{k}$ is a field of characteristic $p > 0$, and $f \in \mathbb{k}[x_1, \dots, x_n]$, then $\frac{\partial f}{\partial x_i} = 0$ iff $f \in \mathbb{k}[x_1, \dots, x_i^p, \dots, x_n]$. Conclude that all the $\frac{\partial f}{\partial x_i}$ are zero iff $f \in \mathbb{k}[x_1^p, \dots, x_n^p]$. □

By allowing infinitely many terms instead of just finitely many, we pass from polynomials to formal power series:

Remark-Definition 1.1.4. Let $\mathbb{k}$ be a field. A **formal power series** in the variables $x_1, \dots, x_n$ with coefficients in $\mathbb{k}$ is an expression of type
$$\sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha, \quad \text{with all } \ a_\alpha \in \mathbb{k}.$$
These expressions form a ring, with algebraic operations
$$\sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha + \sum_{\alpha \in \mathbb{N}^n} b_\alpha x^\alpha = \sum_{\alpha \in \mathbb{N}^n} (a_\alpha + b_\alpha) x^\alpha \ \text{ and}$$
$$\sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha \cdot \sum_{\alpha \in \mathbb{N}^n} b_\alpha x^\alpha = \sum_{\gamma \in \mathbb{N}^n} \Big(\sum_{\alpha + \beta = \gamma} a_\alpha b_\beta \Big) x^\alpha.$$
This ring, denoted $\mathbb{k}[[x_1, \dots, x_n]]$, is called the **ring of formal power series** in n variables $x_1, \dots, x_n$ with coefficients $\mathbb{k}$. Note that $\mathbb{k}[x_1, \dots, x_n]$ is naturally contained in $\mathbb{k}[[x_1, \dots, x_n]]$ as a subring. □

1.2 Algebraic Sets

Let $\mathbb{k}$ be any field. The **affine n-space** over $\mathbb{k}$ is defined to be the set
$$\mathbb{A}^n(\mathbb{k}) := \big\{ (a_1, \dots, a_n) \mid a_1, \dots, a_n \in \mathbb{k} \big\}.$$
An element $p = (a_1, \dots, a_n) \in \mathbb{A}^n(\mathbb{k})$ is called a **point**, and the a_i are called the **coordinates** of p. We say that $\mathbb{A}^1(\mathbb{k})$ and $\mathbb{A}^2(\mathbb{k})$ are the **affine line** and the **affine plane** over $\mathbb{k}$, respectively.

If $\mathbb{k}[x_1, \dots, x_n]$ is the ring of polynomials in n variables with coefficients in $\mathbb{k}$, then each element $f \in \mathbb{k}[x_1, \dots, x_n]$ defines a function
$$f : \mathbb{A}^n(\mathbb{k}) \to \mathbb{k}, \ (a_1, \dots, a_n) \mapsto f(a_1, \dots, a_n).$$

We will refer to such a function as a **polynomial function** on $\mathbb{A}^n(\Bbbk)$, with values in $\Bbbk$. Particular examples are the **coordinate functions** $x_i : \mathbb{A}^n(\Bbbk) \to \Bbbk, (a_1, \dots, a_n) \mapsto a_i$.

Considering a polynomial $f \in \Bbbk[x_1, \dots, x_n]$ as a function on $\mathbb{A}^n(\Bbbk)$ allows us to talk about its **locus of zeros** (or **vanishing locus**) in $\mathbb{A}^n(\Bbbk)$, namely

$$\mathrm{V}(f) := \{p \in \mathbb{A}^n(\Bbbk) \mid f(p) = 0\}.$$

Exercise* 1.2.1. Let $\Bbbk$ be an infinite field, and let $f \in \Bbbk[x_1, \dots, x_n]$ be a polynomial. If f is nonzero, show that the complement $\mathbb{A}^n(\Bbbk) \setminus \mathrm{V}(f)$ is an infinite set. Conclude that f is the zero polynomial iff the polynomial function $f : \mathbb{A}^n(\Bbbk) \to \Bbbk$ is zero.
Hint. Proceed by induction on the number n of variables. To begin with, recall that a nonzero polynomial in one variable has at most finitely many roots. □

Exercise 1.2.2. Let $\mathbb{F}_2$ be the field with two elements. Find a polynomial in $\mathbb{F}_2[x_1, \dots, x_n]$ involving all the x_i and vanishing at each point of $\mathbb{A}^n(\mathbb{F}_2)$. □

Definition 1.2.3. A subset $A \subset \mathbb{A}^n(\Bbbk)$ is called a **hypersurface** in $\mathbb{A}^n(\Bbbk)$ if $A = \mathrm{V}(f)$ for some nonconstant polynomial $f \in \Bbbk[x_1, \dots, x_n]$. □

A hypersurface defined by a degree-1 polynomial

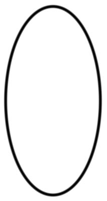
$$f = a_1x_1 + \cdots + a_nx_n - b \in \Bbbk[x_1, \dots, x_n]$$

is called a **hyperplane**. A hypersurface in $\mathbb{A}^2(\Bbbk)$ is called an **affine plane curve**. In showing first examples of such curves, we choose $\Bbbk = \mathbb{R}$ as our ground field so that we can draw pictures:

Example 1.2.4. 1. A **conic** in $\mathbb{A}^2(\mathbb{R})$ is defined by a degree-2 equation

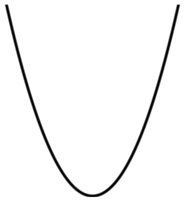
$$ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

where $a, \dots, f \in \mathbb{R}$ are scalars. The **nondegenerate conics**, whose study goes back to the ancient Greek mathematicians, are ellipses, parabolas, and hyperbolas. For instance:



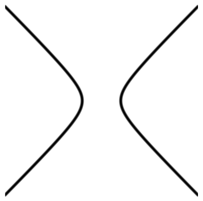
$x^2 + \frac{1}{4}y^2 = 1$

ellipse



$y = x^2$

parabola

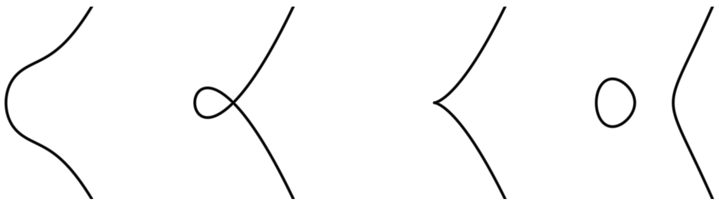


$4x^2 - 4y^2 = 1$

hyperbola

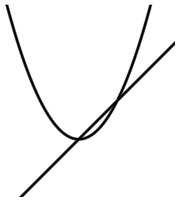
In addition, there are peculiar cases of conics such as the pair of lines with equation $xy = 0$. Can you find other peculiar cases?

2. A **cubic curve** in $\mathbb{A}^2(\mathbb{R})$ is defined by a degree-3 equation. Such curves were systematically investigated by Newton (1666). Here are some particular examples:



$y^2 = x^3 + x + 1$ $y^2 = x^3 + x^2$ $y^2 = x^3$ $y^2 = x^3 - x$

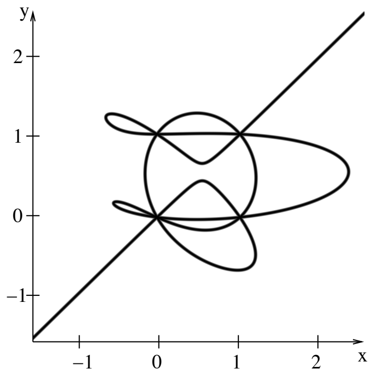
The cubic curve with equation $y^2 = xy + x^2y - x^3$ is the union of a parabola and a line:



3. If $f \in \mathbb{R}[x, y]$ is the degree-seven polynomial

$$\begin{aligned} f = {} & 11\,y^7 + 7\,y^6x + 8\,y^5x^2 - 3\,y^4x^3 - 10\,y^3x^4 - 10\,y^2x^5 - x^7 - 33\,y^6 \\ & - 29\,y^5x - 13\,y^4x^2 + 26\,y^3x^3 + 30\,y^2x^4 + 10\,yx^5 + 3\,x^6 + 33\,y^5 \\ & + 37\,y^4x - 8\,y^3x^2 - 33\,y^2x^3 - 20\,yx^4 - 3\,x^5 - 11\,y^4 - 15\,y^3x \\ & + 13\,y^2x^2 + 10\,yx^3 + x^4, \end{aligned}$$

the curve $C = \mathrm{V}(f) \subset \mathbb{A}^2(\mathbb{R})$ has three triple points and one quadruple point:



□

Exercise 1.2.5. Let $f \in \mathbb{R}[x,y]$ and $C = \mathrm{V}(f) \subset \mathbb{A}^2(\mathbb{R})$ be as in the preceeding example, and let $\mathbb{R}(t) = \mathrm{Q}(\mathbb{R}[t])$ be the field of rational functions in one variable t with coefficients in $\mathbb{R}$. If $x(t), y(t) \in \mathbb{R}(t)$ are the rational functions

$$x\,(t) = \frac{121\,t^7 - 253\,t^6 - 133\,t^5 + 364\,t^4 + 39\,t^3 - 92\,t^2 + 10\,t}{121\,t^7 - 127\,t^6 - 114\,t^5 + 29\,t^4 + 54\,t^3 + 106\,t^2 - 20\,t + 1}\,,$$

$$y\,(t) = \frac{-77\,t^7 + 72\,t^6 + 246\,t^5 - 192\,t^4 - 138\,t^3 + 116\,t^2 - 20\,t + 1}{121\,t^7 - 127\,t^6 - 114\,t^5 + 29\,t^4 + 54\,t^3 + 106\,t^2 - 20\,t + 1}\,,$$

compute that $f(x(t), y(t)) = 0 \in \mathbb{R}(t)$. Hence, there is a well-defined map

$$\varphi : U \to C,\ a \mapsto (x(a), y(a)),$$

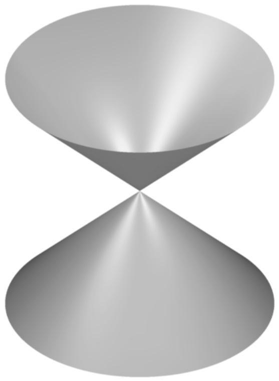
where U consists of all points of $\mathbb{A}^1(\mathbb{R})$ except the real roots of the denominator of $x(t)$ and $y(t)$.
Hint. The coefficients of f, $x(t)$, and $y(t)$ are rational numbers (in fact, integers). Thus, the actual computation takes place in $\mathbb{Q}(t)$. Rather than doing the computation bare-handed, use your favorite computer algebra system. □

Remark 1.2.6. Rational parametrizations such as the map φ in the exercise above will be treated systematically in Section 2.6. In the second half of the book, we will discuss how to decide whether a given curve admits such a parametrization (actually, “most” curves don’t). And, we will present a method for computing rational parametrizations of plane curves in cases where such parametrizations exist. □

Hypersurfaces in affine 3-space provide examples of surfaces:

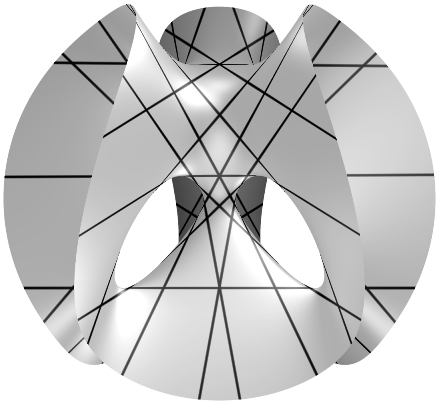
Example 1.2.7. Let $\Bbbk = \mathbb{R}$.

1. The surface
$$\mathrm{V}(x^2 + y^2 - z^2) \subset \mathbb{A}^3(\mathbb{R})$$
is a cone with vertex at the origin:



Note that the ancient Greeks (most notably, Apollonius) realized the nondegenerate conics as sections of cones by planes (see Kline (1972) for some historical remarks).

2. **Clebsch's diagonal cubic** in $\mathbb{A}^3(\mathbb{R})$ is a surface containing precisely 27 real lines (see Clebsch (1871), §16):



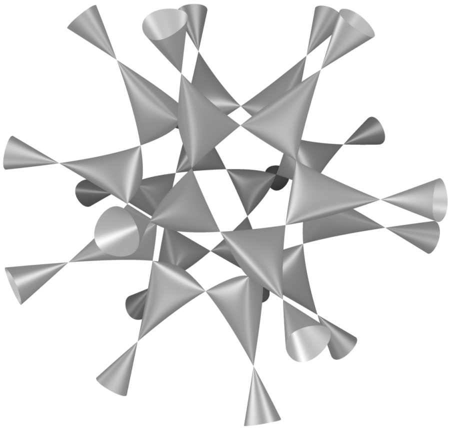
It is defined by the equation

$$(p^3 + q^3 + r^3 - s^3) - (p + q + r - s)^3 = 0,$$

where

$$p = 1 - z - cx,\ q = 1 - z + cx,\ r = 1 + z + cy,$$
$$s = 1 + z - cy, \quad \text{with} \quad c = \sqrt{2}\,.$$

3. **Barth's sextic** in $\mathbb{A}^3(\mathbb{R})$ is a surface with 50 nodes (see Barth (1996)):



It is defined by the equation

$$\begin{aligned}&(8c+4)x^2y^2z^2 - c^4(x^4y^2+y^4z^2+x^2z^4)+c^2(x^2y^4+y^2z^4+x^4z^2)\\ &-\tfrac{2c+1}{4}(x^2+y^2+z^2-1)^2=0, \quad \text{where} \quad c=\tfrac{1+\sqrt{5}}{2} \quad \text{is the golden section.} \ \square\end{aligned}$$

In general, we are not only concerned with hypersurfaces, but also with sets defined by more than one polynomial equation: If $T \subset \Bbbk[x_1,\ldots,x_n]$ is any subset, its **locus of zeros** (or **vanishing locus**) is the set

$$\mathrm{V}(T) = \{p \in \mathbb{A}^n(\Bbbk) \mid f(p) = 0 \text{ for all } f \in T\}.$$

If $T = \{f_1,\ldots,f_r\}$ is finite, we write $\mathrm{V}(f_1,\ldots,f_r) = \mathrm{V}(T)$.

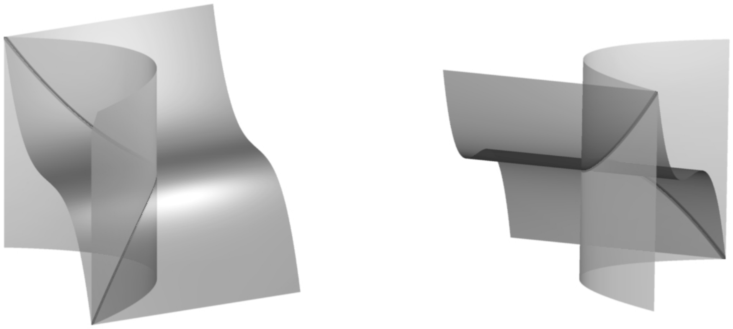
Definition 1.2.8. A subset $A \subset \mathbb{A}^n(\Bbbk)$ is called an **algebraic subset**, or simply an **algebraic set**, if $A = \mathrm{V}(T)$ for some subset $T \subset \Bbbk[x_1,\ldots,x_n]$. An **affine algebraic set** is an algebraic subset of some $\mathbb{A}^n(\Bbbk)$. □

Since $\mathrm{V}(T) = \bigcap_{f\in T} \mathrm{V}(f)$, a subset of $\mathbb{A}^n(\Bbbk)$ is algebraic iff it can be written as the intersection of hypersurfaces.

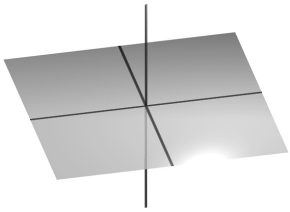
Example 1.2.9. The intersection of hyperplanes is the set of solutions of a system of linear equations as studied in linear algebra. We will refer to such a set as a **linear subvariety** of $\mathbb{A}^n(\Bbbk)$. □

Example 1.2.10. Let $\Bbbk = \mathbb{R}$.

1. The intersection of the two hypersurfaces $\mathrm{V}(y-x^2)$ and $\mathrm{V}(z-x^3)$ in $\mathbb{A}^3(\mathbb{R})$ is called the **twisted cubic curve** in $\mathbb{A}^3(\mathbb{R})$:



2. Intersecting the hypersurfaces $\mathrm{V}(xz)$ and $\mathrm{V}(yz)$ in $\mathbb{A}^3(\mathbb{R})$ gives the union of the xy-plane and the z-axis:



□

Exercise 1.2.11. Use your favorite system(s) for visualization to draw your own pictures of the algebraic sets in Examples 1.2.4, 1.2.7 and 1.2.10. □

Remark-Definition 1.2.12. Given a polynomial $f \in \Bbbk[x_1, \dots, x_n]$, we write

$$\mathrm{D}(f) := \mathbb{A}^n(\Bbbk) \setminus \mathrm{V}(f).$$

We have $\mathrm{D}(1) = \mathbb{A}^n(\Bbbk)$, and if $f, g \in \Bbbk[x_1, \dots, x_n]$, then $\mathrm{D}(f) \cap \mathrm{D}(g) = \mathrm{D}(fg)$. Hence, the $\mathrm{D}(f)$ form the basis for a topology on $\mathbb{A}^n(\Bbbk)$ whose closed sets are the algebraic subsets of $\mathbb{A}^n(\Bbbk)$. This topology is called the **Zariski topology** on $\mathbb{A}^n(\Bbbk)$, and the $\mathrm{D}(f)$ are called the **distinguished open sets**. □

In this book, if not otherwise mentioned, the affine n-space $\mathbb{A}^n(\Bbbk)$ will be endowed with the Zariski topology. Subsets of $\mathbb{A}^n(\Bbbk)$ will carry the induced topology which is called the **Zariski topology** on the subset. Topological notions such as open, closed, dense, or neighborhood will refer to this topology. If $A \subset \mathbb{A}^n(\Bbbk)$ is a subset, then $\overline{A}$ will denote its closure in the Zariski topology.

Remark 1.2.13. If $\Bbbk = \mathbb{R}$ or $\Bbbk = \mathbb{C}$, every subset of $\mathbb{A}^n(\Bbbk)$ which is open in the Zariski topology is also open in the usual Euclidean topology. Indeed, polynomial functions on $\mathbb{A}^n(\Bbbk)$ are continuous in the Euclidean topology. □

As we know from the linear case, the equations describing an algebraic set are by no means unique. In fact, we usually solve a system of linear equations by transforming it to an equivalent system from which the solutions can be read off. Each new equation is obtained as a linear combination of the original ones, using scalars as coefficients. In the more general situation here, we consider linear combinations of polynomial equations with polynomials instead of just scalars as coefficients. For instance, considering $1\cdot(z-x^3)-x\cdot(y-x^2) = z-xy$, we see that the twisted cubic curve $\mathrm{V}(y - x^2, z - x^3)$ may also be described as the intersection of the hypersurfaces $\mathrm{V}(y - x^2)$ and $\mathrm{V}(z - xy)$:

In general, if $T \subset \Bbbk[x_1, \dots, x_n]$ is any set of polynomials, and $\langle T\rangle$ is the set of all $\Bbbk[x_1, \dots, x_n]$-linear combinations $g_1 f_1 + \cdots + g_r f_r$, where $f_1, \dots, f_r \in T$, then $\mathrm{V}(T) = \mathrm{V}(\langle T\rangle)$. In the language of ideals, which we quickly recall in the following section, this means that in definining the algebraic set $\mathrm{V}(T)$, we may replace T by the ideal generated by T.

1.3 Ideals

Let R be a ring.

Definition 1.3.1. An **ideal** of R is an additive subgroup I of R such that if $f \in R$ and $g \in I$, then $fg \in I$. □

If T is any nonempty subset of R, the set of all R-linear combinations of elements of T, written $\langle T \rangle$, is an ideal of R. In fact, it is the smallest ideal of R containing T. We refer to it as the **ideal generated by** $\boldsymbol{T}$. If $T = \{f_1, \ldots, f_r\}$ is finite, we write $\langle f_1, \ldots, f_r \rangle$ for $\langle T \rangle$. By convention, the ideal generated by the empty subset of R is $\langle 0 \rangle$.

If $I \subset R$ is an ideal, any subset T of I satisfying $I = \langle T \rangle$ is called a **set of generators** for I. We say that I is **finitely generated** if it admits a finite set of generators. It is **principal** if it can be generated by a single element.

Exercise* 1.3.2. 1. If $\{I_\lambda\}$ is a family of ideals of R, show that the intersection $\bigcap_\lambda I_\lambda$ is also an ideal of R.

2. If $I_1, \ldots, I_s$ are ideals of R, their **product** $I_1 \cdots I_s$ is the ideal generated by the elements $f_1 \cdots f_s$, where $f_k \in I_k$ for all k. Prove that $I_1 \cdots I_s \subset \bigcap_{k=1}^s I_k$, and give an example showing that the inclusion may be strict. □

The union of a family $\{I_\lambda\}$ of ideals of R is not necessarily an ideal. The **sum** of the I_λ, written $\sum_\lambda I_\lambda$, is the ideal generated by the union $\bigcup_\lambda I_\lambda$.

If I, J are two ideals of R, the set

$$I : J = \{f \in R \mid fg \in I \text{ for all } g \in J\}$$

is an ideal of R containing I. It is called the **ideal quotient** of I by J. If g is a single element of R, we usually write $I : g$ instead of $I : \langle g \rangle$.

Exercise* 1.3.3. Let I, I_k, J, J_k, K be ideals of R, $1 \le k \le s$, and let $g \in R$. Show:

1. $$I : J = R \iff J \subset I.$$

2. $$\left(\bigcap_{k=1}^s I_k\right) : J = \bigcap_{k=1}^s (I_k : J).$$

3. $$I : \left(\sum_{k=1}^s J_k\right) = \bigcap_{k=1}^s (I : J_k).$$

4. $$(I : J) : K = I : JK.$$

5. $$I : g^m = I : g^{m+1} \Longrightarrow I = (I : g^m) \cap \langle I, g^m \rangle.$$ □

We say that an ideal I of R is a **proper ideal** if $I \neq R$. A proper ideal $\mathfrak{p}$ of R is a **prime ideal** if $f, g \in R$ and $fg \in \mathfrak{p}$ implies $f \in \mathfrak{p}$ or $g \in \mathfrak{p}$. A proper ideal $\mathfrak{m}$ of R is a **maximal ideal** if there is no ideal I of R such that $\mathfrak{m} \subsetneq I \subsetneq R$.

Exercise* **1.3.4.** Show:

1. Every maximal ideal of R is a prime ideal of R.
2. If $I_1, \dots, I_s \subset R$ are ideals, and $\mathfrak{p} \subset R$ is a prime ideal containing the product $I_1 \cdots I_s$, then $\mathfrak{p}$ contains one of the I_k.
3. (**Prime Avoidance**) If $\mathfrak{p}_1, \dots, \mathfrak{p}_s \subset R$ are ideals, and $I \subset R$ is an ideal contained in the union $\bigcup_{k=1}^{s} \mathfrak{p}_k$, then I is contained in one of the $\mathfrak{p}_k$. □

Conditions on an ideal I of R may also be expressed as conditions on the quotient ring R/I. We briefly recall the definition of the quotient ring:

Remark-Definition 1.3.5. Let $I \subset R$ be an ideal. Two elements f, g of R are said to be **congruent modulo *I***, written

$$f \equiv g \mod I,$$

if $f - g \in I$. The relation on R defined by congruence modulo I is an equivalence relation. We usually write $\overline{f} = f + I$ for the equivalence class of $f \in R$, and call it the **residue class** of f modulo I. The set of all residue classes becomes a ring, with algebraic operations

$$\overline{f} + \overline{g} = \overline{f + g} \text{ and } \overline{f} \cdot \overline{g} = \overline{f \cdot g}.$$

We refer to this ring as the **quotient ring** R/I, and to the map

$$R \to R/I,\ f \mapsto \overline{f},$$

as the **canonical projection** onto R/I. □

Any homomorpism of rings $\phi : R \to S$ gives rise to a monomorphism

$$R/\ker\phi \to S,\ \overline{f} \mapsto \phi(f),$$

which is an isomorphism iff ϕ is an epimorphism. The proof of this fact, which is known as the **homomorphy theorem**, will be postponed to Exercise 1.10.5, where we will treat the theorem in a more general setting.

Exercise* **1.3.6.** Let I be an ideal of R. Show:

1. I is prime $\iff$ R/I is an integral domain.
2. I is maximal $\iff$ R/I is a field. □

The relationship between ideals of R/I and ideals of R will be discussed in Exercise 1.5.11.

Definition 1.3.7. A ring R is called a **local ring** if it has exactly one maximal ideal. If $\mathfrak{m}$ is this ideal, we also say that $(R, \mathfrak{m})$ is a local ring, and refer to $R/\mathfrak{m}$ as the **residue field** of R. □

Remark 1.3.8. The name local comes from geometry (see Section 4.2). Note that a ring R is local iff its nonunits form a (maximal) ideal. □

Two ideals $I, J \subset R$ are called **coprime** if $I + J = \langle 1 \rangle$.

Exercise* 1.3.9 (Chinese Remainder Theorem). Let $I_1, \dots, I_s$ be ideals of R. Consider the natural ring homomorphism

$$\phi : R \to \prod_{k=1}^{s} R/I_k,\ f \mapsto (f + I_1, \dots, f + I_s).$$

Show:

1. The kernel of ϕ is $\ker \phi = \bigcap_{k=1}^{s} I_k$.
2. If the I_k are pairwise coprime, then $I_1 \cdots I_s = \bigcap_{k=1}^{s} I_k$.
3. The I_k are pairwise coprime iff ϕ is surjective.

If the I_k are pairwise coprime, conclude from the homomorphy theorem that

$$R/(\bigcap_{k=1}^{s} I_k) \cong \prod_{k=1}^{s} R/I_k.$$

□

Remark 1.3.10. Let $\phi : R \to S$ be a homomorphism of rings. If J is an ideal of S, then $\phi^{-1}(J)$ is an ideal of R. In contrast, if I is an ideal of R, then $\phi(I)$ is not necessarily an ideal of S (consider, for instance, the inclusion $\mathbb{Z} \subset \mathbb{Q}$ and any nonzero ideal I of $\mathbb{Z}$). We usually write $IS = \phi(I)S$ for the ideal generated by $\phi(I)$ in S. □

1.4 Hilbert's Basis Theorem

Representing algebraic sets by ideals will allow us to bring algebra into the study of algebraic sets. Corollary 1.4.2 below is an example of how this works.

Theorem 1.4.1 (Hilbert's Basis Theorem). *Every ideal of* $\Bbbk[x_1, \dots, x_n]$ *has a finite set of generators.* □

Corollary 1.4.2. *Every algebraic subset of* $\mathbb{A}^n(\Bbbk)$ *can be expressed as the vanishing locus of finitely many polynomials.*

Proof (of the corollary). Given $\mathrm{V}(T)$, apply the basis theorem to the ideal $\langle T \rangle \subset \Bbbk[x_1, \dots, x_n]$. □

Remark 1.4.3. In terms of the Zariski topology, this means that every open subset of $\mathbb{A}^n(\Bbbk)$ is the *finite* union of distinguished open sets. □

All known proofs of the basis theorem proceed by induction on the number of variables, starting with the univariate case which is particularly easy:

Remark 1.4.4. The polynomial ring $\Bbbk[x]$ in one variable x is a **principal ideal domain** (**PID** for short). That is, every ideal I of $\Bbbk[x]$ is principal. Indeed, if $f \in I$ is a nonzero polynomial of minimal degree, use Euclidean division with remainder to show that $I = \langle f \rangle$. □

Hilbert's original proof of the basis theorem can be found in the first of his two famous papers on invariant theory (1890, 1893). These papers contain further fundamental results which will play a prominent role in this book: the Nullstellensatz 1.6.2, the Syzygy Theorem 2.8.11, and Theorem **??** on the polynomial nature of what is nowadays called the Hilbert function.

Note that Hilbert and his contemporaries used the word "basis" as another name for a "(finite) set of generators". In Chapter 2, we will encounter special bases, nowadays called Gröbner bases, which are well-suited for computational purposes. Historically, these bases were already considered by Gordan (1899) who used them to give his own proof of the basis theorem. We refer to Exercise 2.1.2 and Corollary 2.3.3 for this proof.

The general theory of rings in which every ideal is finitely generated was developed by Emmy Noether (1921), a student of Gordan. In particular, Noether realized the importance of the ascending chain condition (see Exercise 1.4.5 below). From this condition, she derived the existence of primary decompositions (we will treat this in Section 1.8).

Exercise* 1.4.5. Show that the following conditions on a ring R are equivalent:

1. **(Finiteness condition)** Every ideal of R is finitely generated.
2. **(Ascending chain condition)** Every chain
$$I_1 \subset I_2 \subset I_3 \subset \dots$$
of ideals of R is eventually stationary. That is,
$$I_m = I_{m+1} = I_{m+2} = \dots \quad \text{for some} \quad m \geq 1.$$
3. **(Maximal condition)** Every nonempty set of ideals of R has a maximal element with respect to inclusion. □

Definition 1.4.6. A ring satisfying the equivalent conditions above is called a **Noetherian ring.** □

The following exercise shows how the ascending chain condition can be used to prove the basis theorem:

Exercise* 1.4.7 (Hilbert's Basis Theorem, General Version). If R is a Noetherian ring, show that $R[x]$ is Noetherian. Conclude that the polynomial rings $\mathbb{Z}[x_1, \dots, x_n]$ and $\Bbbk[x_1, \dots, x_n]$ are Noetherian.
Hint. Suppose that there is an ideal $I \subset R[x]$ which is not finitely generated. Let $f_1 \in I$ be a nonzero polynomial of minimal degree, and let $a_1 \in R$ be its leading coefficient (that is, the coefficient of the term of highest degree). Construct an ascending chain of ideals
$$\langle a_1 \rangle \subsetneq \langle a_1, a_2 \rangle \subsetneq \dots \subsetneq R.$$
□

1.5 The Correspondences V and I

By taking vanishing loci, we get a map V which sends subsets of $\mathbb{k}[x_1,\dots,x_n]$ to algebraic subsets of $\mathbb{A}^n(\mathbb{k})$. We summarize some properties of this map:

Proposition 1.5.1. *Let $R=\mathbb{k}[x_1,\dots,x_n]$. Then:*

1. $\mathrm{V}(0)=\mathbb{A}^n(\mathbb{k})$. $\quad\mathrm{V}(1)=\emptyset$.
2. *If $I\subset J$ are subsets of R, then $\mathrm{V}(I)\supset\mathrm{V}(J)$.*
3. *If I,J are ideals of R, then*

$$\mathrm{V}(I)\cup V(J)=\mathrm{V}(I\cdot J)=\mathrm{V}(I\cap J).$$

4. *If $\{I_\lambda\}$ is a family of ideals of R, then*

$$\bigcap_\lambda \mathrm{V}(I_\lambda)=\mathrm{V}\left(\sum_\lambda I_\lambda\right).$$

5. *If $a_1,\dots,a_n\in\mathbb{k}$, then*

$$\mathrm{V}(x_1-a_1,\dots,x_n-a_n)=\{(a_1,\dots,a_n)\}.$$

□

Exercise* 1.5.2. Prove Proposition 1.5.1. □

Now, proceeding in the other direction, we define a map I which sends subsets of $\mathbb{A}^n(\mathbb{k})$ to ideals in $\mathbb{k}[x_1,\dots,x_n]$:

Remark-Definition 1.5.3. If $A\subset\mathbb{A}^n(\mathbb{k})$ is any subset, the set

$$\mathrm{I}(A):=\{f\in\mathbb{k}[x_1,\dots,x_n]\mid f(p)=0 \text{ for all } p\in A\}$$

is an ideal of $\mathbb{k}[x_1,\dots,x_n]$. It is called the **vanishing ideal** of A. □

Exercise 1.5.4. 1. Show that every polynomial $f\in\mathbb{k}[x,y,z]$ has a representation of type

$$f=g_1(y-x^2)+g_2(z-x^3)+h,$$

where $g_1,g_2\in\mathbb{k}[x,y,z]$ and $h\in\mathbb{k}[x]$.

2. Let $\mathbb{k}$ be infinite, and let $C=\mathrm{V}(y-x^2,z-x^3)\subset\mathbb{A}^3(\mathbb{k})$ be the **twisted cubic curve** in $\mathbb{A}^3(\mathbb{k})$. Show that

$$\mathrm{I}(C)=\langle y-x^2,z-x^3\rangle.$$

Hint. To obtain the representation in part 1, first suppose that f is a monomial. For part 2, use that C can be parametrized:

$$C=\{(a,a^2,a^3)\mid a\in\mathbb{k}\}.$$

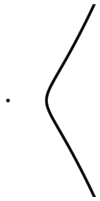
□

The expression for f in terms of $y - x^2$ and $z - x^3$ in the exercise above can be computed in a more systematic way, using a general algorithm for division with remainder. We will come back to this in Exercise 2.2.15.

Exercise 1.5.5. Let $\mathbb{k} = \mathbb{R}$, and let

$$C = \{(a^2 + 1, a^3 + a) \mid a \in \mathbb{R}\} \subset \mathbb{A}^2(\mathbb{R}).$$

Show that $\mathrm{I}(C) = \langle y^2 - x^3 + x^2 \rangle$, and conclude that $\overline{C} = C \cup \{(0, 0)\}$.



Hint. For the second statement, consider lines through $o = (0, 0)$. □

Remark 1.5.6. The computations in both exercises above make use of a parametrization of the given curve. In general, no such parametrization exists, and it can be a difficult task to compute $\mathrm{I}(A)$ (see Remark 2.4.13 for hints on algorithms). A method which often allows one to decide whether a given set of polynomials defining A actually generates $\mathrm{I}(A)$ can be deduced from the Jacobian Criterion 4.1.12 (see Corollaries 4.1.13 and 4.1.14). □

In the following proposition, we summarize some properties of I, and start examining how V and I are related:

Proposition 1.5.7. *Let $R = \mathbb{k}[x_1, \dots, x_n]$. Then:*

1. $\mathrm{I}(\emptyset) = R$. *If* $\mathbb{k}$ *is infinite, then* $\mathrm{I}(\mathbb{A}^n(\mathbb{k})) = \langle 0 \rangle$.
2. *If* $A \subset B$ *are subsets of* $\mathbb{A}^n(\mathbb{k})$, *then* $\mathrm{I}(A) \supset \mathrm{I}(B)$.
3. *If* A, B *are subsets of* $\mathbb{A}^n(\mathbb{k})$, *then*
$$\mathrm{I}(A \cup B) = \mathrm{I}(A) \cap \mathrm{I}(B).$$
4. *If* $(a_1, \dots, a_n) \in \mathbb{A}^n(\mathbb{k})$ *is a point, then*
$$\mathrm{I}(\{(a_1, \dots, a_n)\}) = \langle x_1 - a_1, \dots, x_n - a_n \rangle.$$
5. *For any subset* $A \subset \mathbb{A}^n(\mathbb{k})$, *we have*
$$\mathrm{V}(\mathrm{I}(A)) \supset A,$$
with equality occuring iff A *is algebraic. In any case,* $\mathrm{V}(\mathrm{I}(A)) = \overline{A}$.
6. *For any subset* $I \subset R$, *we have*
$$\mathrm{I}(\mathrm{V}(I)) \supset I,$$
with equality occuring iff I *is the vanishing ideal of a subset of* $\mathbb{A}^n(\mathbb{k})$. □

Exercise* 1.5.8. Prove Proposition 1.5.7. □

Not every ideal $I \subset \mathbb{k}[x_1, \dots, x_n]$ can occur as a vanishing ideal $\mathrm{I}(A)$. That is, the inclusion $\mathrm{I}(\mathrm{V}(I)) \supset I$ may well be strict. To put it yet in another way, the map V is not injective. In fact, there are two different ways in which distinct ideals can represent the same algebraic set. The following example indicates one possibility:

$$\{0\} = \mathrm{V}(x) = \mathrm{V}(x^2) = \mathrm{V}(x^3) = \ \cdots \subset \mathbb{A}^1(\mathbb{k}).$$

In general, if a power f^m of a polynomial f vanishes on a subset $A \subset \mathbb{A}^n(\mathbb{k})$, then f itself vanishes on A. Thus, vanishing ideals have a property not shared by all ideals – they are radical ideals in the following sense:

Remark-Definition 1.5.9. Let R be a ring, and let $I \subset R$ be an ideal. Then the set

$$\operatorname{rad} I := \{f \in R \mid f^m \in I \text{ for some } m \geq 1\}$$

is an ideal of R: use the binomial theorem to show that if $r, s \in R$ and $f, g \in \operatorname{rad} I$, then $rf + sg \in \operatorname{rad} I$. We call $\operatorname{rad} I$ the **radical** of I. Clearly, $\operatorname{rad} I \supset I$. If $\operatorname{rad} I = I$, then I is called a **radical ideal**. □

Example 1.5.10. If R is a UFD, the radical of every principal ideal of R is again a principal ideal. In fact, if $f \in R$ is a nonzero nonunit, decompose f into its distinct irreducible factors:

$$f = u \cdot f_1^{\mu_1} \cdots f_s^{\mu_s}.$$

Here, u is a unit, the μ_i are integers ≥ 1, and the f_i are irreducible and pairwise coprime. Then

$$\operatorname{rad} \langle f \rangle = \langle f_1 \cdots f_s \rangle.$$

The product $f_1 \cdots f_s$, which is uniquely determined by f up to multiplication by a unit, is called the **square-free part** of f. If all the μ_i are 1, we say that f is **square-free**, or **reduced**, or **without multiple factors**. □

If R is any ring, the ideal

$$\operatorname{rad} \langle 0 \rangle = \{f \in R \mid f^m = 0 \text{ for some } m \geq 1\}$$

is called the **nilradical** of R, and its elements the **nilpotent** elements of R. We say that R is a **reduced ring** if $\operatorname{rad} \langle 0 \rangle = \langle 0 \rangle$. Clearly, a quotient ring R/I is reduced iff I is a radical ideal.

Exercise* 1.5.11. Let I be an ideal of a ring R, and let $\pi : R \to R/I$ be the canonical projection. Show:

1. There is a one-to-one correspondence between the ideals J of R/I and the ideals of R containing I, obtained by sending J to $\pi^{-1}(J)$.

2. Under this correspondence, radical ideals correspond to radical ideals. Similarly for prime and maximal ideals.

Conclude that if R is Noetherian (local), then R/I is Noetherian (local) as well. □

Exercise* 1.5.12. Let I, J be ideals of a ring R. Show:

1. $$\operatorname{rad}(IJ) = \operatorname{rad}(I \cap J) = \operatorname{rad} I \cap \operatorname{rad} J.$$
2. $$\operatorname{rad}(I + J) = \operatorname{rad}(\operatorname{rad} I + \operatorname{rad} J).$$
3. $$\operatorname{rad} I = \langle 1 \rangle \iff I = \langle 1 \rangle.$$
4. If $\operatorname{rad} I, \operatorname{rad} J$ are coprime, then I, J are coprime as well. □

1.6 Hilbert's Nullstellensatz

Even for radical ideals, it may happen that distinct ideals give the same algebraic set:

$$\mathrm{V}(1 + x^2) = \mathrm{V}(1) = \emptyset \subset \mathbb{A}^1(\mathbb{R}).$$

Here, we face a problem which is caused by properties of the ground field. Passing from $\mathbb{R}$ to the field $\mathbb{C}$ of complex numbers, the problem will disappear. Indeed, by the fundamental theorem of algebra, $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$. And, if $\Bbbk$ is any field, and $\overline{\Bbbk}$ is its algebraic closure, then every nonconstant polynomial in $\Bbbk[x]$ has a root in $\overline{\Bbbk}$ (by the very definition of $\overline{\Bbbk}$). In terms of ideals $I \subset \Bbbk[x]$, since $\Bbbk[x]$ is a PID, we conclude that the locus of zeros of I in $\mathbb{A}^1(\overline{\Bbbk})$ is empty iff $1 \in I$. This result extends to polynomials in more than one variable:

Theorem 1.6.1 (Hilbert's Nullstellensatz, Weak Version). *Let I be an ideal of $\Bbbk[x_1, \dots, x_n]$, and let $\overline{\Bbbk}$ be the algebraic closure of $\Bbbk$. Then the following are equivalent:*

1. *The locus of zeros of I in $\mathbb{A}^n(\overline{\Bbbk})$ is empty.*
2. *$1 \in I$.* □

We will prove this version of the Nullstellensatz in Section 3.1. Now, we discuss some consequences. To begin with, we deduce a strong version of the Nullstellensatz which implies that the correspondence between algebraic sets and ideals is well behaved if we restrict our attention to radical ideals, and if we work over an algebraically closed field:

Theorem 1.6.2 (Hilbert's Nullstellensatz, Strong Version). *Let $\Bbbk = \overline{\Bbbk}$ be algebraically closed, and let*

$$I \subset \Bbbk[x_1, \dots, x_n]$$

be an ideal. Then

$$\mathrm{I}(\mathrm{V}(I)) = \operatorname{rad} I.$$

Proof. If $f \in \operatorname{rad} I$, then $f^m \in I$ for some $m \geq 1$. This implies that f^m and, hence, f vanish on $\mathrm{V}(I)$. We conclude that

$$\operatorname{rad} I \subset \mathrm{I}(\mathrm{V}(I)).$$

For the opposite inclusion, let $f \in \mathrm{I}(\mathrm{V}(I))$, and let $f_1, \dots, f_r$ be polynomials generating I. Then f vanishes on $\mathrm{V}(I)$, and we have to show that $f^m = g_1 f_1 + \ldots + g_r f_r$ for some $m \geq 1$ and some $g_1, \dots, g_r \in \Bbbk[x_1, \dots, x_n]$.

For this, we use the trick of Rabinowitch. Consider the ideal

$$J := \langle f_1, \dots, f_r, yf - 1 \rangle \subset \Bbbk[x_1, \dots, x_n, y],$$

where y is an extra variable. Proceeding in two steps, we will show in Step 1 that $\mathrm{V}(J) \subset \mathbb{A}^{n+1}(\Bbbk)$ is empty. Then, in Step 2, we will apply the weak version of the Nullstellensatz to conclude that $1 \in J$. The result will follow from a representation of 1 as a $\Bbbk[x_1, \dots, x_n, y]$-linear combination of $f_1, \dots, f_r, yf - 1$.

Step 1. Consider a point $p = (a_1, \dots, a_{n+1}) \in \mathbb{A}^{n+1}(\Bbbk)$. To show that $p \notin \mathrm{V}(J)$, we distinguish two cases. If $(a_1, \dots, a_n) \in \mathrm{V}(I)$, then $f(a_1, \dots, a_n) = 0$ since $f \in \mathrm{I}(\mathrm{V}(I))$. Evaluating $yf - 1$ in $(a_1, \dots, a_{n+1})$ gives

$$a_{n+1} f(a_1, \dots, a_n) - 1 = -1 \neq 0,$$

so that $p = (a_1, \dots, a_{n+1}) \notin \mathrm{V}(J)$. If $(a_1, \dots, a_n) \notin \mathrm{V}(I)$, then $f_k(a_1, \dots, a_n)$ is nonzero for some k. Since $f_k \in J$, we, again, find that $p \notin \mathrm{V}(J)$. We conclude that $\mathrm{V}(J) = \emptyset$.

Step 2. By Step 1 and the weak version of the Nullstellensatz, we have $1 \in J$. Hence, there are polynomials $h_1, \dots, h_r,\ h \in \Bbbk[x_1, \dots, x_n, y]$ such that

$$1 = \sum_{i=1}^{r} h_i(x_1, \dots, x_n, y) f_i + h(x_1, \dots, x_n, y)(yf - 1).$$

Let y^m be the highest power of y appearing in any of the h_i. Multiplying by f^m and reducing modulo $\langle yf - 1 \rangle$, we get polynomials $g_i \in \Bbbk[x_1, \dots, x_n]$ such that

$$f^m \equiv \sum_{i=1}^{r} g_i f_i \mod \langle yf - 1 \rangle.$$

Since the natural homomorphism

$$\Bbbk[x_1, \dots, x_n] \to \Bbbk[x_1, \dots, x_n, y]/\langle yf - 1 \rangle,\ x_i \mapsto \overline{x}_i,$$

is injective, we actually have

$$f^m = \sum_{i=1}^{r} g_i f_i \in \Bbbk[x_1, \dots, x_n]. \quad \square$$

Corollary 1.6.3. *If $\mathbb{k} = \overline{\mathbb{k}}$ is algebraically closed, then* I *and* V *define a one-to-one correspondence*

$$\{\textit{algebraic subsets of } \mathbb{A}^n(\mathbb{k})\} \underset{\mathrm{V}}{\overset{\mathrm{I}}{\rightleftarrows}} \{\textit{radical ideals of } \mathbb{k}[x_1, \dots, x_n]\}.$$

□

The weak version of the Nullstellensatz adresses the basic **problem of solvability**: Given $f_1, \dots, f_r \in \mathbb{k}[x_1, \dots, x_n]$, the system

$$f_1(x_1, \dots, x_n) = 0, \dots, f_r(x_1, \dots, x_n) = 0$$

fails to have a solution over the algebraic closure $\overline{\mathbb{k}}$ iff $1 \in \langle f_1, \dots, f_r \rangle$. The trick of Rabinowitch allows us to discuss a related problem:

Corollary 1.6.4 (Radical Membership). *Let $\mathbb{k}$ be an arbitrary field, let $I \subset \mathbb{k}[x_1, \dots, x_n]$ be an ideal, and let $f \in \mathbb{k}[x_1, \dots, x_n]$ be a polynomial. Then:*

$$f \in \operatorname{rad} I \iff 1 \in J = \langle I, yf - 1 \rangle \subset \mathbb{k}[x_1, \dots, x_n, y],$$

where y is an extra variable.

Proof. The implication from right to left is clear from Step 2 of the proof of Theorem 1.6.2. For the converse implication, let $f \in \operatorname{rad} I$. Then $f^m \in I \subset J$ for some $m \geq 1$. Since $yf - 1 \in J$ as well, we get, as desired:

$$1 = y^m f^m - (y^m f^m - 1) = y^m f^m - (yf - 1) \sum_{i=1}^{m-1} y^i f^i \in J.$$

□

Hilbert's Nullstellensatz is fundamental to the geometry-algebra dictionary. We will apply it to translate geometric statements into statements on ideals $I \subset \mathbb{k}[x_1, \dots, x_n]$ or, in turn, statements on quotient rings $\mathbb{k}[x_1, \dots, x_n]/I$. Here is, for instance, a result which extends the weak version of the Nullststellensatz in that it characterizes systems of polynomial equations with at most finally many solutions:

Exercise* 1.6.5. Let $I \subset \mathbb{k}[x_1, \dots, x_n]$ be an ideal, and let $\overline{\mathbb{k}}$ be the algebraic closure of $\mathbb{k}$. Show that the following are equivalent:

1. The locus of zeros of I in $\mathbb{A}^n(\overline{\mathbb{k}})$ is a finite set of points (or empty).
2. For each i, $1 \leq i \leq n$, there is a nonzero polynomial in $I \cap \mathbb{k}[x_i]$.
3. The $\mathbb{k}$-vector space $\mathbb{k}[x_1, \dots, x_n]/I$ has finite dimension. □

How to decide algorithmically whether an ideal $I \subset \mathbb{k}[x_1, \dots, x_n]$ contains 1 or whether it satisfies conditions 2 and 3 above will be explained in Sections 2.3 and 2.4.

Exercise 1.6.6. Show that every algebraic subset of $\mathbb{A}^n(\mathbb{R})$ can be defined by a single polynomial equation. Give examples of ideals $I \subset \mathbb{R}[x_1, \dots, x_n]$ whose locus of zeros in $\mathbb{A}^n(\mathbb{R})$ is finite though $\dim_{\mathbb{R}} \mathbb{R}[x_1, \dots, x_n]/I = \infty$. □

See also Exercise 1.12.2.

1.7 Irreducible Components

The algebraic set $\mathrm{V}(xz, yz) \subset \mathbb{A}^3(\mathbb{R})$ in Example 1.2.10 decomposes as the union of the xy-plane $\mathrm{V}(z)$ and the z-axis $\mathrm{V}(x, y)$ which are, again, algebraic sets. In this section, we will show that every algebraic set is the union of finitely many algebraic sets which "cannot be decomposed any further".

Definition 1.7.1. A nonempty algebraic set $A \subset \mathbb{A}^n(\Bbbk)$ is **reducible** if it can be expressed as the union $A = A_1 \cup A_2$ of algebraic sets $A_1, A_2 \subset \mathbb{A}^n(\Bbbk)$ properly contained in A. Otherwise, A is called **irreducible**, or a **subvariety** of $\mathbb{A}^n(\Bbbk)$. The empty set is not considered to be irreducible. An **affine variety** is a subvariety of some $\mathbb{A}^n(\Bbbk)$. □

Proposition 1.7.2. *let $A \subset \mathbb{A}^n(\Bbbk)$ be an algebraic set. Then the following conditions are equivalent:*

1. A is irreducible.
2. $\mathrm{I}(A)$ is a prime ideal.
3. $\Bbbk[x_1, \dots, x_n]/\mathrm{I}(A)$ is an integral domain.

Proof. $1 \implies 2$: Suppose that A is irreducible. Then $A \neq \emptyset$, so that $\mathrm{I}(A)$ is a proper ideal. Let $f, g \in \Bbbk[x_1, \dots, x_n]$ such that $fg \in \mathrm{I}(A)$. Then

$$A = (A \cap \mathrm{V}(f)) \cup (A \cap \mathrm{V}(g)).$$

Since A is irreducible, we have either $A = A \cap \mathrm{V}(f)$ or $A = A \cap \mathrm{V}(g)$. Hence, either $f \in \mathrm{I}(A)$ or $g \in \mathrm{I}(A)$.

$2 \implies 1$: Now, suppose that $\mathrm{I}(A)$ is a prime ideal. Then $\mathrm{I}(A)$ is a proper ideal, so that $A = \mathrm{V}(\mathrm{I}(A)) \neq \emptyset$. Let $A_1, A_2 \subset \mathbb{A}^n(\Bbbk)$ be algebraic sets such that $A = A_1 \cup A_2$. Then $\mathrm{I}(A) = \mathrm{I}(A_1) \cap \mathrm{I}(A_2)$. Since $\mathrm{I}(A)$ is a prime ideal, we have either $\mathrm{I}(A) = \mathrm{I}(A_1)$ or $\mathrm{I}(A) = \mathrm{I}(A_2)$ (apply part 2 of Exercise 1.3.4). Hence, either $A = A_1$ or $A = A_2$.

$3 \iff 2$: This is a special case of Exercise 1.3.6, 2. □

Clearly, every prime ideal is a radical ideal.

Corollary 1.7.3. *If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, then* I *and* V *define a one-to-one correspondence*

$$\begin{tikzcd} \{\textit{subvarieties of } \mathbb{A}^n(\Bbbk)\} \arrow[r, shift left, "\mathrm{I}"] & \{\textit{prime ideals of } \Bbbk[x_1, \dots, x_n]\}. \arrow[l, shift left, "\mathrm{V}"] \end{tikzcd}$$

□

Example 1.7.4. If $\Bbbk$ is infinite, then $\mathrm{I}(\mathbb{A}^n(\Bbbk)) = \langle 0 \rangle$ by Exercise 1.2.1. In particular, $\mathrm{I}(\mathbb{A}^n(\Bbbk))$ is a prime ideal, so that $\mathbb{A}^n(\Bbbk)$ is irreducible. In contrast, if $\mathbb{F}_q$ is the finite field with q elements, then $\mathbb{A}^n(\mathbb{F}_q)$ is reducible since it consists of finitely many points. Accordingly, the ideal $\mathrm{I}(\mathbb{A}^n(\Bbbk))$ is not prime. In fact, as we will show in Exercise 2.9.1,

$$\mathrm{I}(\mathbb{A}^n(\mathbb{F}_q)) = \langle x_1^q - x_1, \dots, x_n^q - x_n \rangle.$$

□

Example 1.7.5. If $\Bbbk$ is infinite, every linear subvariety A of $\mathbb{A}^n(\Bbbk)$ is irreducible. Indeed, in this case, $\Bbbk[x_1, \ldots, x_n]/\mathrm{I}(A) \cong \Bbbk[x_{i_1}, \ldots, x_{i_d}]$ for some d and some $i_1, \ldots, i_d$. □

Example 1.7.6. Let $\Bbbk$ be infinite. Using its parametrization, we show that the twisted cubic curve

$$C = \mathrm{V}(y - x^2, z - x^3) = \{(a, a^2, a^3) \mid a \in \Bbbk\} \subset \mathbb{A}^3(\Bbbk),$$

is irreducible. In fact, we show that the vanishing ideal of C is prime. For this, if $f, g \in \Bbbk[x, y, z]$ such that $f \cdot g \in \mathrm{I}(C)$, set

$$F(t) = f(t, t^2, t^3) \text{ and } G(t) = g(t, t^2, t^3) \in \Bbbk[t].$$

Since $\Bbbk$ is infinite, we have $F \cdot G = 0$, so that either $F = 0$ or $G = 0$. Hence, either $f \in \mathrm{I}(C)$ or $g \in \mathrm{I}(C)$. □

Since a set consisting of a single point is irreducible, its vanishing ideal is a prime ideal. In fact, even more is true:

Remark 1.7.7. If $p = (a_1, \ldots, a_n) \in \mathbb{A}^n(\Bbbk)$ is a point, every polynomial $f \in \Bbbk[x_1, \ldots, x_n]$ can be written as a polynomial in the $x_i - a_i$:

$$f = f(p) + \text{terms of degree} \geq 1 \text{ in the } x_i - a_i. \tag{1.2}$$

Indeed, this is the **Taylor expansion of f at p** which is obtained by substituting the $(x_i - a_i) + a_i$ for the x_i in f and expanding the resulting expression. It is clear from (1.2) that the vanishing ideal

$$\mathrm{I}(p) := \mathrm{I}(\{p\}) = \langle x_1 - a_1, \ldots, x_n - a_n \rangle$$

is the kernel of the evaluation map

$$\Bbbk[x_1, \ldots, x_n] \to \Bbbk, \; f \mapsto f(p).$$

Hence, $\Bbbk[x_1, \ldots, x_n]/\mathrm{I}(p) \cong \Bbbk$ by the homomorphy theorem. In particular, $\Bbbk[x_1, \ldots, x_n]/\mathrm{I}(p)$ is a field, so that $\mathrm{I}(p)$ is a maximal ideal. □

Conversely, the following holds:

Proposition 1.7.8. *If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, every maximal ideal of $\Bbbk[x_1, \ldots, x_n]$ is of the form $\langle x_1 - a_1, \ldots, x_n - a_n \rangle$ for some $a_1, \ldots, a_n \in \Bbbk$.*

Proof. Let $\mathfrak{m} \subsetneq \Bbbk[x_1, \ldots, x_n]$ be a maximal ideal. Then $\mathrm{V}(\mathfrak{m}) \neq \emptyset$ by the weak version of the Nullstellensatz. If $p = (a_1, \ldots, a_n) \in \mathrm{V}(\mathfrak{m})$ is a point, we have $\mathfrak{m} \subset \mathrm{I}(p) = \langle x_1 - a_1, \ldots, x_n - a_n \rangle$. In fact, $\mathfrak{m} = \mathrm{I}(p)$ since $\mathfrak{m}$ is maximal. □

Corollary 1.7.9. *If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, then* I *and* V *define a one-to-one correspondence*

$$\{\textit{points of } \mathbb{A}^n(\Bbbk)\} \underset{\mathrm{V}}{\overset{\mathrm{I}}{\rightleftarrows}} \{\textit{maximal ideals of } \Bbbk[x_1, \ldots, x_n]\}.$$

□

If $\Bbbk$ is not necessarily algebraically closed, the maximal ideals of $\Bbbk[x_1, \dots, x_n]$ can be described as follows:

Exercise* 1.7.10. Let $\overline{\Bbbk}$ be the algebraic closure of $\Bbbk$, and let G be the Galois group of $\overline{\Bbbk}$ over $\Bbbk$. Show:

1. Let $p = (a_1, \dots, a_n) \in \mathbb{A}^n(\overline{\Bbbk})$ be a point, and let $\mathfrak{m}_p$ be the kernel of the evaluation map
$$\Bbbk[x_1, \dots, x_n] \to \overline{\Bbbk},\ f \mapsto f(p).$$
Then $\mathfrak{m}_p$ is a maximal ideal of $\Bbbk[x_1, \dots, x_n]$. Moreover, its locus of zeros in $\mathbb{A}^n(\overline{\Bbbk})$ is the orbit of p under the natural action of G on $\mathbb{A}^n(\overline{\Bbbk})$. We, then, say that the points of this locus are **pairwise conjugate** over $\Bbbk$.
2. Every maximal ideal of $\Bbbk[x_1, \dots, x_n]$ is of type $\mathfrak{m}_p$ for some $p \in \mathbb{A}^n(\overline{\Bbbk})$. □

Example 1.7.11. The principal ideal generated by x^2+1 in $\mathbb{R}[x]$ is maximal, and its locus of zeros in $\mathbb{A}^1(\mathbb{C})$ is $\{\pm i\}$. □

We, now, establish the main result of this section:

Theorem-Definition 1.7.12. *Every algebraic set $A \subset \mathbb{A}^n(\Bbbk)$ can be written as a finite union*
$$A = V_1 \cup \dots \cup V_s$$
*of irreducible algebraic sets V_i. We may, in fact, achieve that this decomposition is **minimal** in the sense that $V_i \not\supset V_j$ for $i \neq j$. The V_i are, then, uniquely determined up to order and are called the **irreducible components** of A.*

Proof. The *existence* part of the proof is a typical example of Noetherian induction. Expressed in geometric terms, the maximal condition for ideals in the Noetherian ring $\Bbbk[x_1, \dots, x_n]$ reads that every nonempty collection of algebraic subsets of $\mathbb{A}^n(\Bbbk)$ has a minimal element with respect to inclusion. Using this, we show that the collection Γ of all algebraic subsets of $\mathbb{A}^n(\Bbbk)$ which cannot be written as a finite union of irreducible algebraic sets is empty.

Suppose that $\Gamma \neq \emptyset$. Then Γ has a minimal element A which, by the very definition of Γ, must be reducible. That is, $A = A_1 \cup A_2$ for some algebraic sets $A_1, A_2 \subsetneq A$. Due to the minimality of A, both A_1 and A_2 can be written as a finite union of irreducible algebraic sets. Then the same is true for A, a contradiction to $A \in \Gamma$.

We conclude that every algebraic set $A \subset \mathbb{A}^n(\Bbbk)$ can be written as a finite union of irreducible algebraic sets. Throwing away superfluous sets if necessary, we get a minimal decomposition, as required.

To show *uniqueness*, let
$$A = V_1 \cup \dots \cup V_s = V_1' \cup \dots \cup V_t'$$
be two minimal decompositions. Then, for each i, we have

$$V_i = V_i \cap A = V_i \cap (V_1' \cup \cdots \cup V_t') = (V_i \cap V_1') \cup \cdots \cup (V_i \cap V_t').$$

Since V_i is irreducible, we must have $V_i = V_i \cap V_j'$ for some j, so that $V_i \subset V_j'$. The same argument yields an inclusion $V_j' \subset V_k$ for some k. By minimality, $i = k$ and, thus, $V_i = V_j'$. So every V_i occurs as one of the V_j' which implies that $s \leq t$. Similarly, we get $t \leq s$. Uniqueness up to order follows. □

Exercise* 1.7.13. Show:

1. Every proper algebraic subset of $\mathbb{A}^1(\Bbbk)$ is a finite set of points (or empty).
2. If $f, g \in \Bbbk[x, y]$ are polynomials without a common factor, then
$$\mathrm{V}(f, g) = \mathrm{V}(f) \cap \mathrm{V}(g) \subset \mathbb{A}^2(\Bbbk)$$
is a finite set of points (or empty).
Hint. Prove that f and g are coprime in the PID $\Bbbk(x)[y]$, and deduce that there exist $a, b \in \Bbbk(x)[y]$ such that $af + bg = 1$.
3. Every proper algebraic subset of $\mathbb{A}^2(\Bbbk)$ is a finite union of points and (irreducible) curves (or empty). □

1.8 Primary Decomposition

The Nullstellensatz allows us to rephrase Theorem 1.7.12 in algebraic terms as follows: If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, the radical of every ideal $I \subset \Bbbk[x_1, \ldots, x_n]$ has a unique **minimal prime decomposition**. That is, rad I can be uniquely written as the intersection of finitely many prime ideals:

$$\operatorname{rad} I = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_s,$$

where $\mathfrak{p}_i \not\supset \mathfrak{p}_j$ for $i \neq j$. This is a purely algebraic result which, in fact, can be proved by purely algebraic means (there is no need to translate the Noetherian condition into geometry and apply the Nullstellensatz). In what follows, we present the original argument of Emmy Noether which works for any Noetherian ring R. In fact, the argument applies to arbitrary ideals of R and not just to radical ideals. The resulting decomposition has, then, to be of a more general type, however, since the intersection of prime ideals is necessarily a radical ideal.

Definition 1.8.1. A proper ideal $\mathfrak{q}$ of a ring R is a **primary ideal** if $f, g \in R$ and $fg \in \mathfrak{q}$ implies $f \in \mathfrak{q}$ or $g \in \operatorname{rad} \mathfrak{q}$. □

Clearly, every prime ideal is primary.

Proposition 1.8.2. *Let R be a ring.*

1. *If $\mathfrak{q}$ is a primary ideal of R, then $\mathfrak{p} := \operatorname{rad} \mathfrak{q}$ is the smallest prime ideal containing $\mathfrak{q}$. We refer to this fact by saying that $\mathfrak{q}$ is **$\mathfrak{p}$-primary**.*
2. *A finite intersection of $\mathfrak{p}$-primary ideals is $\mathfrak{p}$-primary.* □

Exercise* 1.8.3. Prove Proposition 1.8.2. □

Definition 1.8.4. Let I be an ideal of a ring R. A **primary decomposition** of I is an expression of I as a finite intersection of primary ideals, say

$$I = \mathfrak{q}_1 \cap \cdots \cap \mathfrak{q}_t.$$

The decomposition is called **minimal** if the radicals $\operatorname{rad} \mathfrak{q}_i$ are all distinct, and $\mathfrak{q}_i \not\supset \bigcap_{j \neq i} \mathfrak{q}_j$ for all i. □

Theorem 1.8.5. *Every proper ideal of a Noetherian ring R has a minimal primary decomposition.*

Proof. We proceed in three steps.

Step 1. In analogy to Definition 1.7.1, we say that an ideal of R is *irreducible* if it is not the intersection of two strictly larger ideals. The algebraic version of the proof of Theorem 1.7.12 shows that every ideal of the Noetherian ring R can be written as a finite intersection of irreducible ideals.

Step 2. Let $I \subsetneq R$ be an irreducible ideal. We prove that I is primary. For this, let $f, g \in R$ such that $fg \in I$ and $f \notin I$. To show that $g \in \operatorname{rad} I$, observe that we have a chain of ideals

$$I : g \subset I : g^2 \subset \cdots .$$

By the ascending chain condition, $I : g^m = I : g^{m+1}$ for some $m \geq 1$. Then

$$I = (I : g^m) \cap \langle I, g^m \rangle$$

by Exercise 1.3.3. Since $fg \in I$, also $fg^m \in I$, so that $f \in I : g^m$. This implies that $I \neq I : g^m$ since $f \notin I$. Taking into account that I is irreducible, we must have $I = \langle I, g^m \rangle$, so that $g^m \in I$. Hence, $g \in \operatorname{rad} I$.

Step 3. Let $I \subsetneq R$ be an arbitrary ideal. By Steps 1 and 2, there is a primary decomposition of I. If two of the primary ideals occuring in this decomposition have the same radical, we may replace them by their intersection which is primary by Proposition 1.8.2. Continuing in this way, all primary ideals will eventually have distinct radicals. Throwing away superfluous primary ideals if necessary, we get a minimal primary decomposition of I. □

Not all the ideals occuring in a primary decomposition of an ideal I are uniquely determined by I:

Example 1.8.6. The ideal $\langle xy, y^2 \rangle \subset \Bbbk[x, y]$ admits, for instance, the following minimal primary decompositions:

$$\langle xy, y^2 \rangle = \langle y \rangle \cap \langle x, y^2 \rangle = \langle y \rangle \cap \langle x^2, xy, y^2 \rangle.$$

Note that both $\langle x, y^2 \rangle$ and $\langle x^2, xy, y^2 \rangle$ are $\langle x, y \rangle$-primary. Furthermore, the prime ideal $\langle x, y \rangle$ contains the prime ideal $\langle y \rangle$. □

Theorem 1.8.7 (1st Uniqueness Theorem). *Let I be a proper ideal of a Noetherian ring R, and let $I = \bigcap_{i=1}^t \mathfrak{q}_i$ be a minimal primary decomposition of I. Then the radicals $\mathfrak{p}_i = \operatorname{rad} \mathfrak{q}_i$ are precisely the prime ideals occuring in the set of ideals $I : f$, $f \in R$.*

Proof. See Exercise 1.9.4. □

Remark-Definition 1.8.8. In the situation of the 1st uniqueness theorem, we see, in particular, that the $\mathfrak{p}_i$ only depend on I (and not on the particular minimal primary decomposition). We call each $\mathfrak{p}_i$ an **associated prime** of I. We say that $\mathfrak{p}_i$ is a **minimal associated prime** of I if $\mathfrak{p}_i \not\supset \mathfrak{p}_j$ for all $j \neq i$. Otherwise, $\mathfrak{p}_i$ is called an **embedded associated prime** of I. If, say, $\mathfrak{p}_1 \dots, \mathfrak{p}_s$ are the mimimal associated primes of I, then

$$\operatorname{rad} I = \mathfrak{p}_1 \cap \dots \cap \mathfrak{p}_s$$

is the uniquely determined **minimal prime decomposition** of rad I (defined as at the beginning of this section).

Any primary ideal occurring in one of the minimal primary decompositions of I is called a **primary component** of I. It is called an **isolated component** of I if its radical is a minimal associated prime of I, and an **embedded component** of I, otherwise. □

The names isolated and embedded come from the geometric interpretation: If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, the minimal associated primes of an ideal $I \subset \Bbbk[x_1, \dots, x_n]$ correspond to the irreducible components of $\mathrm{V}(I)$, and the embedded associated primes to subvarieties of these.

Theorem 1.8.9 (2nd Uniqueness Theorem). *Let I be a proper ideal of a Noetherian ring. Then the isolated primary components of I are uniquely determined by I.*

Proof. We will show this in Exercise 4.5.6. □

Example 1.8.10. If R is a UFD, and $f \in R$ is a nonzero nonunit, then all the associated primes of $\langle f \rangle$ are minimal. Indeed, if

$$f = u \cdot f_1^{\mu_1} \cdots f_s^{\mu_s}$$

is the decomposition of f into distinct irreducible factors, the minimal primary decomposition is

$$\langle f \rangle = \langle f_1^{\mu_1} \rangle \cap \dots \cap \langle f_s^{\mu_s} \rangle.$$

Note that historically, the concept of primary decomposition grew out from the search for some useful generalization of unique factorization. See Eisenbud (1995), Section 1.1. □

If I is a proper ideal of a ring R, and $\mathfrak{p} \subset R$ is a prime ideal containing I, we say that $\mathfrak{p}$ is a **minimal prime of *I*** if there is no prime ideal $\mathfrak{q}$ of R such that $I \subset \mathfrak{q} \subsetneq p$. A minimal prime of the zero ideal of R is also called a **minimal prime of *R***.

Proposition 1.8.11. *Let I be a proper ideal of a Noetherian ring R. Then every prime ideal containing I contains a minimal associated prime of I. Thus, the minimal associated primes of I are precisely the minimal primes of I.*

Proof. Let $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ be the minimal associated primes of I. If $\mathfrak{p} \supset I$ is a prime ideal, then $\mathfrak{p} = \operatorname{rad} \mathfrak{p} \supset \operatorname{rad} I = \bigcap_{i=1}^{s} \mathfrak{p}_i$. Hence, we must have $\mathfrak{p} \supset \mathfrak{p}_i$ for some i by part 2 of Exercise 1.3.4. □

1.9 Removing Algebraic Sets

The ideal $I = \langle xz, yz \rangle \subset \mathbb{R}[x, y, z]$ in Example 1.2.10 is the intersection of the prime ideals $\langle z \rangle$ and $\langle x, y \rangle$. In particular, I is a radical ideal. Geometrically, I defines the union of the xy-plane and the z-axis. If we remove the xy-plane, the remaining set is a punctured line, which is not an algebraic subset of $\mathbb{A}^3(\Bbbk)$ (see Exercise 1.7.13). In this section, we show how to describe the smallest algebraic set containing the difference of two algebraic sets.

We need the following notation. If I, J are two ideals of a ring R, the set

$$I : J^\infty := \{f \in R \mid fJ^m \subset I \text{ for some } m \geq 1\} = \bigcup_{m=1}^{\infty} (I : J^m)$$

is an ideal of R. It is called the **saturation** of I with respect to J. If g is a single element of R, we usually write $I : g^\infty$ instead of $I : \langle g \rangle^\infty$.

In any case, we have an ascending chain of ideals

$$I : J \subset I : J^2 \subset I : J^3 \subset \cdots \subset I : J^\infty.$$

Thus, if R is Noetherian, we have $I : J^m = I : J^{m+1} = I : J^\infty$ for some $m \geq 1$ by the ascending chain condition.

Theorem 1.9.1. *Let $\Bbbk = \overline{\Bbbk}$ be algebraically closed, and let I, J be ideals of $\Bbbk[x_1, \ldots, x_n]$. Then*

$$\overline{\mathrm{V}(I) \setminus \mathrm{V}(J)} = \mathrm{V}(I : J^\infty).$$

If I is a radical ideal, then

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

Proof. For the first statement, let $I = \bigcap_{i=1}^{t} \mathfrak{q}_i$ be a primary decomposition. To show the desired equality, we proceed in four steps, writing $I : J^\infty$ as the intersection of the $\mathfrak{q}_i : J^\infty$.

Step 1. If $J^m \subset \mathfrak{q}_i$ for some $m \geq 1$, then $\mathfrak{q}_i : J^\infty = \Bbbk[x_1, \ldots, x_n]$ by part 1 of Exercise 1.3.3. If $J^m \not\subset \mathfrak{q}_i$ for all $m \geq 1$, then $\mathfrak{q}_i : J^\infty = \mathfrak{q}_i$. Indeed, if $f \in \mathfrak{q}_i : J^\infty$, then $fJ^k \subset \mathfrak{q}_i$ for some $k \geq 1$, so that $f \in \mathfrak{q}_i$ by part 2 of Proposition 1.9.2 below. This shows that $\mathfrak{q}_i : J^\infty \subset \mathfrak{q}_i$. The opposite inclusion is clear.

Step 2. We have $J^m \not\subset \mathfrak{q}_i$ for all $m \geq 1$ iff $\mathrm{V}(J) \not\supset \mathrm{V}(\mathfrak{q}_i)$. Indeed, if $\mathrm{V}(J) \supset \mathrm{V}(\mathfrak{q}_i)$, then $J \subset \operatorname{rad} J \subset \operatorname{rad} \mathfrak{q}_i$ by Hilbert's Nullstellensatz, so that $J^m \subset \mathfrak{q}_i$ for some $m \geq 1$ by part 1 of Proposition 1.9.2 below. This shows the implication from left to right. The converse implication is clear since $\mathrm{V}(J^m) = \mathrm{V}(J)$ for all $m \geq 1$.

Step 3. If $\mathrm{V}(J) \not\supset \mathrm{V}(\mathfrak{q}_i)$, then

$$\mathrm{V}(\mathfrak{q}_i) = \overline{\mathrm{V}(\mathfrak{q}_i) \setminus \mathrm{V}(J)} \cup (\mathrm{V}(\mathfrak{q}_i) \cap \mathrm{V}(J)) = \overline{\mathrm{V}(\mathfrak{q}_i) \setminus \mathrm{V}(J)}$$

since $\mathrm{V}(\mathfrak{q}_i) = \mathrm{V}(\operatorname{rad} \mathfrak{q}_i)$ is irreducible.

Step 4. By Exercise 1.3.3 and Steps 1 and 2,

$$\begin{aligned} I : J^\infty = \bigcap_{i=1}^{t} (\mathfrak{q}_i : J^\infty) &= \left(\bigcap_{\substack{J^m \subset \mathfrak{q}_i \\ \text{for some } m \geq 1}} (\mathfrak{q}_i : J^\infty) \right) \cap \left(\bigcap_{\substack{J^m \not\subset \mathfrak{q}_i \\ \text{for all } m \geq 1}} (\mathfrak{q}_i : J^\infty) \right) \\ &= \bigcap_{\substack{J^m \not\subset \mathfrak{q}_i \\ \text{for all } m \geq 1}} \mathfrak{q}_i = \bigcap_{\mathrm{V}(J) \not\supset \mathrm{V}(q_i)} \mathfrak{q}_i. \end{aligned}$$

Hence, the first statement follows from Step 3:

$$\begin{aligned} \mathrm{V}(I : J^\infty) = \bigcup_{\mathrm{V}(J) \not\supset \mathrm{V}(\mathfrak{q}_i)} \mathrm{V}(\mathfrak{q}_i) &= \bigcup_{\mathrm{V}(J) \not\supset \mathrm{V}(\mathfrak{q}_i)} \overline{\mathrm{V}(\mathfrak{q}_i) \setminus \mathrm{V}(J)} \\ &= \bigcup_{i=1}^{t} \overline{(\mathrm{V}(\mathfrak{q}_i) \setminus \mathrm{V}(J))} = \overline{\mathrm{V}(I) \setminus \mathrm{V}(J)}. \end{aligned}$$

For the second statement, note that a *radical* ideal I can be written as the intersection of *prime* ideals $\mathfrak{q}_i$. The same arguments as above show, then, that

$$\overline{\mathrm{V}(I) \setminus \mathrm{V}(J)} = \mathrm{V}(I : J^\infty) = \mathrm{V}(I : J).$$

□

Proposition 1.9.2. *Let I be an ideal of a Noetherian ring R. Then:*

1. *I contains a power of its radical.*
2. *If $\mathfrak{q} \subset R$ is a primary ideal, and $f \in R$, then $fI \subset \mathfrak{q}$ implies $f \in \mathfrak{q}$ or $I^m \subset \mathfrak{q}$ for some $m \geq 1$.*

Proof. 1. Since R is Noetherian, $\operatorname{rad} I$ is finitely generated, say $\operatorname{rad} I = \langle f_1, \ldots, f_r \rangle$. For each i, we may choose an integer $m_i \geq 1$ such that $f_i^{m_i} \in I$. Let $m = \sum_{i=1}^{r}(m_i - 1) + 1$. Then $(\operatorname{rad} I)^m$ is generated by the products $f_1^{k_1} \cdots f_r^{k_r}$, where $\sum_{i=1}^{r} k_i = m$. From the definition of m, we must have $k_i \geq m_i$ for at least one i. Hence, all the products lie in I. This shows that $(\operatorname{rad} I)^m \subset I$.

2. The argument is similar to that in part 1. □

Exercise* 1.9.3. Let R be a Noetherian ring, let $\mathfrak{m}$ be a maximal ideal of R, and let I be any ideal of R. Show that the following are equivalent:

1. I is $\mathfrak{m}$-primary.
2. $\text{rad}\, I = \mathfrak{m}$.
3. $\mathfrak{m} \supset I \supset \mathfrak{m}^k$ for some $k \geq 1$. □

Exercise* 1.9.4. Prove Theorem 1.8.7.
Hint. As a first step, show that the radicals $\mathfrak{p}_i = \text{rad}\, \mathfrak{q}_i$ are precisely the prime ideals occuring in the set of ideals $(\text{rad}\, I) : f,\ f \in R$. □

How to compute ideal quotients and saturation will be a topic of Section 2.4.

1.10 Modules

In this section, we set the geometry-algebra dictionary aside and introduce modules which are to rings what vector spaces are to fields. We will treat some basic operations on modules, including the tensor product.

An ideal I of a ring R and its quotient ring R/I are both examples of modules. By speaking of modules, we may often formulate definitions and results such that they apply to ideals and quotient rings at the same time. Later in the book, we will encounter further examples of modules which arise naturally in algebraic geometry. Most notably, the syzygies introduced in Chapter 2 and the Kähler differentials treated in Chapter 8 form modules.

Let R be a ring.

Definition 1.10.1. A **module** over R, or an ***R*-module**, is an additively written group M, together with a map $R \times M \to M$, written $(r, m) \mapsto rm$, such that for all $r, s \in R$ and $m, n \in M$ the following hold:

$$r(sm) = (rs)m,\ r(m+n) = rm + rn,\ (r+s)m = rm + sm,\ 1m = m.$$ □

Example 1.10.2. 1. If I is an ideal of R, then I and R/I are R-modules. In particular, R itself is an R-module.
2. Every Abelian group G is a $\mathbb{Z}$-module: if $g \in G$, and $n \in \mathbb{Z}$ is positive (or zero or negative), define ng to be $g + \cdots + g$ (or 0 or (-g) $+ \ldots +$ (-g)). □

The notion of a linear map extends from vector spaces to modules:

Definition 1.10.3. Let M and N be R-modules. A map $\phi : M \to N$ is called an ***R*-module homomorphism**, or an ***R*-linear map**, if

$$\phi(m+n) = \phi(m) + \phi(n) \text{ and } \phi(rm) = r\phi(m)$$

for all $r \in R$ and $m, n \in M$. □

As usual, a homomorphism which is injective (or surjective or bijective) is called a **monomorphism** (or **epimorphism** or **isomorphism**). Also, we write $M \cong N$ and call M and N **isomorphic** if there exists an isomorphism $M \to N$. Note that an R-module homomorphism is an isomorphism iff it admits an inverse homomorphism.

Remark 1.10.4. 1. If M and N are R-modules, the set

$$\operatorname{Hom}_R(M,N) := \{R\text{-module homomorphisms from } M \text{ to } N\}$$

ia again an R-module: if $r \in R$, and $\phi, \psi \in \operatorname{Hom}_R(M,N)$, define

$$(\phi+\psi)(m) = \phi(m) + \psi(m) \text{ and } (r\phi)(m) = r\phi(m)$$

for all $m \in M$.

2. Given R-module homomorphisms $\alpha : M' \to M$ and $\beta : N \to N''$, we obtain induced R-module homomorphisms

$$\widetilde{\alpha} : \operatorname{Hom}(M,N) \to \operatorname{Hom}(M',N) \text{ and } \widetilde{\beta} : \operatorname{Hom}(M,N) \to \operatorname{Hom}(M,N'')$$

by setting

$$\widetilde{\alpha}(\phi) = \phi \circ \alpha \text{ and } \widetilde{\beta}(\phi) = \beta \circ \phi$$

for all $\phi \in \operatorname{Hom}(M,N)$. □

Extending the notions of ideals $I \subset R$ and quotient rings R/I, we get the notions of submodules $I \subset M$ and quotient modules M/I. That is, a **submodule** of an R-module M is an additive subgroup I of M such that if $r \in R$ and $m \in I$, then $rm \in I$. Then I inherits an R-module structure from M, and we have the **quotient module** M/I together with the **canonical projection** $M \to M/I$ (obtained as in Definition 1.3.5).

If $\phi : M \to N$ is an R-module homomorphism, its **kernel**

$$\ker \phi := \{m \in M \mid \phi(m) = 0\} \subset M$$

is a submodule of M, and its **image**

$$\operatorname{im} \phi := \phi(M) \subset N$$

is a submodule of N. Its **cokernel**

$$\operatorname{coker} \phi := N/\operatorname{im} \phi$$

is a quotient module of N.

Exercise* 1.10.5 (Homomorphy Theorem). Let $\phi : M \to N$ be an R-module homomorphism. If I is a submodule of M contained in $\ker \phi$ and $\pi : M \to M/I$ is the canonical projection, show that there exists a unique R-module homomorphism $\overline{\phi} : M/I \to N$ such that $\overline{\phi} \circ \pi = \phi$. That is, the following diagram commutes:

$$\begin{tikzcd} M \arrow[rr, "\pi"] \arrow[dr, "\phi"'] & & M/I \arrow[dl, dashed, "\overline{\phi}"] \\ & N & \end{tikzcd}$$

In particular, taking $I = \ker \phi$, we get

$$M/\ker \phi \cong \operatorname{im} \phi.$$

□

If X is any subset of an R-module M, we write $\langle X \rangle$ for the smallest submodule of M containing X. Then $\langle X \rangle$ consists of all R-linear combinations of elements of X and is called the **submodule generated by *X***. The terminology and notation introduced for ideals in this context will be used for modules as well. In particular, we say that M **is finitely generated** if $M = \langle m_1, \dots, m_k \rangle$ for some $m_1, \dots, m_k \in M$.

Most of the operations on ideals considered in Section 1.3 carry over to submodules of M. For instance, the intersection $\bigcap_\lambda I_\lambda$ of a family $\{I_\lambda\}$ of submodules of M is a submodule of M. The **sum** $\sum_\lambda I_\lambda$ of the $\{I_\lambda\}$ is the submodule generated by the union $\bigcup_\lambda I_\lambda$.

Exercise* 1.10.6 (Isomorphy Theorems). Let N_1, N_2 be submodules of an R-module M.

1. If $N_1 \subset N_2$, show that
$$(M/N_1)/(N_2/N_1) \cong M/N_2.$$
2. Show that
$$(N_1 + N_2)/N_1 \cong N_2/(N_1 \cap N_2).$$

□

The **direct sum** of two R-modules M and N is the set

$$M \oplus N := \{(m, n) \mid m \in M, n \in N\},$$

together with the module structure obtained by setting

$$(m_1, n_1) + (m_2, n_2) = (m_1 + m_2, n_1 + n_2) \ \text{ and } \ r(m, n) = (rm, rn).$$

In the same way, we get the direct sum $M_1 \oplus \dots \oplus M_k$ of any finite set of R-modules $M_1, \dots, M_k$. Specifically, R^k denotes the direct sum of k copies of R. More generally, we can define the **direct sum** $\bigoplus_\lambda M_\lambda$ of any family $\{M_\lambda\}$ of R-modules; it consists of the tuples (m_λ) such that $m_\lambda \in M_\lambda$ for all λ and all but finitely many m_λ are zero. In contrast, the **direct product** $\prod_{\lambda \in \Lambda} M_\lambda$ consists of *all* tuples (m_λ) satisfying $m_\lambda \in M_\lambda$ for all λ.

A module F over R is **free** if it is isomorphic to a direct sum of copies of R. Equivalently, F admits a basis in the sense of linear algebra. That is, F admits a set of generators, also called **free generators**, which are R-linearly independent. By convention, also the zero module is free.

As for vector spaces, if F admits a finite basis, the number of basis elements is independent of the choice of basis. It is called the **rank** of F, written $\operatorname{rank} F$. If F is a free R-module of rank k with a fixed basis, we think of it as the free R-module R^k with its canonical basis (formed by the column vectors $(1, 0, \ldots, 0)^t, \ldots, (0, \ldots, 0, 1)^t$). That is, we consider the elements of F as column vectors with entries in R. Furthermore, given two such modules with fixed bases, we may regard each homomorphism between them as a matrix with entries in R.

Example 1.10.7. A nonzero ideal I of R is free iff it is a principal ideal generated by a nonzerodivisor. In fact, if $k \geq 2$ and $f_1, \ldots, f_k$ are nonzero elements of I, then $f_1, \ldots, f_k$ are not R-linearly independent. For instance, there are always the nontrivial relations $f_i f_j - f_j f_i = 0$. □

Note that according to our definitions, an R-module M is finitely generated iff it can be written as a quotient module of type R^k/I. Indeed, if $M = \langle m_1, \ldots, m_k \rangle$, consider the (module) epimorphism

$$R^k \to M,\ e_i \mapsto m_i,$$

where the e_i are the canonical basis vectors of R^k, and take I to be the kernel.

Definition 1.10.8. An R-module M is called **Noetherian** if every submodule of M is finitely generated. □

As in Exercise 1.4.5 one shows that M is Noetherian iff the ascending chain condition (respectively, maximal condition) holds for submodules of M.

Exercise* 1.10.9. Let R be a Noetherian ring. Show that every finitely generated R-module is Noetherian.
Hint. Reduce the general case to the case of free R-modules. For free R-modules, use induction on the rank. □

If I, J are two submodules of M, their **submodule quotient** is the set

$$I : J = \{f \in R \mid fJ \subset I\},$$

which is an ideal of R.

Definition 1.10.10. If M is an R-module, the ideal

$$\operatorname{Ann}(M) = 0 : M = \{r \in R \mid rm = 0 \text{ for all } m \in M\}$$

is called the **annihilator of *M***. If $m \in M$ is any element, we write $\operatorname{Ann}(m)$ for the annihilator of $\langle m \rangle$, and call it the **annihilator of *m***. □

Exercise 1.10.11. 1. Determine the annihilator of the $\mathbb{Z}$-module

$$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}.$$

2. If M, N are R-modules, show that

$$\mathrm{Ann}(M \oplus N) = \mathrm{Ann}(M) \cap \mathrm{Ann}(N).$$

□

Given R-modules M, N, P, we say that a map $\Phi : M \times N \to P$ is ***R*-bilinear** if for each $m \in M$ the induced map $N \to P,\ n \mapsto \Phi(m, n)$, is R-linear, and for each $n \in N$ the induced map $M \to P,\ m \mapsto \Phi(m, n)$, is R-linear. Our next result allows us to interprete R-bilinear maps in terms of R-linear maps:

Theorem 1.10.12. *Let M and N be R-modules. There is an R-module T, together with an R-bilinear map $t : M \times N \to T$, such that the following* ***universal property*** *holds: Given any R-module P and any R-bilinear map $\Phi : M \times N \to P$, there is a unique R-linear map $\phi : T \to P$ such that $\phi \circ t = \Phi$. That is, the following diagram commutes:*

$$\begin{tikzcd}
M \times N \arrow[rr, "t"] \arrow[dr, "\Phi"'] & & T \arrow[dl, dashed, "\phi"] \\
& P &
\end{tikzcd}$$

Furthermore, if (T, t) and (T', t') are two pairs satisfying the universal property, there is a unique isomorphism $\psi : T \to T'$ such that $\psi \circ t = t'$.

Proof. The *uniqueness* part of the proof is an application of the universal property: Since both pairs (T, t) and (T', t') satisfy this property, we get unique R-linear maps $\phi : T \to T'$ and $\phi' : T' \to T$ such that the diagrams

$$\begin{tikzcd}
M \times N \arrow[rr, "t"] \arrow[dr, "t'"'] & & T \arrow[dl, dashed, "\phi"] \\
& T' &
\end{tikzcd}
\quad \text{and} \quad
\begin{tikzcd}
M \times N \arrow[rr, "t'"] \arrow[dr, "t"'] & & T' \arrow[dl, dashed, "\phi'"] \\
& {} &
\end{tikzcd}$$

commute. Applying the universal property twice again, we obtain $\phi' \circ \phi = \mathrm{id}_T$ and $\phi \circ \phi' = \mathrm{id}_{T'}$. Thus, ϕ is an isomorphism.

The *existence* is obtained as follows. Regarding $M \times N$ as a set of indices, pick a copy of R for each $(m, n) \in M \times N$, let F be the direct sum of these copies, and write $e_{(m,n)}$ for the canonical basis vector of F corresponding to the index (m, n). Let $I \subset F$ be the submodule generated by elements of the following types:

$$\begin{gathered}
e_{(m+m',n)} - e_{(m,n)} - e_{(m',n)}, \\
e_{(m,n+n')} - e_{(m,n)} - e_{(m,n')}, \\
e_{(rm,n)} - re_{(m,n)}, \\
e_{(m,rn)} - re_{(m,n)},
\end{gathered}$$

where $m, m' \in M$, $n, n' \in N$, and $r \in R$. Let $T = F/I$, and let $t : M \times N \to T$ be the map sending (m, n) to the residue class of $e_{(m,n)} \in F$ modulo I. Then, by construction, t is R-bilinear.

Given an R-module P and a map $\Phi : M \times N \to P$, consider the R-linear map $\widetilde{\Phi} : F \to P$ defined by sending $e_{(m,n)}$ to $\Phi(m,n)$. If Φ is R-bilinear, then $\widetilde{\Phi}$ vanishes on I and induces, thus, an R-linear map $\phi : T \to P$ such that $\phi \circ t = \Phi$. In fact, this condition determines ϕ uniquely. We conclude that the pair (T, t) has the desired properties. □

Definition 1.10.13. In the situation of the theorem, we call T the **tensor product** of M and N over R, denoted

$$M \otimes N := M \otimes_R N := T.$$

Furthermore, we write $m \otimes n$ for the image of $(m,n) \in M \times N$ under t. □

Corollary 1.10.14. *If M, N are R-modules, each element $w \in M \otimes_R N$ can be written as a sum of type*

$$w = \sum_{i=1}^{k} m_i \otimes n_i.$$

Proof. Using the notation of the proof of the theorem, let $f \in F$ be an element representing $w \in F/I$. Then f is a (finite) R -linear combination of the basis vectors $e_{(m,n)}$ of F. The result follows. □

Remark 1.10.15. Given sets of generators X and Y for M and N, respectively, the tensor product $M \otimes N$ is generated by the elements of type $x \otimes y$, where $x \in X$ and $y \in Y$. In particular, if M and N are finitely generated, then so is $M \otimes N$. □

From this point on, we do not make use of the explicit construction of the tensor product. Instead, we work with its universal property. In the same way, we deal with the **tensor product** $M_1 \otimes \dots \otimes M_k$ of more than two R-modules $M_1, \dots, M_k$: In analogy to the case of two R-modules, this tensor product is defined by asking a **universal property** for ***k*-linear** maps over R.

Proposition 1.10.16. *Let M, N, P be R-modules. There are unique isomorphisms*

1. $M \otimes N \to N \otimes M$,
2. $(M \otimes N) \otimes P \to M \otimes (N \otimes P) \to M \otimes N \otimes P$,
3. $(M \oplus N) \otimes P \to (M \otimes P) \oplus (N \otimes P)$, *and*
4. $R \otimes M \to M$

such that, respectively,

1. $m \otimes n \mapsto n \otimes m$,
2. $(m \otimes n) \otimes p \mapsto m \otimes (n \otimes p) \mapsto m \otimes n \otimes p$,
3. $(m \oplus n) \otimes p \mapsto (m \otimes p, n \otimes p)$, *and*
4. $r \otimes m \mapsto m$.

Proof. All isomorphisms are obtained by applying the universal property. As an example, we show 3.

The map $(M \oplus N) \times P \to (M \otimes P) \oplus (N \otimes P)$, $((m,n),p) \mapsto (m \otimes p, n \otimes p)$ is R-bilinear in (m,n) and p. It induces, thus, an R-module homomorphism $(M \oplus N) \otimes P \to (M \otimes P) \oplus (N \otimes P)$ such that $(m,n) \otimes p \mapsto (m \otimes p, n \otimes p)$. An inverse to this homomorphism is constructed by similar arguments. □

Exercise* 1.10.17. Complete the proof of Proposition 1.10.16. □

Exercise 1.10.18. Show:

1. $\mathbb{Z}/3\mathbb{Z} \otimes \mathbb{Z}/2\mathbb{Z} = 0$.
2. $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q} \cong \mathbb{Q}$. □

Remark 1.10.19 (Tensor Product of Homomorphisms). If $\phi : M \to N$ and $\phi' : M' \to N'$ are homomorphisms of R-modules, the map $M \times M' \to N \otimes N'$, $(m, m') \mapsto \phi(m) \otimes \phi'(m')$, is R-bilinear. It induces, thus, an R-module homomorphism

$$\phi \otimes \phi' : M \otimes M' \to N \otimes N'$$

such that

$$m \otimes m' \mapsto \phi(m) \otimes \phi'(m').$$ □

Our final remarks in this section deal with algebras. Given a ring homomorphism $\phi : R \to S$, we make S into an R-module by setting $rs := \phi(r)s$ for all $r \in R$ and $s \in S$. This R-module structure is compatible with the ring structure on S in the sense that

$$(rs)s' = r(ss').$$

We refer to this fact by saying that S is an ***R*-algebra**. A **subalgebra** of S is a subring S' of S contained in the image of ϕ.

Remark 1.10.20. With notation as above, let $R = \Bbbk$ be a field (and suppose that S nonzero). Then ϕ is necessarily a monomorphism. Identifying $\Bbbk$ with its image in S by means of ϕ, we see that a $\Bbbk$-algebra is nothing but a ring S containing $\Bbbk$ as a subring. A particular example is the polynomial ring $\Bbbk[x_1, \dots, x_n]$. □

An ***R*-algebra homomorphism** between two R-algebras S and T is a ring homomorphism $S \to T$ which is also an R-module homomorphism. Mono-, epi-, and isomorphisms of R-algebras are defined in the usual way.

Exercise* 1.10.21 (Tensor Product of Algebras). Let S and T be R-algebras, defined by maps $\phi : R \to S$ and $\psi : R \to T$. Use the universal property of the tensor product and Proposition 1.10.16 to establish a multiplication on $S \otimes_R T$ such that

$$(s \otimes t)(s' \otimes t') = ss' \otimes tt'.$$

Show that this multiplication turns $S \otimes_R T$ into a (commutative) ring (with multiplicative identity $1 \otimes 1$). Furthermore, show that $S \otimes_R T$ is an R-algebra: the map

$$R \to S \otimes_R T,\ r \mapsto \phi(r) \otimes \psi(r),$$

is a ring homomorphism. □

We say that an ***R*-algebra *S* is finitely generated** if there are elements $s_1, \dots, s_n \in S$ such that every element of S is a polynomial expression in the s_i with coefficients in R. This means that S can be written as a quotient ring of type $R[x_1, \dots, x_n]/I$. Indeed, consider the (R-algebra) epimorphism

$$R[x_1, \dots, x_n] \to S,\ x_i \mapsto s_i,$$

and take I to be the kernel. In combining this with the general version of Hilbert's basis theorem and Exercise 1.5.11, we see that a finitely generated algebra over a Noetherian ring is again a Noetherian ring. In particular, every finitely generated $\Bbbk$-algebra is a Noetherian ring. We refer to such a $\Bbbk$-algebra as an **affine $\Bbbk$-algebra**, or simply as an **affine ring**. An **affine domain** is an affine ring without zerodivisors. This terminology is justified by Proposition 1.11.4 below.

1.11 Coordinate Rings and Morphisms

In this section, we will return to the geometry-algebra dictionary. We will take up a theme familiar from other courses in mathematics: The study of a given class of mathematical objects usually requires that we understand the natural maps, or morphisms, between these objects. In linear algebra, for instance, we deal with linear maps between vector spaces, and in topology, we consider continuous maps between topological spaces. In algebraic geometry, where the objects of study are given by polynomials, the morphisms are also given by polynomials. In discussing this, we will first treat the polynomial functions on an algebraic set. Then, we will use these functions to define the morphisms.

Definition 1.11.1. Let $A \subset \mathbb{A}^n(\Bbbk)$ be a (nonempty) algebraic set. A **polynomial function** on A is the restriction of a polynomial function on $\mathbb{A}^n(\Bbbk)$ to A. □

The set $\Bbbk[A]$ of all polynomial functions on A is made into a ring, with algebraic operations defined by adding and multiplying values in $\Bbbk$: If $f, g \in \Bbbk[A]$, for all $p \in A$ set

$$(f + g)(p) = f(p) + g(p) \ \text{ and } \ (f \cdot g)(p) = f(p) \cdot g(p).$$

We regard $\Bbbk$ as the subring of all constant functions and, thus, $\Bbbk[A]$ as a $\Bbbk$-algebra. Since two polynomials in $\Bbbk[x_1, \dots, x_n]$ define the same element of $\Bbbk[A]$ iff their difference vanishes on A, we have a natural isomorphism

$$\Bbbk[x_1, \ldots, x_n]/\mathrm{I}(A) \cong \Bbbk[A]$$

which allows us to identify the two rings. Accordingly, the elements of $\Bbbk[A]$ may be viewed in two ways – as residue classes of polynomials modulo $\mathrm{I}(A)$, or as polynomial functions on A.

Definition 1.11.2. The **coordinate ring** of a (nonempty) algebraic set $A \subset \mathbb{A}^n(\Bbbk)$ is the $\Bbbk$-algebra $\Bbbk[A]$ defined above. □

The notion reflects the fact that $\Bbbk[A]$ is the $\Bbbk$-algebra of functions on A generated by the coordinate functions on A.

Exercise 1.11.3. Let $\Bbbk$ be a finite field, and let $A \subset \mathbb{A}^n(\Bbbk)$ be an algebraic set. Show that $\Bbbk[A]$ is the ring of *all* $\Bbbk$-valued functions on A. □

According to our definitions, coordinate rings are specific examples of affine $\Bbbk$-algebras. In particular, they are Noetherian. Furthermore, they are reduced since vanishing ideals are radical ideals. Somewhat conversely, we have:

Proposition 1.11.4. *If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, every reduced affine $\Bbbk$-algebra T is of the form $T = \Bbbk[A]$ for some affine algebraic set A.*

Proof. Write T as the quotient of a polynomial ring $\Bbbk[x_1, \ldots, x_n]$ modulo an ideal I. Then I is a radical ideal since T is reduced. The Nullstellensatz implies that $I = \mathrm{I}(\mathrm{V}(I))$, and we may take $A = \mathrm{V}(I)$. □

In the following exercise, we write $\boldsymbol{x} = x_1, \ldots, x_n$ and $\boldsymbol{y} = y_1, \ldots, y_m$.

Exercise* 1.11.5. Let $A \subset \mathbb{A}^n(\Bbbk)$ and $B \subset \mathbb{A}^m(\Bbbk)$ be algebraic sets. Show:

1. The product $A \times B \subset \mathbb{A}^n(\Bbbk) \times \mathbb{A}^m(\Bbbk) = \mathbb{A}^{n+m}(\Bbbk)$ is an algebraic set.
2. If $\mathrm{I}(A) \subset \Bbbk[\boldsymbol{x}]$ and $\mathrm{I}(B) \subset \Bbbk[\boldsymbol{y}]$ are the vanishing ideals of A and B, then
$$\mathrm{I}(A \times B) = \mathrm{I}(A)\,\Bbbk[\boldsymbol{x}, \boldsymbol{y}] + \mathrm{I}(B)\,\Bbbk[\boldsymbol{x}, \boldsymbol{y}] \subset \Bbbk[\boldsymbol{x}, \boldsymbol{y}].$$
3. For the cooordinate rings, we have
$$\Bbbk[A \times B] \cong \Bbbk[A] \otimes_{\Bbbk} \Bbbk[B].$$
□

Our next exercise shows that the idea of relating algebraic sets to ideals still works nicely if we replace $\mathbb{A}^n(\Bbbk)$ and $\Bbbk[x_1, \ldots, x_n]$ by an arbitrary algebraic set $A \subset \mathbb{A}^n(\Bbbk)$ and its coordinate ring $\Bbbk[A]$. We use the following notation:

Definition 1.11.6. Let $A \subset \mathbb{A}^n(\Bbbk)$ be an algebraic set.

1. If $J \subset \Bbbk[A]$ is a subset, its **locus of zeros** in A is the set
$$\mathrm{V}_A(J) := \{p \in A \mid f(p) = 0 \text{ for all } f \in J\}.$$
2. If $B \subset A$ is a subset, its **vanishing ideal** in $\Bbbk[A]$ is the ideal
$$\mathrm{I}_A(B) := \{f \in \Bbbk[A] \mid f(p) = 0 \text{ for all } p \in B\}.$$

3. An **algebraic subset** of A is an algebraic subset of $\mathbb{A}^n(\Bbbk)$ contained in A. A **subvariety** of A is a subvariety of $\mathbb{A}^n(\Bbbk)$ contained in A. □

Exercise* 1.11.7. Let $A \subset \mathbb{A}^n(\Bbbk)$ be an algebraic set. Show:

1. A subset $B \subset A$ is algebraic iff $B = \mathrm{V}_A(J)$ for some ideal $J \subset \Bbbk[A]$.
2. If $B \subset A$ is a subset, then $\mathrm{I}_A(B)$ is indeed an ideal of $\Bbbk[A]$.
3. The algebraic subsets of A are the closed sets of the Zariski topology on A. The **distinguished open sets**

$$\mathrm{D}_A(f) := A \setminus \mathrm{V}_A(f),\ f \in \Bbbk[A],$$

   form a basis for this topology.
4. If $B \subset A$ is an algebraic subset, then

$$\mathrm{V}_A(\mathrm{I}_A(B)) = B.$$

5. **(Nullstellensatz in $\Bbbk[A]$)** If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, and $J \subset \Bbbk[A]$ is an ideal, then

$$\mathrm{I}_A(\mathrm{V}_A(J)) = \operatorname{rad} J.$$

   Hint. Deduce this from Hilbert's Nullstellensatz by passing from ideals in $\Bbbk[A] = \Bbbk[x_1, \dots, x_n]/\mathrm{I}(A)$ to ideals in $\Bbbk[x_1, \dots, x_n]$ (see Exercise 1.5.11).
6. If $\Bbbk = \overline{\Bbbk}$, then I_A and V_A define a one-to-one inclusion-reversing correspondence

$$\begin{tikzcd}\{\text{algebraic subsets of } A\} \arrow[r, shift left, "\mathrm{I}_A"] & \{\text{radical ideals of } \Bbbk[A]\}. \arrow[l, shift left, "\mathrm{V}_A"]\end{tikzcd}$$

   Under this correspondence, subvarieties correspond to prime ideals, and points to maximal ideals. □

Recall that the usual Euclidean topology over the real or complex numbers is Hausdorff. In contrast, the Zarisky topology on an affine variety V is not Hausdorff, except when V consists of a single point:

Proposition 1.11.8. *Let $A \subset \mathbb{A}^n(\Bbbk)$ be an algebraic set. Then the following conditions are equivalent:*

1. *A is irreducible.*
2. *Any two nonempty open subsets of A have a nonempty intersection.*
3. *Any nonempty open subset of A is dense in A.*

Proof. Since the intersection of two subsets of A is empty iff the union of their complements equals A, condition 2 is just a restatement of the defining condition of irreducibility. Condition 3, in turn, is a restatement of condition 2 since a subset of a topological space is dense iff it meets every nonempty open subset. □

We summarize some properties of the Zariski topology for later use:

Exercise* 1.11.9. Let $A \subset \mathbb{A}^n(\Bbbk)$ be an algebraic set. Show:

1. The Zariski topology on A is **quasicompact**. That is, every open cover of A has a finite subcover.
2. An open subset of A is dense in A iff it meets every irreducible component of A.
3. Every open dense subset of A contains a distinguished open dense subset of A.
4. A distinguished open set $\mathrm{D}_A(f)$ is dense in A iff f is a nonzerodivisor of $\Bbbk[A]$. □

Exercise* 1.11.10. Let $A \subset \mathbb{A}^n(\Bbbk)$ and $B \subset \mathbb{A}^m(\Bbbk)$ be algebraic sets.

1. Show that the product $A \times B$ is irreducible iff A and B are irreducible.
2. Give an example showing that the Zariski topology on $A \times B$ may not be the product of the Zariski topologies on A and B. □

We, now, come to the morphisms of affine algebraic sets. In our discussion, we will denote the coordinates on $\mathbb{A}^n$ and $\mathbb{A}^m$ by $x_1, \dots, x_n$ and $y_1, \dots, y_n$, respectively.

Definition 1.11.11. Let $A \subset \mathbb{A}^n(\Bbbk)$ and $B \subset \mathbb{A}^m(\Bbbk)$ be (nonempty) algebraic sets. A map $\varphi : A \to B$ is a **polynomial map**, or a **morphism**, if its components are polynomial functions on A. That is, there exist polynomials $f_1, \dots, f_m \in \Bbbk[x_1, \dots, x_n]$ such that $\varphi(p) = (f_1(p), \dots, f_m(p))$ for all $p \in A$. □

Proposition 1.11.12. *A map $\varphi : A \to B$ of affine algebraic sets is a polynomial map iff for all $f \in \Bbbk[B]$, the composition $f \circ \varphi$ is in $\Bbbk[A]$.*

Proof. If the condition on the right hand side is fulfilled, then, in particular, the $y_j \circ \varphi$ are polynomial functions on A. That is, φ is a polynomial map.

Conversely, let φ be given by polynomials $f_1, \dots, f_m \in \Bbbk[x_1 \dots, x_n]$. If $q \mapsto g(q)$ is a polynomial function on B, represented by a polynomial $g(y_1, \dots, y_m) \in \Bbbk[y_1 \dots, y_m]$, then $p \mapsto g(\varphi(p))$ is a polynomial function on A, represented by the polynomial $g(f_1, \dots, f_m) \in \Bbbk[x_1 \dots, x_n]$. □

Theorem 1.11.13. *Let $A \subset \mathbb{A}^n(\Bbbk)$ and $B \subset \mathbb{A}^m(\Bbbk)$ be algebraic sets.*

1. *Every polynomial map induces a $\Bbbk$-algebra homomorphism*

$$\varphi^* : \Bbbk[B] \to \Bbbk[A],\ g \mapsto g \circ \varphi.$$

2. *Conversely, if $\phi : \Bbbk[B] \to \Bbbk[A]$ is a $\Bbbk$-algebra homomorphism, there exists a unique polynomial map $\varphi : A \to B$ such that $\phi = \varphi^*$.*

3. If $\varphi : A \to B$ and $\psi : B \to C$ are polynomial maps, their composition $(\psi \circ \varphi) : A \to C$ is a polynomial map as well, and

$$(\psi \circ \varphi)^* = \varphi^* \circ \psi^*.$$

Furthermore,

$$\mathrm{id}_A^* = \mathrm{id}_{\Bbbk[A]}. \qquad \square$$

Exercise* 1.11.14. Prove Theorem 1.11.13. To show part 2, for $j = 1, \dots, m$, choose polynomials $f_j \in \Bbbk[x_1, \dots, x_n]$ representing $\phi(y_j + \mathrm{I}(B))$, and consider the polynomial map $\varphi : A \to \mathbb{A}^m(\Bbbk)$ defined by the f_j. □

Remark 1.11.15. According to Definition 6.3.12, every morphism of affine algebraic sets is induced by a morphism of the ambient affine spaces. That is, with notation as in the definition, there is a commutative diagram

$$\begin{tikzcd}
\mathbb{A}^n(\Bbbk) \arrow[r, "F=(f_1,\dots,f_m)"] & \mathbb{A}^m(\Bbbk) \\
A \arrow[u] \arrow[r, "\varphi"] & B \arrow[u] .
\end{tikzcd}$$

By Theorem 1.11.13, there is a corresponding commutative diagram of $\Bbbk$-algebra homomorphisms, with all arrows reversed:

$$\begin{tikzcd}
\Bbbk[x_1, \dots, x_n] \arrow[d] & \Bbbk[y_1, \dots, y_m] \arrow[l, "F^*"'] \arrow[d] \\
\Bbbk[A] & \Bbbk[B] \arrow[l, "\varphi^*"'] .
\end{tikzcd}$$

Here, F^* is obtained by substituting the f_j for the y_j. Any $\Bbbk$-algebra homomorphism $\phi : \Bbbk[y_1, \dots, y_m] \to \Bbbk[x_1, \dots, x_n]$ is a substitution homomorphism, and the images $f_j := \phi(y_j)$ define a morphism $\mathbb{A}^n(\Bbbk) \to \mathbb{A}^m(\Bbbk)$ whose restriction to A is a morphism $A \to \mathbb{A}^m(\Bbbk)$. Note, however, that this morphism maps A to B only if $\phi(\mathrm{I}(B)) \subset \mathrm{I}(A)$. Thus, there are always plenty of morphisms $A \to \mathbb{A}^m(\Bbbk)$, but quite often only constant morphisms $A \to B$. See Exercise 1.11.20 for an example. □

Remark 1.11.16. Every morphism $A \to B$ of affine algebraic sets is continuous with regard to the respective Zariski topologies. Indeed, if $\mathrm{D}_B(g) \subset B$ is a distinguished open set, then $\varphi^{-1}(\mathrm{D}_B(g)) = \mathrm{D}_A(\varphi^*(g)) \subset A$ is a distinguished open set as well. □

As ususal, an isomorphism is a morphism admitting an inverse morphism:

Definition 1.11.17. A morphism $\varphi : A \to B$ of affine algebraic sets is called an **isomorphism** if there is a morphism $\psi : B \to A$ such that $\psi \circ \varphi = \mathrm{id}_A$ and $\varphi \circ \psi = \mathrm{id}_B$. We say that A and B are **isomorphic**, written $A \cong B$, if there is an isomorphism $A \to B$. □

Theorem 1.11.13 implies:

Corollary 1.11.18. *A morphism* $\varphi : A \to B$ *of affine algebraic sets is an isomorphism iff* $\varphi^* : \Bbbk[B] \to \Bbbk[A]$ *is an isomorphism of* $\Bbbk$*-algebras. Two affine algebraic sets are isomorphic iff their coordinate rings are isomorphic.* □

Exercise 1.11.19. Let $\Bbbk$ be infinite. Show:

1. The parametrization

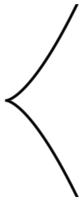
$$\mathbb{A}^1(\Bbbk) \to V(y - x^2, z - x^3) \subset \mathbb{A}^3(\Bbbk),\ a \mapsto (a, a^2, a^3),$$

of the twisted cubic curve is an isomorphism.

2. The map

$$\mathbb{A}^1(\Bbbk) \to \mathrm{V}(y^2 - x^3) \subset \mathbb{A}^2(\Bbbk),\ a \mapsto (a^2, a^3),$$

is a bijective morphism, but not an isomorphism.



□

How to decide algorithmically whether a given morphism of affine algebraic sets is an isomorphism will be explained in Section 2.5.

Exercise 1.11.20. If $\Bbbk$ is infinite, show that every morphism from the parabola $A = \mathrm{V}(y - x^2) \subset \mathbb{A}^2(\Bbbk)$ to the hyperbola $B = \mathrm{V}(xy - 1) \subset \mathbb{A}^2(\Bbbk)$ is constant. In particular, A and B are not isomorphic. □

The image of an affine algebraic set under an arbitrary morphism needs not be Zariski closed (we postpone a discussion of this failure to Section 2.6). Under an isomorphism $A \to B$, however, algebraic subsets of A correspond to algebraic subsets of B:

Exercise* **1.11.21.** Let $\varphi : A \to B$ be an isomorphism of affine algebraic sets, and let $A_1 \subset A$ be an algebraic subset. Show that $B_1 := \varphi(A_1)$ is an algebraic subset of B, and that φ restricts to an isomorphism of A_1 with B_1. *Hint.* If $A_1 = \mathrm{V}_A(f_1, \ldots, f_r)$, where $f_1, \ldots, f_r \in \Bbbk[A]$, show that $B_1 = \mathrm{V}_B(\psi^*(f_1), \ldots, \psi^*(f_r))$, where $\psi = \varphi^{-1}$. □

We usually think of isomorphic affine algebraic sets as the same geometric object, embedded in possibly different ways in affine spaces of possibly different dimensions. On the algebraic side, the vanishing ideal depends on the embedding, but the coordinate ring does not – the coordinate ring is invariant under isomorphism.

Definition 1.11.22. An isomorphism of an affine algebraic set A with itself is called an **automorphism** of A. □

The automorphisms of A form a group under composition which acts on A in the natural way. We write $\mathrm{Aut}(A)$ for this group.

Lemma 1.11.23. *A morphism* $F = (f_1, \dots, f_n) : \mathbb{A}^n(\Bbbk) \to \mathbb{A}^n(\Bbbk)$ *is an automorphism of* $\mathbb{A}^n(\Bbbk)$ *iff*

$$\Bbbk[x_1, \dots, x_n] = \Bbbk[f_1, \dots, f_n].$$

Proof. The condition $\Bbbk[x_1, \dots, x_n] = \Bbbk[f_1, \dots, f_n]$ means that the the $\Bbbk$-algebra homomorphism F^* induced by F is surjective. But, then, F^* is injective as well since, otherwise, the transcendence degree of the quotient field of $\Bbbk[x_1, \dots, x_n]/\ker F^* \cong \Bbbk[x_1, \dots, x_n]$ over $\Bbbk$ would be smaller than n. □

If $F = (f_1, \dots, f_n)$ is an automorphism of $\mathbb{A}^n(\Bbbk)$, we will speak of $f_1, \dots, f_n$ as a **coordinate system** of $\mathbb{A}^n(\Bbbk)$, and regard F as transforming $x_1, \dots, x_n$ into the new cooordinates $f_1, \dots, f_n$. The image of any algebraic subset $A \subset \mathbb{A}^n(\Bbbk)$ under F can, then, be thought of as the the original set A viewed using the new coordinates.

Definition 1.11.24. An automorphism of $\mathbb{A}^n(\Bbbk)$ is called a **change of coordinates** of $\mathbb{A}^n(\Bbbk)$. □

Example 1.11.25. We consider two types of automorphisms of $\mathbb{A}^n(\Bbbk)$ which are both preserved under taking the inverse (in fact, the automorphisms of either type form a subgroup of $\mathrm{Aut}(\mathbb{A}^n(\Bbbk))$):

1. An **affine change of coordinates** of $\mathbb{A}^n(\Bbbk)$ is given by degree-1 polynomials

   $$f_i = a_{i1}x_1 + \dots + a_{in}x_n + b_i \in \Bbbk[x_1, \dots, x_n],\ i = 1, \dots, n,$$

   where (a_{ij}) is an invertible $n \times n$ matrix with entries in $\Bbbk$, and where $b = (b_1, \dots, b_n) \in \Bbbk^n$. We speak of a **linear change of coordinates** if b is zero, and of a **translation** if (a_{ij}) is the identity matrix.
2. A **triangular change of coordinates** of $\mathbb{A}^n(\Bbbk)$ is given by polynomials of type

   $$f_i = x_i + g_i(x_1, \dots, x_{i-1}),\ i = 1, \dots, n,$$

   where $g_i \in \Bbbk[x_1, \dots, x_{i-1}]$ for all i (in particular, $g_1 = 0$). □

Remark 1.11.26. 1. By results of Jung (1942) and van der Kulk (1953), who treat the cases $\operatorname{char} \Bbbk = 0$ and $\operatorname{char} \Bbbk > 0$, respectively, $\mathrm{Aut}(\mathbb{A}^2(\Bbbk))$ is generated by affine and triangular changes of coordinates. It is not known, whether the analogous result holds in dimensions $n \geq 3$.

2. Given a morphism $F : \mathbb{A}^n(\Bbbk) \to \mathbb{A}^n(\Bbbk)$, we can easily write down a necessary condition for F to be an isomorphism. In fact, suppose that F admits an inverse G. Then $G \circ F = \mathrm{id}_{\mathbb{A}^n(\Bbbk)}$, and we may apply the chain rule and take determinants to conclude that the determinant of the **Jacobian matrix** $\left(\frac{\partial f_i}{\partial x_j}\right)$ is a nonzero constant. In case $\operatorname{char} \Bbbk = 0$, the famous **Jacobian conjecture** suggests that the condition on the determinant is also sufficient. Recently, quite a number of false proofs for this conjecture have been published – at least as e-prints (see `http://xxx.lanl.gov/archive/math`).

See van den Essen (2000) for further reading. □

Exercise 1.11.27. 1. Show that the Jacobian conjecture is true if $n = 1$.
2. Show by example that the condition on the Jacobian determinant may not be sufficient if $\operatorname{char} \Bbbk > 0$. □

Remark 1.11.28. 1. In the language of categories (see, for instance, Mac Lane (1990) for categories), Theorem 1.11.13 can be rephrased as follows. Over an algebraically closed field $\Bbbk$, the functor $A \to \Bbbk[A]$ induces an arrow-reversing equivalence between the category of affine algebraic sets over $\Bbbk$ and the category of reduced affine $\Bbbk$-algebras. The subcategory of affine varieties over $\Bbbk$ corresponds to that of affine domains over $\Bbbk$.
2. Grothendieck's concept of affine schemes gives a geometric interpretation of the full category of rings (commutative, and with a multiplicative identity). See Hartshorne (1977) and Eisenbud and Harris (2000). The concept of schemes, which will not be treated in this book, is fundamental to modern algebraic geometry. □

1.12 Additional Exercises

Exercise 1.12.1. Let $A \subset \mathbb{A}^n(\Bbbk)$ be a finite set. Show that A is an algebraic set which can be defined by n polynomial equations.
Hint. Use interpolation.

Exercise 1.12.2. If $\Bbbk$ is not algebraically closed, show that every algebraic subset of $\mathbb{A}^n(\Bbbk)$ can be defined by a single polynomial equation (see Exercise 1.6.6 for the case $\Bbbk = \mathbb{R}$).
Hint. Consider the case of the origin in $\mathbb{A}^2(\Bbbk)$ first.

Exercise 1.12.3. Describe all ideals of the quotient ring R/I for $R = \mathbb{R}[x]$ and $I = \langle x^3 - 2x^2 - x + 2\rangle$.

Exercise 1.12.4. If $\operatorname{char} \Bbbk = p > 0$, show that the map

$$\mathbb{A}^1(\Bbbk) \to \mathbb{A}^1(\Bbbk),\ a \mapsto a^p,$$

is a bijective morphism, but not an isomorphism. This map is called the **Frobenius morphism**. □

Chapter 2

Gröbner Bases

Our goal in this chapter is to tackle the computational problems arising from the geometry-algebra dictionary. For a guiding example, recall from Section 1.6 that both the problem of solvability and the problem of radical membership ask for a method to determine whether 1 belongs to a given ideal. Here, we encounter a special instance of a problem which is known as the **ideal membership problem**: Given $g, f_1, \dots, f_r \in \Bbbk[x_1, \dots, x_n]$, decide whether

$$g \in \langle f_1, \dots, f_r \rangle.$$

That is, decide whether there are $g_1, \dots, g_r \in \Bbbk[x_1, \dots, x_n]$ such that

$$g = \sum_{i=1}^{r} g_i f_i. \tag{2.1}$$

We may think of (2.1) as a system of (infinitely many) linear equations in the unknown coefficients of the g_i. To reduce to a finite number of equations (so that the system could be attacked by means of linear algebra), an *a priori* bound on the degree of the g_i is needed, Such a bound was established in the thesis of Grete Hermann (1926), a student of Emmy Noether. Hermann proved that each $g \in \langle f_1, \dots, f_r \rangle$ can be written as a sum $g = \sum_{i=1}^{r} g_i f_i$ such that

$$\deg g_i \leq \deg g + (rd)^{2^n} \text{ for all } i.$$

Here, d is the maximum degree of the f_i. Being doubly exponential in the number of variables, Herrmann's bound is quite large. Unfortunately, as shown by examples due to Mayr and Meyer (1982), the double exponential form of the bound cannot be improved.

It is worth pointing out that the special instance of checking whether 1 is contained in a given ideal and, thus, the radical membership problem admit a bound which is single exponential in the number of variables: If $h \in \operatorname{rad} \langle f_1, \dots, f_r \rangle \subset \Bbbk[x_1, \dots, x_n]$, there is an expression $h^m = \sum_{i=1}^{r} g_i f_i$ such that $m \leq d^n$ and $\deg(g_i f_i) \leq (1 + \deg h)d^n$, where $d = \max\{3, \deg f_i\}$ (see Kollár (1999) for a more precise statement giving an optimal bound).

In developing computational tools, we will not make use of the bounds discussed above. Instead, taking our cue from the case of one variable in which Euclidean division with remainder provides a solution to the ideal membership problem, we will extend the division algorithm to polynomials in more than one variable, allowing at the same time more than one divisor. Due to some undesirable behavior of the extended algorithm, however, this does not provide an immediate solution to the ideal membership problem. To remedy the situation, we introduce Gröbner bases, which are sets of generators for ideals behaving well under division with remainder. The name Gröbner basis was coined in the 1960's by Buchberger to honour his thesis advisor Gröbner. In his thesis, Buchberger used Gröbner bases to give an algorithmic way of computing in affine rings (1965, 1970). In particular, he designed an algorithm which computes Gröbner bases. In subsequent years, this algorithm became the major work horse of computational algebraic geometry. Though there is, again, a worst-case upper bound (on the degree of the elements of a Gröbner basis, see Möller and Mora (1984)) which is doubly exponential in the number of variables, Buchberger's algorithm works well in many examples of interest.

The algorithm is based on a criterion which allows one to check whether a given set of polynomials is a Gröbner basis. The resulting test yields certain $\mathbb{k}[x_1, \dots, x_n]$-linear relations on the elements of a Gröbner basis. These relations play a key role in our proof of Buchberger's criterion.

Given any $\mathbb{k}[x_1, \dots, x_n]$-linear relation

$$g_1 f_1 + \cdots + g_r f_r = 0$$

on polynomials $f_1, \dots, f_r \in \mathbb{k}[x_1, \dots, x_n]$, we think of it as a column vector

$$(g_1, \dots, g_r)^t \in \mathbb{k}[x_1, \dots, x_n]^r,$$

and call it a syzygy on $f_1, \dots, f_r$. It will turn out that the concept of Gröbner bases extends from ideals to submodules of free modules, and that Buchberger's algorithm computes syzygies as well. In fact, if $f_1, \dots, f_r$ form a Gröbner basis, the special syzygies obtained in Buchberger's test form a Gröbner basis for the module of all the syzygies on $f_1, \dots, f_r$. In theoretical terms, this will allow us to give a short proof of Hilbert's syzygy theorem which, following Hilbert, will be used in Section 6.4 to verify the polynomial nature of the Hilbert function. In practical terms, syzygy computations can be used to compute, for instance, ideal intersections and ideal quotients.

Among the fundamental applications of Gröbner bases is the elimination of variables from a given system of polynomial equations. Buchberger's algorithm extends, thus, Gaussian elimination. Geometrically, elimination amounts to projection. More generally, it will allow us to compute the Zariski closure of the image of an algebraic set under an arbitrary morphism

Historically, as already pointed out in Chapter 1, Gröbner bases made their first appearance in Gordan's proof (1899) of Hilbert's basis theorem.

This proof nicely demonstrates the key idea in the use of Gröbner bases which is to reduce questions on arbitrary ideals to questions on monomial ideals and, thus, to questions which are usually much easier.

2.1 Monomials and Monomial Ideals

A **monomial** in $x_1, \dots, x_n$ is a product $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, where $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$. A **monomial ideal** of $\Bbbk[x_1, \dots, x_n]$ is an ideal generated by monomials.

Operations on polynomials become often simpler if we restrict them to monomials. For example, if $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$ is another multiindex, the **least common multiple** of x^α and x^β is

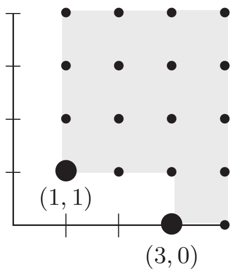
$$\mathrm{LCM}(x^\alpha, x^\beta) = x_1^{\max(\alpha_1,\beta_1)} \cdots x_n^{\max(\alpha_n,\beta_n)},$$

and their **greatest common divisor** is

$$\mathrm{GCD}(x^\alpha, x^\beta) = x_1^{\min(\alpha_1,\beta_1)} \cdots x_n^{\min(\alpha_n,\beta_n)}.$$

Similarly, monomial ideals are easier to handle than arbitrary ideals. As an example, consider the ideal membership problem: If $I \subset \Bbbk[x_1, \dots, x_n]$ is a monomial ideal, given by monomial generators $m_1, \dots, m_r$, a term is contained in I iff it is divisible by at least one of the m_i; an arbitrary polynomial $g \in \Bbbk[x_1, \dots, x_n]$ is contained in I iff all its terms are contained in I.

Example 2.1.1. In the following picture, we visualize the monomials in $\Bbbk[x, y]$ via their exponent vectors. The monomials contained in the ideal $I = \langle x^3, xy \rangle$ correspond to the dots in the shaded region:



The monomials $1, x, x^2$ and all the powers of y are not contained in I. □

The first step in Gordan's proof of Hilbert's Basis Theorem 1.4.1 is to show that monomial ideals are finitely generated (see Corollary 2.3.3 for the remaining part of the proof):

Exercise* 2.1.2 (Gordan's Lemma). By induction on the number of variables, show that any nonempty set of monomials in $\Bbbk[x_1, \dots, x_n]$ has only finitely many minimal elements in the partial order given by divisibility ($x^\alpha \geq x^\beta$ iff $\alpha - \beta \in \mathbb{N}^n$). Conclude that any monomial ideal $I \subset \Bbbk[x_1, \dots, x_n]$ has finitely many monomial generators. □

The minimal elements in the sitation above are obtained from any set of monomial generators by removing those generators which are divisible by others. In this way, we obtain a uniquely determined set of monomial generators for I, to which we refer as the **minimal generators** for I.

Exercise* **2.1.3.** Let I and J be monomial ideals of $\Bbbk[x_1, \dots, x_n]$, given by monomial generators $m_1, \dots, m_r$ and $n_1 \dots, n_s$, respectively, and let m be a monomial in $\Bbbk[x_1, \dots, x_n]$.

1. Show that
$$I \cap J = \langle \mathrm{LCM}(m_i, n_j) \mid 1 \leq i \leq r,\ 1 \leq j \leq s \rangle.$$
2. Show that $I : m$ is generated by the monomials
$$\mathrm{LCM}(m_i, m)/m = m_i / \mathrm{GCD}(m_i, m),\ 1 \leq i \leq r.$$

In particular, $I \cap J$ and $I : m$ are monomial ideals as well. The same is, hence, true for $I : J$ since $I : J = \bigcap_{k=1}^{s} (I : n_k)$ by part 3 of Exercise 1.3.3. □

Most of the terminology used when working with polynomials extends to elements of free modules over polynomial rings. In what follows, let $R = \Bbbk[x_1, \dots, x_n]$, and let F be a free R-module with a fixed basis $\{e_1, \dots, e_s\}$.

Definition 2.1.4. A **monomial** in F, **involving the basis element** e_i, is a monomial in R times e_i. A **term** in F is a monomial in F multiplied by a **coefficient** $c \in \Bbbk$. Every nonzero element $f \in F$ can be uniquely expressed as the sum of finitely many nonzero terms involving distinct monomials. These terms (monomials) are called the **terms** (**monomials**) **of** $\boldsymbol{f}$. □

To give an example, if $F = \Bbbk[x, y]^3$, and $e_1 = (1, 0, 0)^t$, $e_2 = (0, 1, 0)^t$, $e_3 = (0, 0, 1)^t$ are the canonical basis vectors, then
$$f := \begin{pmatrix} x^2y + x^2 \\ 1 \\ x \end{pmatrix} = x^2y \cdot e_1 + x^2 \cdot e_1 + 1 \cdot e_2 + x \cdot e_3 \in F.$$

For terms in F, notions like multiple or divisible are defined in the obvious way. For instance, the nonzero term $cx^\alpha e_i$ is **divisible** by the nonzero term $dx^\beta e_j$, with **quotient** $c/d\ x^{\alpha - \beta} \in R$, if $i = j$ and x^α is divisible by x^β. Furthermore, the **least common multiple** of two nonzero terms involving the same basis element e_i is defined by the formula
$$\mathrm{LCM}(cx^\alpha e_i, dx^\beta e_i) = \mathrm{LCM}(x^\alpha, x^\beta)\, e_i \in F.$$

If $cx^\alpha e_i$ and $dx^\beta e_j$ involve distinct basis elements, we set

$$\mathrm{LCM}(cx^\alpha e_i, dx^\beta e_j) = 0.$$

A submodule of F is a **monomial submodule** if it is generated by monomials. It easily follows from Gordan's lemma that every such submodule is generated by finitely many monomials (see Exercise 1.10.9). As in the ideal case, there ie a unique finite set of **minimal generators**. Moreover, membership in monomial submodules can be decided as for monomial ideals.

Exercise* 2.1.5. If I, J are monomial submodules of F, given by monomial generators, and if $m \in F$ is a term, show how to obtain monomial generators for the submodule $I \cap J \subset F$ and the ideal $I : m \subset R$. □

2.2 Division with Remainder

Euclid's division algorithm for polynomials in one variable, which we recall now, relies on the fact that the monomials in $\Bbbk[x]$ and, thus, the terms of every polynomial $f \in \Bbbk[x] \setminus \{0\}$ can be arranged unambiguously by degree. In fact, for the division process, we write the terms of f in *decreasing* order by degree, referring to the term of highest degree as the leading term. In the discussion below, we denote this term by $\mathbf{L}(f)$.

Theorem 2.2.1 (Euclidean Division with Remainder). *Let f be a nonzero polynomial in $\Bbbk[x]$. For every polynomial $g \in \Bbbk[x]$, there are uniquely determined polynomials $g_1, h \in \Bbbk[x]$ such that*

$$g = g_1 f + h \quad \textit{and} \quad \deg h < \deg f.$$

□

Indeed, **Euclid's division algorithm** finds the remainder h and the quotient g_1 by successively using f to cancel leading terms. We write this in pseudocode:

```
1. Set $h := g$ and $g_1 := 0$.
2. while $\big(h \neq 0$ and $\mathbf{L}(h)$ is divisible by $\mathbf{L}(f)\big)$
        • set $h := h - \frac{\mathbf{L}(h)}{\mathbf{L}(f)} f$ and $g_1 := g_1 + \frac{\mathbf{L}(h)}{\mathbf{L}(f)} f$.
3. return$(h, g_1)$.
```

This process must terminate since, at each stage, the degree of the new dividend is smaller than that of the preceeding dividend.

Remark 2.2.2. Euclidean division with remainder also works for univariate polynomials with coefficients in a ring, provided the divisor f is **monic**. That is, the coefficient of the leading term of f is 1. □

Exercise 2.2.3 (Euclid's GCD algorithm). If an expression $g = g_1 f + h$ as in Theorem 2.2.1 is given, show that $\mathrm{GCD}(f, g) = \mathrm{GCD}(f, h)$ (here, GCD refers to the monic greatest common divisor). Deduce Euclid's classical algorithm for computing $\mathrm{GCD}(f, g)$. Show how to extend this algorithm such that it computes not only the GCD but also a representation

$$\mathrm{GCD}(f, g) = sf + tg, \quad \text{where } \ s, t \in \Bbbk[x].$$

□

Euclidean division with remainder allows us to decide ideal membership in $\Bbbk[x]$ as follows. If nonzero polynomials $g, f_1, \ldots, f_r \in \Bbbk[x]$ are given, use Euclid's algorithm to compute $f = \mathrm{GCD}(f_1, \ldots, f_r)$. Then $\langle f_1, \ldots, f_r\rangle = \langle f\rangle$, so that $g \in \langle f_1, \ldots, f_r\rangle$ iff the remainder of g on division by f is zero.

To solve the ideal membership problem for polynomials in more than one variable in a similar way, we have to extend the division algorithm. Since for $n \geq 2$ not every ideal of $\Bbbk[x_1, \ldots, x_n]$ is generated by just one element, we ask for an algorithm which divides by several polynomials in $\Bbbk[x_1, \ldots, x_n]$ instead of a single polynomial. As in the case of one variable, we need to impose a total order on the set of monomials in $\Bbbk[x_1, \ldots, x_n]$ which allows us to single out leading terms of polynomials. This has to be done with some care:

Example 2.2.4. If $f_1 = x^2 + xy \in \Bbbk[x, y]$, any polynomial $g \in \Bbbk[x, y]$ can be written in the form $g = g_1 f_1 + h$, where no term of h is a multiple of x^2. Similarly, we may use $f_2 = y^2 + xy \in \Bbbk[x, y]$ to cancel the multiples of y^2. It is not possible, however, to cancel the multiples of x^2 and the multiples of y^2 simultaneously using f_1 and f_2: If every polynomial $g \in \Bbbk[x, y]$ could be written in the form

$$g = g_1 f_1 + g_2 f_2 + h,$$

where no term of h is contained in the ideal $\langle x^2, y^2\rangle$, the monomials $1, x, y, xy$ would represent generators for $\Bbbk[x, y]/\langle f_1, f_2\rangle$ as a $\Bbbk$-vector space. Thus, by Exercise 1.6.5, the locus of zeros of $\langle f_1, f_2\rangle$ in $\mathbb{A}^2(\overline{\Bbbk})$ would be finite. This is impossible since this locus contains the line with equation $x + y = 0$.

The problem with choosing the leading terms x^2 of f_1 and y^2 of f_2 is that this choice is not compatible with the multiplication in $\Bbbk[x, y]$ in the sense of the following definition. □

Definition 2.2.5. A **monomial order** on $\Bbbk[x_1, \ldots, x_n]$ is a total order $>$ on the set of monomials in $\Bbbk[x_1, \ldots, x_n]$ such that if $\alpha, \beta, \gamma \in \mathbb{N}^n$, then

$$x^\alpha > x^\beta \implies x^\gamma x^\alpha > x^\gamma x^\beta.$$

□

Example 2.2.6. The following are monomial orders on $\Bbbk[x_1, \ldots, x_n]$:

1. **(Lexicographic order)** Set

$$x^\alpha >_{\mathrm{lex}} x^\beta \iff \text{the first nonzero entry of } \alpha - \beta \text{ is positive.}$$

2. **(Weight orders)** If $w = (w_1, \ldots, w_n)\colon \mathbb{R}^n \to \mathbb{R}$ is a linear form with $\mathbb{Q}$-linearly independent coefficients w_i, set
$$x^\alpha >_w x^\beta \iff w(\alpha) > w(\beta).$$
In this context, given a term cx^α with $0 \neq c \in \Bbbk$, we will occasionally abuse notation by writing $w(cx^\alpha) = w(\alpha)$. □

Note that we have defined $>_{\text{lex}}$ such that the variables are ordered according to their appearance when writing $\Bbbk[x_1, \ldots, x_n]$. For instance, in $\Bbbk[x, y, z]$,

$$x^3 >_{\text{lex}} xyz >_{\text{lex}} x >_{\text{lex}} y^{25} >_{\text{lex}} y >_{\text{lex}} z.$$

Given a monomial order $>$ on $\Bbbk[x_1, \ldots, x_n]$, we will abuse notation as follows: If $c, d \in \Bbbk \setminus \{0\}$ are scalars and x^α, x^β are monomials in $\Bbbk[x_1, \ldots, x_n]$ such that $x^\alpha > x^\beta$ (or $x^\alpha \geq x^\beta$), we will write $cx^\alpha > dx^\beta$ (or $cx^\alpha \geq dx^\beta$). In the same spirit, we will occasionally speak of the maximum of a finite number of nonzero terms (which is determined up to a scalar).

Definition 2.2.7. Let $>$ be a monomial order on $\Bbbk[x_1, \ldots, x_n]$, and let $f \in \Bbbk[x_1, \ldots, x_n]$ be a nonzero polynomial. The **leading term** (or **initial term**) of f with respect to $>$, written

$$\mathbf{L}_>(f) = \mathbf{L}(f),$$

is the largest term of f with repect to $>$. By convention, $\mathbf{L}_>(0) = \mathbf{L}(0) = 0$. If $\mathbf{L}(f) = cx^\alpha$, with $c \in \Bbbk$, then c is called the **leading coefficient** of f and x^α is called the **leading monomial** of f. □

Remark 2.2.8. Since monomial orders are compatible with multiplication,

$$\mathbf{L}(fg) = \mathbf{L}(f)\mathbf{L}(g)$$

for all $f, g \in \Bbbk[x_1, \ldots, x_n]$. Furthermore, if f, g, and $f + g$ are nonzero, then

$$\max\{\mathbf{L}(f), \mathbf{L}(g)\} \geq \mathbf{L}(f + g).$$

The inequality is strict iff $\mathbf{L}(f)$ and $\mathbf{L}(g)$ cancel each other in $f + g$. □

This shows that if $\mathbf{L}(h)$ is divisible by $\mathbf{L}(f)$, and if we think of computing $h - \frac{\mathbf{L}(h)}{\mathbf{L}(f)} f$ as a single step of a division process, then the new dividend in such a step will be zero, or its leading term will be smaller than that of the preceeding dividend. This does not imply, however, that the process terminates:

Example 2.2.9. In $\Bbbk[x]$, choose the terms of lowest degree as the leading terms. Divide $g = x$ by $f = x - x^2$ using division steps as described above. Then, the successive intermediate dividends are $f = x - x^2, x^2, x^3, \ldots$. □

Proposition 2.2.10. *Let $>$ be a monomial order on $\Bbbk[x_1, \ldots, x_n]$. Then the following are equivalent:*

1. *$>$ is* **Artinian**. *That is, each nonempty set of monomials has a least element with respect to $>$.*
2. *$>$ is* **global**. *That is,*

$$x_i > 1 \quad \textit{for} \;\; i = 1, \dots, n.$$

3. *$>$* **refines the partial order defined by divisibility**. *That is,*

$$x^\alpha \;\textit{divisible by}\; x^\beta \implies x^\alpha > x^\beta.$$

Proof. The only nontrivial part of the proof is to show that condition 3 implies condition 1. If condition 3 holds, and X is a nonempty set of monomials, the monomial ideal $I = \langle X \rangle \subset \Bbbk[x_1, \dots, x_n]$ generated by X is, in fact, generated by a finite subset Y of X due to Gordan's lemma. Hence, every monomial in X is divisible by a monomial in Y, and the least element of Y is the least element of X. □

We use the word global to distinguish the monomial orders considered in this chapter from those used in Section 4.4, where we will explain how to compute in local rings. The lexicographic order is global. A weight order $>_w$ is global iff the coefficients of w are strictly positive.

Exercise* 2.2.11. Let $>$ be a monomial order on $\Bbbk[x_1, \dots, x_n]$, and let X be a finite set of monomials in $\Bbbk[x_1, \dots, x_n]$. Prove that there exists a weight order $>_w$ on $\Bbbk[x_1, \dots, x_n]$ which coincides on X with the given order $>$. If $>$ is global, show that $>_w$ can be chosen to be global as well.
Hint. Consider the set of differences

$$\{\alpha - \beta \mid x^\alpha, x^\beta \in X, x^\alpha > x^\beta\},$$

and show that its convex hull in $\mathbb{R}^n$ does not contain the origin. For the second statement, add $1, x_1, \dots, x_n$ to X if necessary. □

We are, now, ready to extend the division algorithm. In several variables, allowing several divisors, the result of the division process may depend on some choices made in carrying out the process. For instance, if h is some intermediate dividend, and $f_1, \dots, f_r$ are the divisors, it may happen that $\mathbf{L}(h)$ is divisible by more than one of the $\mathbf{L}(f)_i$, and any of these can be used to cancel $\mathbf{L}(h)$. Our first version of the extended division algorithm avoids such ambiguities. For us, this determinate version will be particularly useful in relating Buchberger's algorithm to syzygies (see Corollary 2.3.17).

Theorem 2.2.12 (Division with Remainder, Determinate Version). *Let $>$ be a global monomial order on $R = \Bbbk[x_1, \dots, x_n]$, and let $f_1, \dots, f_r \in R \setminus \{0\}$. For every $g \in R$, there exists a uniquely determined expression*

$$g = g_1 f_1 + \ldots + g_r f_r + h, \;\; \textit{with}\; g_1, \dots, g_r, h \in R,$$

and such that:

(DD1) For $i > j$, no term of $g_i\,\mathbf{L}(f_i)$ is divisible by $\mathbf{L}(f_j)$.
(DD2) For all i, no term of h is divisible by $\mathbf{L}(f_i)$.

We call h the ***remainder*** *of g on determinate division by $f_1, \dots, f_r$.*

Proof. Uniqueness. Given any representation as in the assertion, conditions (DD1) and (DD2) imply that the nonzero terms among the $\mathbf{L}(g_i f_i) = \mathbf{L}(g_i)\mathbf{L}(f_i)$ and $\mathbf{L}(h)$ involve different monomials. Hence, these terms do not cancel with each other on the right hand side of the representation. If two such representations for $g \in R$ are given, their difference is a representation for the zero polynomial satisfying (DD1) and (DD2). According to what we just said, the difference must be the trivial representation.

Existence. The **determinate division algorithm** finds the desired representation for $g \in R$ as follows.

If $f_1, \dots, f_r$ are terms, first remove any multiple of f_1 from g. Then cancel the remaining multiples of f_2. Continue in this way until any multiple of any f_k has been removed.

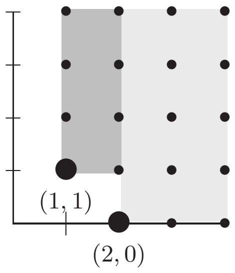
If $f_1, \dots, f_r$ are arbitrary, apply the above to g and $\mathbf{L}(f_1), \dots, \mathbf{L}(f_r)$. If

$$g = \sum_{i=1}^{r} g_i\,\mathbf{L}(f_i) + h$$

is the resulting representation, then either $g^{(1)} := g - \sum_{i=1}^{r} g_i f_i - h$ is zero, and we are done, or $\mathbf{L}(g) > \mathbf{L}(g^{(1)})$. By recursion, since $>$ is Artinian, we may assume in the latter case that $g^{(1)}$ has a representation $g^{(1)} = \sum_{i=1}^{r} g_i^{(1)} f_i + h^{(1)}$ satisfying (DD1) and (DD2). Then $g = \sum_{i=1}^{r}(g_i + g_i^{(1)}) f_i + (h + h^{(1)})$ is a representation for g satisfying (DD1) and (DD2). □

Conditions (DD1) and (DD2) are best understood by considering a partition of the monomials in $\Bbbk[x_1, \dots, x_n]$ as in the following example:

Example 2.2.13. Let $f_1 = x^2, f_2 = xy + x \in \Bbbk[x, y]$ with $>_{\text{lex}}$. Then $\mathbf{L}(f_1) = f_1 = x^2$ and $\mathbf{L}(f_2) = xy$. In the picture below, the monomials divisible by $\mathbf{L}(f_1)$ correspond to the dots in the region which is shaded in light grey:



Given $g \in \Bbbk[x, y]$ and a representation $g = g_1 f_1 + g_2 f_2 + h$, condition (DD1) means that that the monomials of $g_2\mathbf{L}(f_2)$ are represented in the region shaded in dark grey. Condition (DD2), in turn, requires that the monomials of h are represented in the nonshaded region.

Dividing, for instance, $g = x^3 + x^2y^3 + xy^2$ by f_1 and f_2, we get:

$$g = (x + y^3) \cdot \mathbf{L}(f_1) + y \cdot \mathbf{L}(f_2) + 0,$$

$$g^{(1)} = g - (x + y^3) \cdot f_1 - y \cdot f_2 = -xy = 0 \cdot \mathbf{L}(f_1) - 1 \cdot \mathbf{L}(f_2) + 0,$$

$$g^{(2)} = g^{(1)} + f_2 = x = 0 \cdot \mathbf{L}(f_1) + 0 \cdot \mathbf{L}(f_2) + x,$$

and

$$g^{(3)} = g^{(2)} - x = 0.$$

Thus, the desired representation is

$$g = (x + y^3) \cdot f_1 + (y - 1) \cdot f_2 + x.$$

□

It should be particularly clear from the picture in the example above that condition (DD1) makes the order in which $f_1, \dots, f_r$ are listed play a crucial role in the determinate division algorithm. We illustrate this by another example:

Example 2.2.14. Let $f_1 = x^2y - y^3, f_2 = x^3 \in \Bbbk[x, y]$ with $>_{\text{lex}}$. Then $\mathbf{L}(f_1) = x^2y$. For $g = x^3y$, the determinate division algorithm proceeds as follows:

$$x^3y = x \cdot \mathbf{L}(f_1) + 0 \cdot \mathbf{L}(f_2) + 0,$$

$$g^{(1)} = g - x \cdot f_1 = xy^3 = 0 \cdot \mathbf{L}(f_1) + 0 \cdot \mathbf{L}(f_2) + xy^3,$$

and

$$g^{(2)} = g^{(1)} - xy^3 = g - x \cdot f_1 - xy^3 = 0.$$

Thus, the desired representation is

$$x^3y = x \cdot (x^2y - y^3) + 0 \cdot (x^3) + xy^3.$$

If we interchange f_1 and f_2, determinate division yields the expression

$$x^3y = y \cdot (x^3) + 0 \cdot (x^2y - y^3) + 0.$$

□

Exercise 2.2.15. Define a global monomial order on $\Bbbk[x, y, z]$ yielding the leading terms y of $y - x^2$ and z of $z - x^3$, and reconsider part 1 of Exercise 1.5.4. □

Remark 2.2.16 (Division with Remainder, Indeterminate Version). With notation as in Theorem 2.2.12, the steps below describe a version of the division algorithm which is indeterminate: the computed remainder depends on the choices made in the `while` loop (termination follows once more from the fact that a global monomial order is Artinian).

1. Set $h := g$ and $D := \{f_1, \dots, f_r\}$.

```
2. while $\big(h \neq 0$ and $D(h) := \{f \in D \mid \mathbf{L}(h)$ is divisible by $\mathbf{L}(f)\} \neq \emptyset\big)$
       • choose $f \in D(h)$;
       • set $h := h - \frac{\mathbf{L}(h)}{\mathbf{L}(f)} f$.
3. return$(h)$.
```

With some extra bookkeeping as in Euclid's division algorithm, the algorithm also returns polynomials $g_1, \dots, g_r$ such that $g = g_1 f_1 + \dots + g_r f_r + h$. This representation of g satisfies the conditions (ID1) and (ID2) below which are weaker than the conditions (DD1) and (DD2), respectively:

(ID1) $\mathbf{L}(g) \geq \mathbf{L}(g_i f_i)$ whenever both sides are nonzero.
(ID2) If h is nonzero, then $\mathbf{L}(h)$ is not divisible by any $\mathbf{L}(f_i)$.

Each such representation is called a **standard expression** for g with **remainder** h (in terms of the f_i, with respect to $>$).

In practical terms, it is often useful to give up uniqueness and allow choices to be made since some choices are more efficient than others. In fact, there are various possible selection strategies for the division process. It is not clear to us whether there is a "generally best" strategy. Typically, the selection of the strategies depends on the particular application one has in mind.

A version of the division algorithm which is even more indeterminate is discussed in the exercise below. □

Exercise 2.2.17. Show that we still get a division process which terminates if, at each stage, we remove *some* term of the current dividend with the help of some $\mathbf{L}(f_i)$ by which it is divisible, and if we stop as soon as this is no longer possible. Show that the resulting representation $g = g_1 f_1 + \dots + g_r f_r + h$ satisfies the conditions (ID1) and (DD2). □

Remark 2.2.18 (Leading Terms in Standard Expressions). If g is a nonzero polynomial in $\Bbbk[x_1, \dots, x_n]$, and $g = g_1 f_1 + \dots + g_r f_r + h$ is a standard expression, then $\mathbf{L}(g)$ is the maximum nonzero term among the $\mathbf{L}(g_i f_i) = \mathbf{L}(g_i)\mathbf{L}(f_i)$ and $\mathbf{L}(h)$ (the term is determined up to a scalar). Indeed, this follows from condition (ID1) in conjunction with Remark 2.2.8. In particular, if the remainder h is zero, then $\mathbf{L}(g)$ is divisible by one of $\mathbf{L}(f_1), \dots, \mathbf{L}(f_r)$. We, then, write

$$\mathbf{L}(g) = \max\{\mathbf{L}(g_1)\mathbf{L}(f_1), \dots, \mathbf{L}(g_r)\mathbf{L}(f_r)\} \in \langle \mathbf{L}(f_1), \dots, \mathbf{L}(f_r) \rangle. \qquad \square$$

Our goal in this chapter is to develop the computational concepts not only for polynomial rings, but also for free modules over polynomial rings. In extending division with remainder to free modules, we write $R = \Bbbk[x_1, \dots, x_n]$, and consider a free R-module F with a fixed basis $\{e_1, \dots, e_s\}$.

Definition 2.2.19. A **monomial order** on F is a total order $>$ on the set of monomials in F such that if $x^\alpha e_i$ and $x^\beta e_j$ are monomials in F, and x^γ is a monomial in R, then

$$x^\alpha e_i > x^\beta e_j \implies x^\gamma x^\alpha e_i > x^\gamma x^\beta e_j. \qquad \square$$

In this book, we require in addition that

$$x^\alpha e_i > x^\beta e_i \iff x^\alpha e_j > x^\beta e_j \text{ for all } i, j.$$

Then $>$ induces a unique monomial order on R in the obvious way, and we say that $>$ is **global** if the induced order on R is global.

Remark 2.2.20. One way of getting a monomial order on F is to pick a monomial order $>$ on R, and extend it to F. For instance, setting

$$x^\alpha e_i > x^\beta e_j \iff x^\alpha > x^\beta \text{ or } (x^\alpha = x^\beta \text{ and } i > j)$$

gives priority to the monomials in R, whereas the order defined below gives priority to the components of F:

$$x^\alpha e_i > x^\beta e_j \iff i > j \text{ or } (i = j \text{ and } x^\alpha > x^\beta).$$ □

Exercise* 2.2.21 (Division with Remainder in Free Modules). Starting with determinate division with remainder, rewrite our discussion on the division process such that it applies to elements of the free module F. Extend the relevant definitions and results from R to F. □

Exercise 2.2.22. Consider $F = \Bbbk[x, y]^3$ with its canonical basis and the vectors

$$g = \begin{pmatrix} x^2y + x^2 + xy^2 + xy \\ xy^2 - 1 \\ xy + y^2 \end{pmatrix}, \; f_1 = \begin{pmatrix} xy + x \\ 0 \\ y \end{pmatrix}, \; f_2 = \begin{pmatrix} 0 \\ y^2 \\ x + 1 \end{pmatrix} \in F.$$

Extend $>_{\text{lex}}$ on $\Bbbk[x, y]$ to F in the two ways described in Remark 2.2.20. With respect to both orders, find $\mathbf{L}(g)$, $\mathbf{L}(f_1)$, and $\mathbf{L}(f_2)$, and divide g by f_1 and f_2 (use the determinate division algorithm). □

2.3 Gröbner Bases and Buchberger's Algorithm

In Example 2.2.14, with $f_1 = x^2y - y^3$ and $f_2 = x^3$, we computed the standard expressions

$$x^3y = x \cdot f_1 + 0 \cdot f_2 + xy^3$$

and

$$x^3y = y \cdot f_2 + 0 \cdot f_1 + 0,$$

which, in particular, have two different remainders. The problem with the first standard expression is that x^3y and, thus, xy^3 are contained in the ideal $\langle x^2y - y^3, x^3\rangle$, but xy^3 cannot be removed in the division process since it is not divisible by any of the leading terms x^2y and x^3 of the divisors. To decide ideal membership, we need to be able to cancel any leading term of any element of I, using the leading terms of the divisors.

Based on this consideration, we make the following definition:

Definition 2.3.1. Let F be a free $\Bbbk[x_1, \ldots, x_n]$-module with a fixed finite basis, let $>$ be a global monomial order on F, and let $I \subset F$ be a submodule.

1. The **leading submodule** (or **initial submodule**) of I is the monomial submodule
$$\mathbf{L}(I) := \mathbf{L}_>(I) := \big\langle \mathbf{L}_>(f) \bigm| f \in I \big\rangle \subset F.$$
That is, $\mathbf{L}(I)$ is generated by the leading terms of the elements of I. In the special case where I is an ideal of $\Bbbk[x_1, \ldots, x_n]$, we refer to $\mathbf{L}(I)$ as the **leading ideal** (or **initial ideal**) of I.
2. A finite subset $\mathcal{G} = \{f_1, \ldots, f_r\}$ of I is a **Gröbner basis for *I*** if
$$\mathbf{L}_>(I) = \big\langle \mathbf{L}_>(f_1), \ldots, \mathbf{L}_>(f_r) \big\rangle.$$
That is, the leading submodule of I is generated by the leading terms of the elements of $\mathcal{G}$.

For simplicity, we will say that a finite subset $\mathcal{G}$ of F is a **Gröbner basis** if it is a Gröbner basis for the submodule it generates. □

Our terminology in the definition above is somewhat inaccurate in that we should have written leading module with respect to $>$ and Gröbner basis with respect to $>$. Indeed, leading modules depend on the choice of the monomial order. Furthermore, if $\mathcal{G}$ is a Gröbner basis with respect to $>$, and if $>'$ is another monomial order, then $\mathcal{G}$ may fail to be a Gröbner basis with respect to $>'$. See Exercise 2.5.6 below for a simple example. *For the rest of this section, $>$ will be a fixed global monomial order on a free $\Bbbk[x_1, \ldots, x_n]$-module F with a fixed finite basis.*

In contrast to the polynomials f_1, f_2 in Example 2.2.14, the elements of a Gröbner basis behave well under division with remainder and can, thus, be used to decide ideal and submodule membership:

Proposition 2.3.2. *Let $\{f_1, \ldots, f_r\} \subset F \setminus \{0\}$ be a Gröbner basis for the submodule $I := \langle f_1, \ldots, f_r\rangle \subset F$. If $g = \sum_{i=1}^r g_i f_i + h$ is a standard expression for an element $g \in F$, then $g \in I$ iff the remainder h is zero.*

Proof. If h is zero, then clearly $g \in I$. Conversely, if $g \in I$, then $h \in I$, which implies that $\mathbf{L}(h) \in \mathbf{L}(I) = \langle \mathbf{L}(f_1), \ldots, \mathbf{L}(f_r)\rangle$. So $\mathbf{L}(h)$ and, thus, h are zero by condition (ID2) on the remainder of a standard expression. □

Corollary 2.3.3 (Gordan). *Every submodule $I \subset F$ has a Gröbner basis. Furthermore, the elements of any such basis generate I. In particular, $\Bbbk[x_1, \ldots, x_n]$ is Noetherian.*

Proof. As remarked earlier, it follows from Gordan's lemma that every monomial submodule of F is generated by finitely many monomials. In particular, there are finitely many elements $f_1, \ldots, f_r \in I$ such that $\mathbf{L}(I) = \langle \mathbf{L}(f_1), \ldots, \mathbf{L}(f_r)\rangle$. That is, $f_1, \ldots, f_r$ form a Gröbner basis for I. If $\mathcal{G} \subset F \setminus \{0\}$ is any such basis, and $g \in I$ is any element, division with remainder yields a standard expression for g in terms of the elements of $\mathcal{G}$ whose remainder is zero (apply Proposition 2.3.2). In particular, I is generated by $\mathcal{G}$. □

Remark 2.3.4. Gordan's proof has as every other proof of Hilbert's basis theorem two ingredients, namely induction on the number of variables (here used to verify Gordan's Lemma 2.1.2) and division with remainder. The advantage of Gordan's proof is that it separates these ingredients. □

Macaulay (1927) used the idea of obtaining information on an ideal from information on its leading ideal to classify Hilbert functions (see Section 6.4 for Hilbert functions). On his way, he proved the following crucial result:

Theorem-Definition 2.3.5 (Macaulay). *If $I \subset F$ is a submodule, the monomials not contained in $\mathbf{L}_>(I)$ represent a $\Bbbk$-vector space basis for F/I. We refer to these monomials as* **standard monomials** *(for I, with respect to $>$).*

Proof. Let

$$\mathcal{B} := \{m + I \mid m \in F \text{ a standard monomial}\} \subset F/I.$$

To show that the elements of $\mathcal{B}$ *are $\Bbbk$-linearly independent*, consider a $\Bbbk$-linear combination g of standard monomials such that the residue class $g+I$ is zero. Then $g \in I$, so that $\mathbf{L}(g) \in \mathbf{L}(I)$. Since $\mathbf{L}(g)$ is a scalar times a standard monomial, this implies $0 = \mathbf{L}(g) = g$ by the very definition of the standard monomials.

To show that the elements of $\mathcal{B}$ *generate F/I as a $\Bbbk$-vector space*, consider any element $g \in F$. Choose elements $f_1, \dots, f_r \in F\setminus\{0\}$ which form a Gröbner basis for I, and let $g = \sum_{i=1}^r g_i f_i + h$ be a standard expression satisfying condition (DD2) of determinate division with remainder. Then no term of h is in $\langle \mathbf{L}(f_1), \dots, \mathbf{L}(f_r)\rangle = \mathbf{L}(I)$. Hence, the residue class $g + I = h + I$ is a $\Bbbk$-linear combination of the elements of $\mathcal{B}$, as desired. □

Remark-Definition 2.3.6. In the situation of Macaulay's theorem, given $g \in F$, the remainder h in a standard expression $g = \sum_{i=1}^r g_i f_i + h$ satisfying (DD2) is uniquely determined by g, I, and $>$ (and does not depend on the choice of Gröbner basis). It represents the residue class $g + I \in F/I$ in terms of the standard monomials. We write $\mathrm{NF}(g, I) = h$ and call $\mathrm{NF}(g, I)$ the **canonical representative** of $g + I \in F/I$ (or the **normal form** of g mod I), with respect to $>$. □

If a Gröbner basis for an ideal I of $\Bbbk[x_1, \dots, x_n]$ is given, we may use normal forms to perform the sum and product operations in $\Bbbk[x_1, \dots, x_n]/I$ (this is Buchberger's original application of Gröbner bases):

Exercise* 2.3.7. Let $I \subset \Bbbk[x_1, \dots, x_n]$ be an ideal. If $f, g \in \Bbbk[x_1, \dots, x_n]$, show that

$$\mathrm{NF}(f + g, I) = \mathrm{NF}(f, I) + \mathrm{NF}(g, I), \text{ and}$$
$$\mathrm{NF}(f \cdot g, I) = \mathrm{NF}(\mathrm{NF}(f, I) \cdot \mathrm{NF}(g, I), I).$$

□

Following these first indications of the usefulness of Gröbner bases, we, now, treat their computation.

In principle, finding a Gröbner basis for a submodule $I = \langle f_1, \dots, f_r \rangle \subset F$ amounts to adding suitable elements of I to $f_1, \dots, f_r$ such that, eventually, the leading terms of the resulting set of generators for I generate $\mathbf{L}(I)$. A possible approach to detecting new leading terms is to form $\Bbbk[x_1, \dots, x_n]$-linear combinations of $f_1, \dots, f_r$ and divide them by $f_1, \dots, f_r$. Then the remainder is an element of I, and is either zero, or its leading term is not divisible by any of the $\mathbf{L}(f_i)$. In the simplest possible case, we face combinations $g_i f_i + g_j f_j$ involving just two of the generators. To increase our chances of getting a nonzero remainder in this case, we choose g_i and g_j such that $\mathbf{L}(g_i f_i)$ and $\mathbf{L}(g_j f_j)$ cancel each other in $g_i f_i + g_j f_j$:

Definition 2.3.8. Let $f_1, \dots, f_r \in F$ be nonzero polynomial vectors. For each pair of indices i, j, the **S-vector** $\mathrm{S}(f_i, f_j) \in F$ is defined by setting

$$\mathrm{S}(f_i, f_j) = m_{ji} f_i - m_{ij} f_j \in F,$$

where

$$m_{ij} = \mathrm{LCM}(\mathbf{L}(f_i), \mathbf{L}(f_j))/\mathbf{L}(f_j) \in \Bbbk[x_1, \dots, x_n].$$

In the special case where F is the polynomial ring, we say that $\mathrm{S}(f_i, f_j)$ is an **S-polynomial**. □

As it turns out, the division of S-vectors by $f_1, \dots, f_r$ suffices to decide whether $f_1, \dots, f_r$ form a Gröbner basis. Since $\mathrm{S}(f_i, f_j) = -\mathrm{S}(f_j, f_i)$ for all i, j, we only need to consider the $\mathrm{S}(f_i, f_j)$ with $j < i$. In fact, we can do even better: For $i = 2, \dots, r$, let M_i be the monomial ideal

$$M_i = \langle \mathbf{L}(f_1), \dots, \mathbf{L}(f_{i-1}) \rangle : \mathbf{L}(f_i) \subset \Bbbk[x_1, \dots, x_n].$$

Then, by Exercises 2.1.3 and 2.1.5, M_i is generated by the terms

$$m_{ji} = \mathrm{LCM}(\mathbf{L}(f_j), \mathbf{L}(f_i))/\mathbf{L}(f_i),\ j < i.$$

For every i and every *minimal* monomial generator x^α for M_i, choose an index $j = j(i, \alpha) < i$ such that $m_{ji} = c x^\alpha$ for some nonzero scalar $c \in \Bbbk$. Moreover, choose a standard expression for $\mathrm{S}(f_i, f_j)$ in terms of the f_k with remainder $h_{i,\alpha}$ (we suppress the index j in our notation).

Theorem 2.3.9 (Buchberger's Criterion). *Let $f_1, \dots, f_r \in F$ be nonzero polynomial vectors. With notation as above, $f_1, \dots, f_r$ form a Gröbner basis iff all remainders $h_{i,\alpha}$ are zero.* □

In the situation of the criterion, we refer to the selection of the indices $j = j(i, \alpha)$ together with the computation of the remainders $h_{i,\alpha}$ as **Buchberger's test**. It is clear from the criterion that the amount of computation needed for the test depends in a crucial way on the order in which we list $f_1, \dots, f_r$.

Before proving the criterion, we illustrate it by an example. Also, we show how to use the criterion for computing Gröbner bases. For the example, recall that a $k \times k$ **minor** of a matrix is the determinant of a $k \times k$ submatrix.

Example 2.3.10. Consider the ideal generated by the 3×3 minors of the matrix

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ y_1 & y_2 & y_3 & y_4 & y_5 \\ z_1 & z_2 & z_3 & z_4 & z_5 \end{pmatrix}$$

and the lexicographic order on $\Bbbk[x_1, \ldots, z_5]$. The leading terms of the minors and the minimal generators for the corresponding monomial ideals M_i are:

$x_1y_2z_3$	
$x_1y_2z_4$	$M_2 = \langle z_3 \rangle$
$x_1y_3z_4$	$M_3 = \langle y_2 \rangle$
$x_2y_3z_4$	$M_4 = \langle x_1 \rangle$
$x_1y_2z_5$	$M_5 = \langle z_3, z_4 \rangle$
$x_1y_3z_5$	$M_6 = \langle y_2, z_4 \rangle$
$x_2y_3z_5$	$M_7 = \langle x_1, z_4 \rangle$
$x_1y_4z_5$	$M_8 = \langle y_2, y_3 \rangle$
$x_2y_4z_5$	$M_9 = \langle x_1, y_3 \rangle$
$x_3y_4z_5$	$M_{10} = \langle x_1, x_2 \rangle$

So only 15 out of $\binom{10}{2} = 45$ S-vectors are needed in Buchberger's test. The test shows that the minors form a Gröbner basis (we will work this out in Exercise 2.3.21): □

The proof of our next result consists of **Buchberger's algorithm** for computing Gröbner bases:

Corollary 2.3.11. *Given polynomial vectors $f_1, \ldots, f_r \in F \setminus \{0\}$, a Gröbner basis for $I := \langle f_1, \ldots, f_r \rangle \subset F$ can be computed in finitely many steps.*

Proof. If $f_1, \ldots, f_r$ satisfy Buchberger's criterion, we are done. Otherwise, Buchberger's test yields a remainder $0 \neq h \in I$ with $\mathbf{L}(h) \notin \langle \mathbf{L}(f_1), \ldots, \mathbf{L}(f_r) \rangle$. That is, $\langle \mathbf{L}(f_1), \ldots, \mathbf{L}(f_r) \rangle \subsetneq \langle \mathbf{L}(f_1), \ldots, \mathbf{L}(f_r), \mathbf{L}(h) \rangle$. In this case, add $f_{r+1} := h$ to the set of generators, and start over again. After finitely steps, the resulting process must terminate with a Gröbner basis $f_1, \ldots, f_r, f_{r+1}, \ldots, f_{r'}$ for I. Indeed, as a consequence of Gordan's lemma, every ascending chain of (monomial) submodules of F is eventually stationary. □

Example 2.3.12. Let $f_1 = x^2$, $f_2 = xy - y^2 \in \Bbbk[x, y]$ with $>_{\text{lex}}$. Then $\mathbf{L}(f_2) = xy$ and $M_2 = \langle x \rangle$. We compute the standard expression

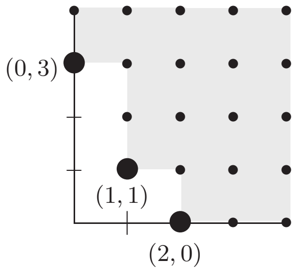
$$\mathrm{S}(f_2, f_1) = x \cdot f_2 - y \cdot f_1 = -xy^2 = 0 \cdot f_1 - y \cdot f_2 - y^3,$$

and add the nonzero remainder $f_3 := -y^3$ to the set of generators. Then $M_3 = \langle x^2, x \rangle = \langle x \rangle$. Computing the standard expression

$$\mathrm{S}(f_3, f_2) = x \cdot f_3 + y^2 \cdot f_2 = -y^4 = 0 \cdot f_1 + 0 \cdot f_2 + y \cdot f_3$$

with remainder zero, we find that f_1, f_2, f_3 form a Gröbner basis for the ideal $I = \langle f_1, f_2 \rangle$.

We visualize, once more, the monomials in $\Bbbk[x, y]$:



The dots in the shaded region correspond to the monomials in the ideal $\mathbf{L}(I)$ which is minimally generated by y^3, xy, and x^2. The monomials $1, x, y, y^2$ respresented outside the shaded region are the standard monomials. Due to Macaulay's Theorem 2.3.5, their residue classes form a $\Bbbk$-vector space basis for $\Bbbk[x, y]/I$. Hence, every class $g + I \in \Bbbk[x, y]/I$ is canonically represented by a uniquely determined $\Bbbk$-linear combination $a + bx + cy + dy^2$ (see Remark 2.3.6). To add and multiply residue classes, we add and multiply the canonical representatives according to the rules in Exercise 2.3.7. The multiplication in $\Bbbk[x, y]/I$ is, thus, determined by the following table (we write $\overline{f} = f + I$):

$\cdot$	1	$\overline{x}$	$\overline{y}$	$\overline{y}^2$
1	1	$\overline{x}$	$\overline{y}$	$\overline{y}^2$
$\overline{x}$	$\overline{x}$	0	$\overline{y}^2$	0
$\overline{y}$	$\overline{y}$	$\overline{y}^2$	$\overline{y}^2$	0
$\overline{y}^2$	$\overline{y}^2$	0	0	0

□

Exercise 2.3.13. Let $f_1 = x^2y - y^3$, $f_2 = x^3 \in \Bbbk[x, y]$ with $>_{\text{lex}}$ as in Example 2.2.14. Compute a Gröbner basis for the ideal $I = \langle f_1, f_2 \rangle$. Visualize the monomials in $\mathbf{L}(I)$, and compute a multiplication table for $\Bbbk[x, y]/I$. □

In general, the products in a multiplication table as above are not represented by terms only:

Exercise 2.3.14. A polynomial in $\Bbbk[x_1, \ldots, x_n]$ is called a **binomial** if it has at most two terms. An ideal of $\Bbbk[x_1, \ldots, x_n]$ is called a **binomial ideal** if it is generated by binomials.

Now let $>$ be any global monomial order on $\Bbbk[x_1, \ldots, x_n]$. Show that the following conditions on an ideal $I \subset \Bbbk[x_1, \ldots, x_n]$ are equivalent:

1. I is a binomial ideal.
2. I has a **binomial Gröbner basis**, that is, a Gröbner basis consisting of binomials.
3. The normal form mod I of any monomial is a term.
4. The multiplication table of $\Bbbk[x_1, \ldots, x_n]/I$ consists of terms only.

See Eisenbud and Sturmfels (1996) for more on binomial ideals. □

We, next, prove Buchberger's criterion. For this, recall that the S-vectors are designed to cancel leading terms:

$$m_{ji}\mathbf{L}(f_i) - m_{ij}\mathbf{L}(f_j) = 0. \tag{2.2}$$

Rewriting the standard expressions

$$\mathrm{S}(f_i, f_j) = g_1^{(ij)} f_1 + \ldots + g_r^{(ij)} f_r + 0$$

with remainder zero as

$$-g_1^{(ij)} f_1 - \cdots + (-m_{ij} - g_j^{(ij)}) f_j - \cdots + (m_{ji} - g_i^{(ij)}) f_i - \cdots - g_r^{(ij)} f_r = 0, \tag{2.3}$$

we may rephrase Buchberger's criterion by saying that $f_1, \ldots, f_r$ form a Gröbner basis iff every relation of type (2.2) considered in Buchberger's test "lifts" to a relation of type (2.3) such that $\mathbf{L}(\mathrm{S}(f_i, f_j)) \geq \mathbf{L}(g_k^{(ij)} f_k)$ whenever both sides are nonzero.

In general, we think of a relation

$$g_1 f_1 + \cdots + g_r f_r = 0 \in F$$

as a column vector $(g_1, \ldots, g_r)^t \in \Bbbk[x_1, \ldots, x_n]^r$, and call it a syzygy on $f_1, \ldots, f_r$:

Definition 2.3.15. Let R be a ring, let M be an R-module, and let $f_1, \ldots, f_r \in M$. A **syzygy** on $f_1, \ldots, f_r$ is an element of the kernel of the homomorphism

$$\phi : R^r \to M, \ \epsilon_i \mapsto f_i,$$

where $\{\epsilon_1, \ldots, \epsilon_r\}$ is the canonical basis of R^r. We call $\ker \phi$ the (first) **syzygy module** of $f_1, \ldots, f_r$, written

$$\mathrm{Syz}\,(f_1, \ldots, f_r) = \ker \phi.$$

If $\mathrm{Syz}\,(f_1, \ldots, f_r)$ is finitely generated, we regard the elements of a given finite set of generators for it as the columns of a matrix which we call a **syzygy matrix** of $f_1, \ldots, f_r$. □

Exercise 2.3.16. Determine a syzygy matrix of $x, y, z \in \Bbbk[x, y, z]$. □

To handle the syzygies on the elements $f_1, \ldots, f_r$ of a Gröbner basis, we consider the free module $F_1 = \Bbbk[x_1, \ldots, x_n]^r$ with its canonical basis $\{\epsilon_1, \ldots, \epsilon_r\}$ and the **induced monomial order** $>_1$ on F_1 defined by setting

$$\begin{aligned} x^\alpha \epsilon_i >_1 x^\beta \epsilon_j \iff \ & x^\alpha \mathbf{L}(f_i) > x^\beta \mathbf{L}(f_j), \ \text{ or} \\ & x^\alpha \mathbf{L}(f_i) = x^\beta \mathbf{L}(f_j) \ \text{ (up to a scalar)} \ \text{ and } \ i > j. \end{aligned}$$

Note that $>_1$ is global if this is true for $>$ (what we suppose, here).

Proof of Buchberger's criterion. Write $R = \Bbbk[x_1, \dots, x_n]$ and $I = \langle f_1, \dots, f_r \rangle \subset F$. If $f_1, \dots, f_r$ form a Gröbner basis for I, all remainders $h_{i,\alpha}$ are zero by Proposition 2.3.2. Indeed, the S-vectors are contained in I.

Conversely, suppose that all the $h_{i,\alpha}$ are zero. Then, for every pair (i, α), we have a standard expression of type

$$\mathrm{S}(f_i, f_j) = g_1^{(ij)} f_1 + \ldots + g_r^{(ij)} f_r + 0,$$

where $j = j(i, \alpha) < i$ is as selected in Buchberger's test. Let

$$G^{(i,\alpha)} := (-g_1^{(ij)}, \dots, -m_{ij} - g_j^{(ij)}, \dots, m_{ji} - g_i^{(ij)}, \dots, -g_r^{(ij)})^t \in F_1 = R^r$$

be the corresponding syzygy on $f_1, \dots, f_r$ (we suppress the index j in our notation on the left hand side). On F_1, we consider the induced monomial order. The leading term of $G^{(i,\alpha)}$ with respect to this order is

$$\mathbf{L}(G^{(i,\alpha)}) = m_{ji} \epsilon_i.$$

Indeed,

$$m_{ji} \mathbf{L}(f_i) = m_{ij} \mathbf{L}(f_j), \text{ but } i > j,$$

and

$$m_{ji} \mathbf{L}(f_i) > \mathbf{L}(\mathrm{S}(f_i, f_j)) \geq \mathbf{L}(g_k^{(ij)}) \mathbf{L}(f_k)$$

whenever these leading terms are nonzero.

To prove that the f_k form a Gröbner basis for I, let g be any nonzero element of I, say $g = a_1 f_1 + \ldots + a_r f_r$, where $a_1, \dots, a_r \in \Bbbk[x_1, \dots, x_n]$. The key point of the proof is to replace this representation of g in terms of the f_k by a standard expression $g = \sum_{k=1}^r g_k f_k$ (with remainder zero). The result, then, follows by applying Remark 2.2.18 on leading terms in standard expressions:

$$\mathbf{L}(g) = \max\{\mathbf{L}(g_1)\mathbf{L}(f_1), \dots, \mathbf{L}(g_r)\mathbf{L}(f_r)\} \in \langle \mathbf{L}(f_1), \dots, \mathbf{L}(f_r) \rangle.$$

To find the desired standard expression, we go back and forth between $F_1 = R^r$ and F: Consider the polynomial vector $A := (a_1, \dots, a_r)^t \in R^r$, and let $G = (g_1, \dots, g_r)^t \in R^r$ be the remainder of A under determinate division by the $G^{(i,\alpha)}$ (listed in some order). Then

$$g = a_1 f_1 + \ldots + a_r f_r = g_1 f_1 + \ldots + g_r f_r \tag{2.4}$$

since the $G^{(i,\alpha)}$ are syzygies on $f_1, \dots, f_r$. We show that the right hand side of (2.4) satisfies condition (DD1) of determinate division by the f_k (in particular, it is a standard expression). Suppose the contrary. Then there is a pair $k < i$ such that one of the terms of $g_i \, \mathbf{L}(f_i)$ is divisible by $\mathbf{L}(f_k)$. In turn, one of the terms of g_i is contained in the monomial ideal

$$M_i = \langle \mathbf{L}(f_1), \ldots, \mathbf{L}(f_{i-1}) \rangle : \mathbf{L}(f_i) \subset R$$

which is generated by the m_{ji} selected in Buchberger's test. In F_1, this means that one of the terms of G is divisible by some $m_{ji}\epsilon_i = \mathbf{L}(G^{(i,\alpha)})$, contradicting the fact that according to how we found G, the terms of G satisfy condition (DD2) of determinate division by the $G^{(i,\alpha)}$ in F_1. □

Corollary 2.3.17. *If $f_1, \ldots, f_r \in F \setminus \{0\}$ form a Gröbner basis with respect to $>$, the $G^{(i,\alpha)}$ considered in the proof of Buchberger's criterion form a Gröbner basis for the syzygy module* Syz $(f_1, \ldots, f_r)$ *with respect to the induced monomial order. In particular, the $G^{(i,\alpha)}$ generate the syzygies on $f_1, \ldots, f_r$.*

Proof. Let $A \in R^r$ be an arbitrary syzygy on $f_1, \ldots, f_r$, and let $G = (g_1, \ldots, g_r) \in R^r$ be the remainder of A under determinate division by the $G^{(i,\alpha)}$ (listed in some order). Then, since A and the $G^{(i,\alpha)}$ are syzygies on $f_1, \ldots, f_r$, the same must be true for G:

$$0 = g_1 f_1 + \ldots + g_r f_r.$$

Furthermore, as shown in the proof of Buchberger's criterion, the g_i satisfy condition (DD1) of determinate division by $f_1, \ldots, f_r$. Since standard expressions under determinate division are uniquely determined, the g_i and, thus, G must be zero. Taking, once more, Remark 2.2.18 into account, we find that $\mathbf{L}(A)$ is divisible by some $\mathbf{L}(G^{(i,\alpha)})$. The result follows. □

Remark 2.3.18. The S in S-vector stands for syzygy. In fact, the relations

$$m_{ji}\mathbf{L}(f_i) - m_{ij}\mathbf{L}(f_j) = 0 \tag{2.5}$$

corresponding to the S-vectors $\mathrm{S}(f_i, f_j)$ generate Syz $(\mathbf{L}(f_1), \ldots, \mathbf{L}(f_r))$. In our version of Buchberger's test, selecting the m_{ji} for all i means that we select a subspace $X \subset \{\mathrm{S}(f_i, f_j) \mid j < i\}$ such that the relations (2.5) corresponding to the S-vectors in X still generate Syz $(\mathbf{L}(f_1), \ldots, \mathbf{L}(f_r))$. It is this property of X on which our proof of Buchberger's criterion is based. Hence, in stating the criterion, we can choose any set of S-vectors satisfying this property. □

Remark 2.3.19. Let $f_1, \ldots, f_r \in F \setminus \{0\}$, and let $I = \langle f_1, \ldots, f_r \rangle \subset F$. If we compute a Gröbner basis $f_1, \ldots, f_r, f_{r+1}, \ldots, f_{r'}$ for I using Buchberger's algorithm, the syzygies $G^{(i,\alpha)}$ generating Syz $(f_1, \ldots, f_r, f_{r+1}, \ldots, f_{r'})$ are obtained in two ways. Either, $G^{(i,\alpha)}$ arises from a division leading to a new generator f_k, $k \geq r+1$:

$$\mathrm{S}(f_i, f_j) = g_1^{(ij)} f_1 + \ldots + g_{k-1}^{(ij)} f_{k-1} + f_k.$$

Or, $G^{(i,\alpha)}$ arises from a division with remainder zero:

$$\mathrm{S}(f_i, f_j) = g_1^{(ij)} f_1 + \ldots + g_\ell^{(ij)} f_\ell + 0.$$

□

Example 2.3.20. In Example 2.3.12, the matrix

$$\begin{pmatrix} -y & 0 \\ x+y & y^2 \\ -1 & x-y \end{pmatrix}$$

is a syzygy matrix of $f_1 = x^2, f_2 = xy - y^2, f_3 = -y^3 \in \mathbb{k}[x, y]$. □

Exercise 2.3.21. Consider the ideal generated by the 3×3 minors of the matrix

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ y_1 & y_2 & y_3 & y_4 & y_5 \\ z_1 & z_2 & z_3 & z_4 & z_5 \end{pmatrix}$$

and the lexicographic order on $\mathbb{k}[x_1, \dots, z_5]$ as in Example 2.3.10. Prove that the minors form a Gröbner basis, and show that their syzygy module is generated by 15 elements. Referring to the syzygies on these 15 (first order) syzygies as the second order syzygies on the minors, how many elements of $\mathbb{k}[x_1, \dots, z_5]^{15}$ do we need to generate the second order syzygies?
Hint. In this example, lengthy computations can be avoided by using Laplace expansion. □

A Gröbner basis $\{f_1, \dots, f_r\} \subset F$ computed with Buchberger's algorithm quite often contains elements whose leading terms are unneeded generators for $\mathbf{L}(\langle f_1, \dots, f_r\rangle)$. By eliminating superfluous generators and by adjusting constants such that the coefficient of each leading term is 1, we get a **minimal Gröbner basis**, that is, a Gröbner basis whose leading terms are the minimal generators for $\mathbf{L}(\langle f_1, \dots, f_r\rangle)$. In addition, we may "reduce the tail" of each element in the Gröbner basis:

Exercise* 2.3.22. A minimal Gröbner basis $\{f_1, \dots, f_r\} \subset F$ is **reduced** if, for $i \neq j$, no term of f_i is divisible by $\mathbf{L}(f_j)$. Show that if $\langle 0\rangle \neq I \subset F$ is a submodule, there is a uniquely determined reduced Gröbner basis for I with respect to the given monomial order, namely

$$m_1 - \mathrm{NF}(m_1, I), \dots, m_r - \mathrm{NF}(m_r, I),$$

where $m_1, \dots, m_r$ are the minimal generators for $\mathbf{L}(I)$. Explain how to compute the reduced Gröbner basis from any given Gröbner basis. □

Remark 2.3.23. Buchberger's algorithm generalizes both Gaussian elimination and Euclid's GCD algorithm:

1. Given homogeneous degree-1 polynomials

$$f_i = a_{i1}x_1 + \cdots + a_{in}x_n \in \mathbb{k}[x_1, \dots, x_n],\ i = 1, \dots, r,$$

let $>$ be a global monomial order on $\mathbb{k}[x_1, \dots, x_n]$ such that $x_1 > \cdots > x_n$. Computing a minimal Gröbner basis for $\langle f_1, \dots, f_r\rangle$ amounts, then, to transforming the coefficient matrix $A = (a_{ij})$ into a matrix in row echelon form with pivots 1.

2. In the case of one variable x, there is precisely one global monomial order: $\dots > x^2 > x > 1$. Given $f_1, f_2 \in \Bbbk[x]$, the reduced Gröbner basis for $\langle f_1, f_2 \rangle$ with respect to this order consists of exactly one element, namely the greatest common divisor $\mathrm{GCD}(f_1, f_2)$, and Buchberger's algorithm takes precisely the same steps as Euclid's algorithm for computing the GCD. □

2.4 First Applications

As already remarked, division with remainder, Proposition 2.3.2, and Buchberger's algorithm allow us to decide submodule (ideal) membership:

Algorithm 2.4.1 (Submodule Membership). *Given a free module F over $\Bbbk[x_1, \dots, x_n]$ with a fixed finite basis, and given nonzero elements g, $f_1, \dots, f_r \in F$, decide whether*

$$g \in I := \langle f_1, \dots, f_r \rangle \subset F.$$

[If so, express g as a $\Bbbk[x_1, \dots, x_n]$-linear combination

$$g = g_1 f_1 + \ldots + g_r f_r.]$$

1. *Compute a Gröbner basis $f_1, \dots, f_r, f_{r+1}, \dots, f_{r'}$ for I using Buchberger's algorithm. [Store each syzygy arising from a division which leads to a new generator f_k in Buchberger's test.]*
2. *Compute a standard expression for g in terms of $f_1, \dots, f_{r'}$ with remainder h (use the same global monomial order on F as in Step 1).*
3. *If $h = 0$, then $g \in I$. [In this case, for $k = r', \dots, r+1$, successively do the following: in the standard expression computed in Step 2, replace f_k by the expression in terms of $f_1, \dots, f_{k-1}$ given by the syzygy leading to f_k in Step 1.]* □

Example 2.4.2. In Example 2.3.12, we computed the lexicographic Gröbner basis

$$f_1 = x^2, \quad f_2 = xy - y^2, \quad f_3 = -y^3$$

for the ideal $I = \langle f_1, f_2 \rangle \subset \Bbbk[x, y]$. Dividing

$$g = x^3 - x^2 + xy^2$$

by f_1, f_2, f_3, we get the standard expression $g = (x-1) \cdot f_1 + y \cdot f_2 - f_3$ with remainder zero. Hence, $g \in I$. Substituting, then, $(x+y) \cdot f_2 - y \cdot f_1$ for f_3 in the standard expression (see Example 2.3.20), we find that

$$g = (x - 1 + y) \cdot f_1 - x \cdot f_2.$$

□

Exercise* **2.4.3 (Equality of Submodules).** Let F be a free $\Bbbk[x_1, \ldots, x_n]$-module, and let $I, J \subset F$ be submodules. Note that if I, J are monomial submodules, then $I = J$ iff I and J have the same minimal generators. Show that if $>$ is any global monomial order on F, then

$$I \subset J \text{ and } \mathbf{L}(I) = \mathbf{L}(J) \implies I = J.$$

Finally, note that if I, J are arbitrary, then

$$I = J \iff I = I + J \text{ and } J = I + J.$$

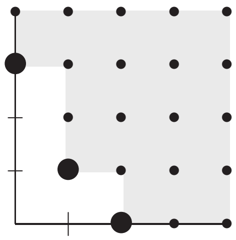
□

As already remarked earlier, Algorithm 2.4.1 can be used to decide solvability. More generally, inspecting the Gröbner basis computed in the first step of the algorithm, we get the following information on the set of solutions (see the Nullstellensatz, Exercise 1.6.5, and Macaulay's Theorem 2.3.5):

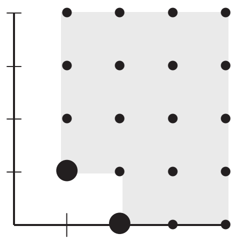
Remark 2.4.4. Let $f_1, \ldots, f_r \in \Bbbk[x_1, \ldots, x_n] \setminus \{0\}$, let $I := \langle f_1, \ldots, f_r \rangle$, and let $\overline{\Bbbk}$ be the algebraic closure of $\Bbbk$. Then we can determine whether the system

$$f_1(x_1, \ldots, x_n) = 0, \ldots, f_r(x_1, \ldots, x_n) = 0$$

has no solution in $\mathbb{A}^n(\overline{\Bbbk})$, at most finitely many solutions in $\mathbb{A}^n(\overline{\Bbbk})$, or infinitely many solutions in $\mathbb{A}^n(\overline{\Bbbk})$ by checking whether any monomial in $x_1, \ldots, x_n$ is contained in $\mathbf{L}(I)$, at most finitely many monomials are not contained in $\mathbf{L}(I)$, or infinitely many monomials are not contained in $\mathbf{L}(I)$. In terms of a Gröbner basis $\mathcal{G}$ for I, the first condition means that at least one element of $\mathcal{G}$ is a nonzero constant. The second condition means that, for any $1 \leq i \leq n$, there is an element of $\mathcal{G}$ whose leading monomial is of type $x_i^{\alpha_i}$ for some $\alpha_i \geq 1$.



Finitely many solutions.



Infinitely many solutions.

Note that though our check gives a result over $\overline{\Bbbk}$, the actual Gröbner basis computation is carried through over $\Bbbk$ (see Section 2.7 for more remarks on the role of the ground field). □

Exercise* **2.4.5.** If the system defined by $f_1, \ldots, f_r \in \Bbbk[x_1, \ldots, x_n]$ has only finitely many solutions in $\mathbb{A}^n(\overline{\Bbbk})$, prove that the number of these solutions is at most $\dim_{\Bbbk} \Bbbk[x_1, \ldots, x_n]/\langle f_1, \ldots, f_r \rangle$. That is, the number of monomials not in $\mathbf{L}(f_1, \ldots, f_r)$ is an upper bound for the number of solutions. Show that these numbers are equal if $\langle f_1, \ldots, f_r \rangle$ is a radical ideal. □

Exercise 2.4.6. If $I \subset \mathbb{Q}[x,y,z]$ is the ideal generated by the polynomials

$$\begin{aligned} f_1 &= 3xz + 4x - 2y - z^2 - 4z,\\ f_2 &= -2x + 3yz - 2y + 2z^2 - z,\\ f_3 &= -3xy + 5x + 3y^2 - y - 2z^2 - 2z, \end{aligned}$$

show that the reduced Gröbner basis for I with respect to $>_{\text{lex}}$ is given by the polynomials

$$\begin{aligned} g_1 &= x - 1/12z^4 + 1/3z^3 + 1/12z^2 - 4/3z,\\ g_2 &= y + 1/3z^4 + 1/6z^3 - 4/3z^2 - 1/6z,\\ g_3 &= z^5 - 5z^3 + 4z. \end{aligned}$$

Deduce from the new set of generators that the locus of zeros $\mathrm{V}(I) \subset \mathbb{A}^3(\overline{\mathbb{Q}})$ consists of precisely five points:

$$(0,0,0),\ (1,1,1),\ (-1,1,-1),\ (1,-1,2),\ (1,1,-2).$$

Note that the number of solutions is exactly $\dim_{\mathbb{Q}} \mathbb{Q}[x,y,z]/I$: the five monomials z^i, $0 \le i \le 4$, represent a $\mathbb{Q}$-vector space basis for $\mathbb{Q}[x,y,z]/I$. □

Exercise 2.4.7. If $I \subset \mathbb{Q}[x,y,z]$ is the ideal generated by the polynomials

$$\begin{aligned} f_1 &= x^3 + y^3 + z^3 - 1,\\ f_2 &= x^2 + y^2 + z^2 - 1,\\ f_3 &= x + y + z - 1, \end{aligned}$$

show that the reduced Gröbner basis for I with respect to $>_{\text{lex}}$ is given by the polynomials

$$\begin{aligned} g_1 &= x + y + z - 1\\ g_2 &= y^2 + yz - y + z^2 - z,\\ g_3 &= z^3 - z^2. \end{aligned}$$

Conclude that $\dim_{\mathbb{Q}} \mathbb{Q}[x,y,z]/I = 6$ though there are only three solutions in $\mathbb{A}^3(\overline{\mathbb{Q}})$:

$$(1,0,0),\ (0,1,0)\,,(0,0,1).$$ □

We already know that Buchberger's algorithm computes the syzygies on the elements of a Gröbner basis (see Corollary 2.3.17). Based on this, we can compute the syzygies on any given set of generators:

Algorithm 2.4.8 (Syzygy Modules). *Given a free $\Bbbk[x_1,\dots,x_n]$-module F with a fixed finite basis and polynomial vectors $f_1,\dots,f_r \in F \setminus \{0\}$, compute a syzygy matrix of $f_1,\dots,f_r$.*

1. *Compute a Gröbner basis $f_1,\dots,f_r,f_{r+1},\dots,f_{r'}$ for $\langle f_1,\dots,f_r\rangle \subset F$ using Buchberger's algorithm. On your way, store each syzygy on $f_1,\dots,f_{r'}$ obtained in Buchberger's test. Let t be the number of these syzygies.*

2. *Arrange the syzygies such that those obtained from a division leading to a new generator f_k are first (and those arising from a division with remainder zero are second). Then the syzygies fit as columns into an $r' \times t$ matrix which has block form $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, where C is an upper triangular square matrix of size $r' - r$ with diagonal entries* 1 *(if signs are adjusted appropriately).*
3. *The $r \times (t - r' + r)$ matrix $B - AC^{-1}D$ is a syzygy matrix of $f_1, \dots, f_r$.*

Proof (of correctness). By Corollary 2.3.17, the columns of $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ generate all the syzygies on $f_1, \dots, f_r, f_{r+1}, \dots, f_{r'}$. We multiply $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ with the invertible $t \times t$ matrix $\begin{pmatrix} E_{r'-r} & -C^{-1}D \\ 0 & E_{t-r'+r} \end{pmatrix}$, where E_j stands for the $j \times j$ identity matrix. As a result, we obtain new generators for the syzygies, namely the columns of the matrix $M = \begin{pmatrix} A & B - AC^{-1}D \\ C & 0 \end{pmatrix}$. A $\Bbbk[x_1, \dots, x_n]$-linear combination of the columns of M defines a syzygy just on $f_1, \dots, f_r$ iff its last $r' - r$ entries are zero. It is, then, a $\Bbbk[x_1, \dots, x_n]$-linear combination of the last $t - r' + r$ columns of M since C has maximal rank. We conclude that $B - AC^{-1}D$ is a syzygy matrix of $f_1, \dots, f_r$. □

Example 2.4.9. Recall from Exercise 2.3.20 how we computed the lexicographic Gröbner basis $f_1 = x^2, f_2 = xy - y^2, f_3 = -y^3$ for the ideal $I = \langle f_1, f_2 \rangle \subset \Bbbk[x, y]$. With conventions as in Algorithm 2.4.8, the resulting syzygy matrix is

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} y & 0 \\ -x - y & y^2 \\ 1 & x - y \end{pmatrix}.$$

Thus, Syz (f_1, f_2) is generated by the single syzygy

$$\begin{pmatrix} 0 \\ y^2 \end{pmatrix} - \begin{pmatrix} y \\ -x - y \end{pmatrix} (x - y) = \begin{pmatrix} -f_2 \\ f_1 \end{pmatrix}.$$

□

We give two examples of how syzygy computations may be used to perform operations on ideals (the same ideas work, more generally, for submodules of free modules). We begin with ideal intersections (geometrically, with the union of algebraic sets). The correctness of the algorithm is obvious.

Algorithm 2.4.10 (Ideal Intersection). *Given ideals $I = \langle f_1, \dots, f_r \rangle$ and $J = \langle g_1, \dots, g_s \rangle$ of $R = \Bbbk[x_1, \dots, x_n]$, compute generators for the intersection*

$$I \cap J.$$

1. *Compute the kernel of the map $R^{r+s+1} \to R^2$ with matrix*

$$\begin{pmatrix} f_1 \dots f_r & 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & g_1 \dots g_s & 1 \end{pmatrix}.$$

   That is, compute a syzygy matrix of the columns of the matrix.
2. *The entries of the last row of the syzygy matrix generate $I \cap J$.* □

Now we deal with ideal quotients and saturation (geometrically, with the Zariski closure of the difference of two algebraic sets):

Exercise* 2.4.11. Let I and J be ideals of $\Bbbk[x_1, \dots, x_n]$. Design algorithms for computing $I : J$ and $I : J^\infty$.
Hint. For $I : J$, if $I = \langle f_1, \dots, f_r \rangle$ and $J = \langle g_1, \dots, g_s \rangle$, consider the matrix

$$\begin{pmatrix} f_1 \dots f_r & 0 & \dots & & \dots & 0 & g_1 \\ 0 & \dots & 0 & f_1 \dots f_r & 0 & \dots & 0 & g_2 \\ \vdots & & & & \ddots & & \vdots \\ 0 & \dots & & \dots & 0 & f_1 & \dots f_r & g_s \end{pmatrix}.$$

For $I : J^\infty$, proceed by iteration. □

The following exercise contains examples of how these algorithms work:

Exercise 2.4.12. Let $\Bbbk$ be infinite. Consider the ideal

$$I = \langle xz - y^2, x^2 - y \rangle \subset \Bbbk[x, y, z].$$

1. Observe that the line $\mathrm{V}(x, y)$ is contained in $\mathrm{V}(I) \subset \mathbb{A}^3(\Bbbk)$.
2. Compute that $I : \langle x, y \rangle = I : \langle x, y \rangle^\infty = \mathrm{I}(C)$, where $\mathrm{I}(C) = \langle x^2 - y, xy - z \rangle$ is the vanishing ideal of the twisted cubic curve $C \subset \mathbb{A}^3(\Bbbk)$.
3. Compute that $I = \langle x, y \rangle \cap \mathrm{I}(C)$. Conclude that this intersection is a primary decomposition of I and that $\mathrm{V}(I) = \mathrm{V}(x, y) \cup C$ is the decomposition of $\mathrm{V}(I)$ into its irreducible components. □

The example in the exercise shows, in particular, that the intersection of two varieties needs not be a variety.

Remark 2.4.13. The arguments used in the exercise are special to the case $I = \langle xz - y^2, x^2 - y \rangle$. A more sytematic approach to decomposing ideals is provided by a number of algorithms for computing radicals and, more generally, primary decomposition. These algorithms are quite involved. Typically, they use Gröbner basis methods (or other means of manipulating ideals) to reduce to the hypersurface case, and algorithms for square-free decomposition and, more generally, polynomial factorization to settle the hypersurface case. We will not discuss any details in this book. See Decker, Greuel, and Pfister (1999) for a survey on algorithms for primary decomposition, and Kaltofen (1982, 1990, 1992, 2003) for the history of polynomial factorization. □

2.5 The Use of Different Monomial Orders

Buchberger's algorithm requires the choice of a global monomial order. The performance of the algorithm and the resulting Gröbner basis depend in a crucial way on this choice. For most applications, in principle, any Gröbner basis and, thus, any order will do. With regard to efficiency, however, the global monomial order defined next appears to be best possible (see the discussion following Example 2.5.2 below and Bayer and Stillman (1987) for some remarks in this direction):

Definition 2.5.1. We define the **degree reverse lexicographic order** on $\Bbbk[x_1, \dots, x_n]$ by setting

$$x^\alpha >_{\rm drlex} x^\beta \iff \begin{array}{l}\deg x^\alpha > \deg x^\beta, \text{ or } (\deg x^\alpha = \deg x^\beta \text{ and the} \\ \text{last nonzero entry of } \alpha - \beta \in \mathbb{Z}^n \text{ is negative}).\end{array}$$

This order is extended to free $\Bbbk[x_1, \dots, x_n]$-modules as in Remark 2.2.20 (we suggest to give priority to the monomials in $\Bbbk[x_1, \dots, x_n]$). □

Note that as in the case of $>_{\rm lex}$, we have defined $>_{\rm drlex}$ such that the variables are ordered according to their appearance when writing $\Bbbk[x_1, \dots, x_n]$. In contrast to $>_{\rm lex}$, however, $>_{\rm drlex}$ refines the partial order by total degree:

$$\deg x^\alpha > \deg x^\beta \Longrightarrow x^\alpha >_{\rm drlex} x^\beta.$$

We will refer to this fact by saying that $>_{\rm drlex}$ is **degree-compatible**.

Example 2.5.2. With respect to $>_{\rm lex}$ and $>_{\rm drlex}$, the monomials of degree 2 in $\Bbbk[x, y, z]$ are ordered as follows:

$$x^2 >_{\rm lex} xy >_{\rm lex} xz >_{\rm lex} y^2 >_{\rm lex} yz >_{\rm lex} z^2$$

and

$$x^2 >_{\rm drlex} xy >_{\rm drlex} y^2 >_{\rm drlex} xz >_{\rm drlex} yz >_{\rm drlex} z^2.$$ □

For monomials of the same degree, the difference between $>_{\rm lex}$ and $>_{\rm drlex}$ is subtle but crucial. The use made of these orders relies on their key properties (which, as we will see in Exercise 2.9.4, characterize $>_{\rm lex}$ and $>_{\rm drlex}$ among all global monomial orders).

The key property of $>_{\rm drlex}$ is: $>_{\rm drlex}$ is degree-compatible, and if $f \in \Bbbk[x_1, \dots, x_n]$ is homogeneous, then

$$\begin{array}{c} >_{\rm drlex} \text{ chooses the leading term of } f \text{ in a subring } \Bbbk[x_1, \dots, x_k] \\ \text{such that } k \text{ is as small as possible.}\end{array}$$

This property has usually the effect that, compared to other global monomial orders, the monomial ideals M_i in Buchberger's test have fewer minimal generators.

The key property of $>_{\rm lex}$ is that the following holds for all $f \in \Bbbk[x_1, \dots, x_n]$:

$$\mathbf{L}(f) \in \Bbbk[x_{k+1}, \dots, x_n] \text{ for some } k \implies f \in \Bbbk[x_{k+1}, \dots, x_n].$$

This makes $>_{\text{lex}}$ useful for eliminating variables, an application of Buchberger's algorithm which requires the computation of special Gröbner bases and, thus, the choice of special monomial orders.

Definition 2.5.3. If $I \subset \Bbbk[x_1, \dots, x_n]$ is an ideal, its ***k*th elimination ideal** is the ideal

$$I_k = I \cap \Bbbk[x_{k+1}, \dots, x_n]. \qquad \square$$

In particular, $I_0 = I$.

Algorithm 2.5.4 (Elimination Using $>_{\text{lex}}$). *Given $I = \langle f_1, \dots, f_r \rangle \subset \Bbbk[x_1, \dots, x_n]$, compute all elimination ideals I_k.*

1. *Compute a Gröbner basis $\mathcal{G}$ for I with respect to $>_{\text{lex}}$ on $\Bbbk[x_1, \dots, x_n]$.*
2. *For any k, the elements $g \in \mathcal{G}$ with $\mathbf{L}(g) \in \Bbbk[x_{k+1}, \dots, x_n]$ form a Gröbner basis for I_k with respect to $>_{\text{lex}}$ on $\Bbbk[x_{k+1}, \dots, x_n]$.*

Proof (of correctness). If $f \in I \cap \Bbbk[x_{k+1}, \dots, x_n]$, then $\mathbf{L}(f)$ is divisible by $\mathbf{L}(g)$ for some $g \in \mathcal{G}$. Since f does not involve $x_1, \dots, x_k$, the same holds for $\mathbf{L}(g)$ and, thus, also for g due to the key property of $>_{\text{lex}}$. $\square$

Example 2.5.5. Let $I = \langle f_1, f_2 \rangle \subset \Bbbk[x, y, z]$, with $f_1 = x^2 - y$, $f_2 = xy - z$. Then I is the vanishing ideal of the twisted cubic curve. We compute a lexicographic Gröbner basis for I. To begin with, $M_2 = \langle x^2 \rangle : xy = \langle x \rangle$, and we have the standard expression

$$\mathrm{S}(f_2, f_1) = x(xy - z) - y(x^2 - y) = -xz + y^2 =: f_3.$$

We add f_3 to the set of generators. Then $M_3 = \langle x^2, xy \rangle : xz = \langle x, y \rangle$, and we have the standard expressions

$$\mathrm{S}(f_3, f_1) = x(-xz + y^2) + z(x^2 - y) = xy^2 - yz = y(xy - z)$$

and

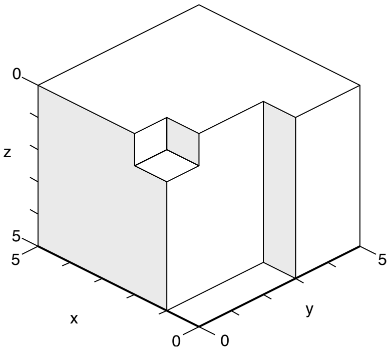
$$\mathrm{S}(f_3, f_2) = y(-xz + y^2) + z(xy - z) = y^3 - z^2 =: f_4.$$

In the next step, $M_4 = \langle x^2, xy, xz \rangle : y^3 = \langle x \rangle$, and

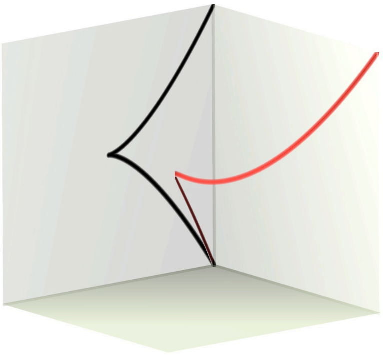
$$\mathrm{S}(f_4, f_2) = x(y^3 - z^2) - y^2(xy - z) = -xz^2 + y^2 z = z(-xz + y^2)$$

is a standard expression with remainder zero. Hence, f_1, f_2, f_3, f_4 form a Gröbner basis for I.

We visualize the monomials in $\mathbf{L}(I)$ via their exponent vectors:



The computaion shows that the elimination ideal $I_1 \subset \mathbb{k}[y, z]$ is generated by the polynomial $f_4 = y^3 - z^2$. As will become clear in Section 2.6 below, the geometric interpretation of this is:



In the yz-plane, the equation $f_4 = 0$ defines the image of the twisted cubic curve $\mathrm{V}(f_1, f_2)$ under the projection which sends (a, b, c) to (b, c).

□

Exercise 2.5.6. In the previous example, $f_1 = x^2 - y$, $f_2 = xy - z$, $-f_3 = xz - y^2$, and $f_4 = y^3 - z^3$ form the reduced lexicographic Gröbner basis for the ideal $I = \langle f_1, f_2 \rangle$. In contrast, show that f_1, f_2, and f_3 form the reduced Gröbner basis with respect to $>_{\text{drlex}}$. □

A single Gröbner basis computation with respect to $>_{\text{lex}}$ yields the whole flag of elimination ideals I_k, $k = 0, \dots, n-1$. If only one of the elimination ideals is needed, other monomial orders are usually more effective.

Definition 2.5.7. A monomial order $>$ on the polynomial ring

$$\mathbb{k}[\boldsymbol{x}, \boldsymbol{y}] = \mathbb{k}[x_1, \dots, x_n, y_1, \dots y_m]$$

is an **elimination order** with respect to $x_1, \dots, x_n$ if the following holds for all $f \in \mathbb{k}[\boldsymbol{x}, \boldsymbol{y}]$:

$$\mathbf{L}(f) \in \mathbb{k}[\boldsymbol{y}] \implies f \in \mathbb{k}[\boldsymbol{y}].$$

□

Example 2.5.8. Let $>_1$ on $\Bbbk[\boldsymbol{x}]$ and $>_2$ on $\Bbbk[\boldsymbol{y}]$ be monomial orders. The **product order** (or **block order**) $> = (>_1, >_2)$ on $\Bbbk[\boldsymbol{x}, \boldsymbol{y}]$ is defined by

$$x^\alpha y^\gamma > x^\beta y^\delta \iff x^\alpha >_1 x^\beta, \text{ or } (x^\alpha = x^\beta \text{ and } y^\gamma >_2 y^\delta).$$

It is an elimination order with respect to $x_1, \dots, x_n$ which is global if $>_1$ and $>_2$ are global. Choosing $>_1$ and $>_2$ to be degree reverse lexicographic is often most efficient. □

As for $>_{\text{lex}}$, one shows:

Proposition 2.5.9 (Elimination). *Let $I \subset \Bbbk[\boldsymbol{x}, \boldsymbol{y}]$ be an ideal, let $>$ be a global elimination order on $\Bbbk[\boldsymbol{x}, \boldsymbol{y}]$ with respect to $x_1, \dots, x_n$, and let $\mathcal{G}$ be a Gröbner basis for I with respect to $>$. Then $\mathcal{G} \cap \Bbbk[\boldsymbol{y}]$ is a Gröbner basis for $I \cap \Bbbk[\boldsymbol{y}]$ with respect to the restriction of $>$ to $\Bbbk[\boldsymbol{y}]$.* □

Remark 2.5.10. Computing a Gröbner basis $\mathcal{G}$ for I with respect to an elimination order may be costly. It is usually much faster to proceed along the following lines. First, compute a Gröbner basis $\mathcal{G}'$ for I with respect to $>_{\text{drlex}}$. Then apply a **Gröbner walk algorithm** which, starting from $\mathcal{G}'$, approaches the target Gröbner basis $\mathcal{G}$ in several steps, "walking" along a path through the **Gröbner fan** of I (see Sturmfels (1996) for the Gröbner fan). In each step, a Gröbner basis with respect to an "intermediate order" is computed. There are several strategies for choosing the path through the Gröbner fan, leading to different variants of the algorithm (see Decker and Lossen (2006) and the references cited there). A completely different approach to computing Gröbner bases with respect to slow orders makes use of Hilbert functions (see Remark 6.4.44). □

As already indicated in Example 2.5.5, the geometric meaning of elimination is projection. We will treat this systematically in Section 2.6 below. Applying projection to the graph of an arbitrary morphism φ, given by polynomials $f_1, \dots, f_m \in \Bbbk[x_1, \dots, x_n]$, we will find a way of computing the Zariski closure of the image of a given algebraic set under φ. The corresponding algebraic result is our next topic in this section. Rather than considering R-module relations as in Definition 2.3.15, we are, now, interested in $\Bbbk$-algebra relations:

Definition 2.5.11. Let S be a $\Bbbk$-algebra, and let $s_1, \dots, s_m$ be elements of S. A **$\Bbbk$-algebra relation** on $s_1, \dots, s_m$ is a polynomial expression of type

$$\sum c_\alpha s_1^{\alpha_1} \cdots s_m^{\alpha_m} = 0 \in S,$$

with coefficients $c_\alpha \in \Bbbk$. Formally, consider a polynomial ring $\Bbbk[y_1, \dots, y_m]$ and think of a $\Bbbk$-algebra relation as an element of the kernel of the homomorphism

$$\phi : \Bbbk[y_1, \dots, y_m] \to S, \; y_i \mapsto s_i.$$

If only the trivial such relation exists, $s_1, \dots, s_m$ are **algebraically independent** over $\Bbbk$. □

Proposition 2.5.12 (Algebra Relations in Affine Rings). *Let I be an ideal of $\mathbb{k}[x_1,\dots,x_n]$, and let $\overline{f}_1 = f_1 + I, \dots, \overline{f}_m = f_m + I \in \mathbb{k}[x_1,\dots,x_n]/I$. Consider the homomorphism*

$$\phi : \mathbb{k}[y_1,\dots,y_m] \to S = \mathbb{k}[x_1,\dots,x_n]/I,\ y_i \mapsto \overline{f}_i.$$

If J is the ideal

$$J = I\,\mathbb{k}[\boldsymbol{x},\boldsymbol{y}] + \langle f_1 - y_1,\dots,f_m - y_m\rangle \subset \mathbb{k}[\boldsymbol{x},\boldsymbol{y}],$$

then

$$\ker\phi = J \cap \mathbb{k}[\boldsymbol{y}].$$

Proof. Let $g \in \mathbb{k}[\boldsymbol{y}] \subset \mathbb{k}[\boldsymbol{x},\boldsymbol{y}]$. To prove the assertion, we have to show:

$$g(f_1,\dots,f_m) \in I \iff g \in J.$$

If $g = h + \sum g_j(f_j - y_j) \in J$, with $h \in I\,\mathbb{k}[\boldsymbol{x},\boldsymbol{y}]$, then $g(f_1,\dots,f_m) = h(x_1,\dots,x_n,f_1,\dots,f_m) \in I\,\mathbb{k}[\boldsymbol{x},\boldsymbol{y}] \cap \mathbb{k}[\boldsymbol{x}] = I$.

For the converse, observe that substituting the $f_j - (f_j - y_j)$ for the y_j in g and expanding gives an expression of type

$$g(y) = g(f_1,\dots,f_m) + \sum g_j(f_j - y_j). \qquad \square$$

Since we already know how to compute in affine rings, a particular application of the proposition is a method for computing in the algebra $\mathbb{k}[\overline{f}_1,\dots,\overline{f}_m] \cong \mathbb{k}[y_1,\dots,y_m]/\ker\phi$. Once we have the required Gröbner basis for J, we know, in particular, whether $\ker\phi = 0$. That is, we can decide whether $\overline{f}_1,\dots,\overline{f}_m$ are algebraically independent over $\mathbb{k}$. In addition, we can check whether ϕ is surjective:

Exercise* 2.5.13 (Subalgebra Membership). With notation as above, let $\overline{g},\overline{f}_1,\dots,\overline{f}_m$ be elements $\mathbb{k}[x_1,\dots,x_n]/I$, and let $>$ be a global elimination order on $\mathbb{k}[\boldsymbol{x},\boldsymbol{y}]$ with respect to $x_1,\dots,x_n$. Show:

1. We have $\overline{g} \in \mathbb{k}[\overline{f}_1,\dots,\overline{f}_m]$ iff the normal form $h = \mathrm{NF}(g,J) \in \mathbb{k}[\boldsymbol{x},\boldsymbol{y}]$ is contained in $\mathbb{k}[\boldsymbol{y}]$. In this case, $\overline{g} = h(\overline{f}_1,\dots,\overline{f}_m)$ is a polynomial expression for $\overline{g}$ in terms of the $\overline{f}_k$.
2. The homomorphism $\phi : \mathbb{k}[y_1,\dots,y_m] \to \mathbb{k}[x_1,\dots,x_n]/I$ is surjective iff $\mathrm{NF}(x_i,J) \in \mathbb{k}[\boldsymbol{y}]$ for $i = 1,\dots,n$. $\square$

Exercise 2.5.14. 1. Compute the algebra relations on the polynomials

$$f_1 = x^2 + y^2,\ f_2 = x^2y^2,\ f_3 = x^3y - xy^3 \in \mathbb{k}[x,y].$$

2. Consider the polynomials

$$g = x^4 + y^4,\ g_1 = x + y,\ g_2 = xy \in \mathbb{k}[x,y].$$

Show that g is contained in the subalgebra $\mathbb{k}[g_1,g_2] \subset \mathbb{k}[x,y]$, and express g as a polynomial in g_1, g_2.

3. Consider the endomorphism ϕ of $\mathbb{k}[x_1, x_2, x_3]$ defined by

$$x_1 \mapsto x_2x_3,\ x_2 \mapsto x_1x_3,\ x_3 \mapsto x_1x_2.$$

Prove that ϕ induces an automorphism of

$$\mathbb{k}[x_1, x_2, x_3]/\langle x_1x_2x_3 - 1\rangle.$$

This means that the variety $A = \mathrm{V}(x_1x_2x_3 - 1) \subset \mathbb{A}^3(\mathbb{k})$ admits a nonlinear automorphism. Determine the fixed points of this automorphism. □

2.6 The Geometry of Elimination

To study the geometry of elimination, we consider the projection map

$$\pi_k : \mathbb{A}^n(\mathbb{k}) \to \mathbb{A}^{n-k}(\mathbb{k}),\ (a_1, \dots, a_n) \mapsto (a_{k+1}, \dots, a_n).$$

Let $I \subset \mathbb{k}[x_1, \dots, x_n]$ be an ideal, let

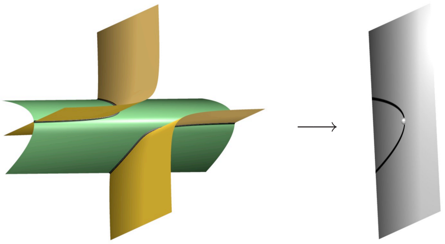
$$I_k = I \cap \mathbb{k}[x_{k+1}, \dots, x_n]$$

be its kth elimination ideal, and let $A = \mathrm{V}(I) \subset \mathbb{A}^n(\mathbb{k})$. Then

$$\pi_k(A) \subset \mathrm{V}(I_k) \subset \mathbb{A}^{n-k}(\mathbb{k}). \tag{2.6}$$

Indeed, every element $f \in I_k \subset I$ vanishes on A and, thus, on $\pi_k(A)$. Note that $\pi_k(A)$ may well be *strictly* contained in $\mathrm{V}(I_k)$. In fact, even over an algebraically closed field, the image $\pi_k(A)$ needs not be Zariski closed:

Exercise 2.6.1. Let $\mathbb{k}$ be any field, and let $I = \langle xy - 1, y^2 - z\rangle \subset \mathbb{k}[x, y, z]$. We project $A = \mathrm{V}(I) \subset \mathbb{A}^3(\mathbb{k})$ to the yz-plane: Apply Algorithm 2.5.4 to show that $I_1 = \langle y^2 - z\rangle \subset \mathbb{k}[y, z]$ is the first elimination ideal of I. Then note that the origin $o = (0, 0)$ is a point of $\mathrm{V}(I_1) \subset \mathbb{A}^2(\mathbb{k})$ which has no preimage in A.



□

In Chapter 6, it will turn out that missing preimage points may be realized as some sort of "points at infinity". In fact, the idea of adding points at infinity will lead us to the introduction of projective algebraic sets, and we will see in Theorem 6.3.26 that the image of a projective algebraic set under a morphism is always Zariski closed – provided we work over an algebraically closed field. In the affine case, we have the following result:

Theorem 2.6.2. *With notation as at the beginning of this section, suppose that $\Bbbk = \overline{\Bbbk}$ is algebraically closed. Then*

$$\overline{\pi_k(A)} = \mathrm{V}(I_k) \subset \mathbb{A}^{n-k}(\Bbbk).$$

That is, $\mathrm{V}(I_k)$ *is the smallest algebraic subset of* $\mathbb{A}^{n-k}(\Bbbk)$ *containing* $\pi_k(A)$*.*

Proof. From (2.6), we have $\pi_k(A) \subset \mathrm{V}(I_k)$, so that also $\overline{\pi_k(A)} \subset \mathrm{V}(I_k)$.

For the opposite inclusion, let $f \in \Bbbk[x_{k+1}, \dots, x_n] \subset \Bbbk[x_1, \dots, x_n]$ be a polynomial vanishing on $\pi_k(A)$ and, thus, on A. Then, by the Nullstellensatz, $f^m \in I \cap \Bbbk[x_{k+1}, \dots, x_n] = I_k$ for some $m \geq 1$. It follows that $\mathrm{I}(\pi_k(A)) \subset \operatorname{rad} I_k$, so that

$$\overline{\pi_k(A)} = \mathrm{V}(\mathrm{I}(\pi_k(A))) \supset \mathrm{V}(\operatorname{rad} I_k) = \mathrm{V}(I_k). \qquad \square$$

The theorem implies the following more general result:

Corollary 2.6.3. *Let $I \subset \Bbbk[\boldsymbol{x}] = \Bbbk[x_1, \dots, x_n]$ be an ideal, let $A = \mathrm{V}(I) \subset \mathbb{A}^n(\Bbbk)$, and let*

$$\varphi : A \to \mathbb{A}^m(\Bbbk),\ p \mapsto (f_1(p), \dots, f_m(p)),$$

be a morphism, given by polynomials $f_1, \dots, f_m \in \Bbbk[\boldsymbol{x}]$. Let J be the ideal

$$J = I\,\Bbbk[\boldsymbol{x}, \boldsymbol{y}] + \langle f_1 - y_1, \dots, f_m - y_m \rangle \subset \Bbbk[\boldsymbol{x}, \boldsymbol{y}],$$

where $\boldsymbol{y}$ stands for the coordinate functions $y_1, \dots, y_m$ on $\mathbb{A}^m(\Bbbk)$. If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, then

$$\overline{\varphi(A)} = \mathrm{V}(J \cap \Bbbk[\boldsymbol{y}]) \subset \mathbb{A}^m(\Bbbk).$$

Proof. If $\Bbbk$ is any field, the locus of zeros of J in $\mathbb{A}^{n+m}(\Bbbk)$ is the graph of φ:

$$\mathrm{V}(J) = \{(p, \varphi(p)) \mid p \in A\} \subset \mathbb{A}^{n+m}(\Bbbk).$$

Thus, if $\pi : \mathbb{A}^{n+m}(\Bbbk) \to \mathbb{A}^m(\Bbbk)$, $(p, q) \mapsto q$, is projection onto the $\boldsymbol{y}$-components, then

$$\varphi(A) = \pi(\mathrm{V}(J)) \subset \mathrm{V}(J \cap \Bbbk[\boldsymbol{y}]) \subset \mathbb{A}^m(\Bbbk).$$

If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, Theorem 2.6.2 implies that

$$\overline{\varphi(A)} = \overline{\pi(\mathrm{V}(J))} = \mathrm{V}(J \cap K[\boldsymbol{y}]). \qquad \square$$

If $\Bbbk$ is not algebraically closed, the conclusions of Theorem 2.6.2 and Corollary 2.6.3 may fail (for instance, consider $\mathrm{V}(x^2 + 1, y) \subset \mathbb{A}^2(\mathbb{R})$ and project to the y-axis). They hold, however, under an additional hypothesis:

Corollary 2.6.4. *Let I, A, φ, and J be as in the preceeding corollary. If $\Bbbk$ is not algebraically closed, suppose that A is Zariski dense in the locus of zeros of I in $\mathbb{A}^n(\overline{\Bbbk})$. Then*

$$\overline{\varphi(A)} = \mathrm{V}(J \cap \Bbbk[\boldsymbol{y}]) \subset \mathbb{A}^m(\Bbbk).$$

Proof. We write $\overline{\mathrm{V}}$ for taking loci of zeros over $\overline{\Bbbk}$ and $\overline{\varphi}$ for the morphism $\overline{\mathrm{V}}(I) \to \mathbb{A}^m(\overline{\Bbbk})$ given by $f_1, \dots, f_m$.

From the proof of the preceeding corollary, we already know that

$$\varphi(A) \subset \mathrm{V}(J \cap K[\boldsymbol{y}]) \subset \mathbb{A}^m(\Bbbk).$$

To show that $\mathrm{V}(J \cap K[\boldsymbol{y}])$ is the smallest algebraic subset of $\mathbb{A}^m(\Bbbk)$ containing $\varphi(A)$, let $g \in \Bbbk[y_1, \dots, y_m]$ be any polynomial vanishing on $\varphi(A)$. Then the polynomial $g(f_1, \dots, f_m) \in \Bbbk[x_1, \dots, x_n]$ vanishes on A and, thus, on $\overline{\mathrm{V}}(I) \subset \mathbb{A}^n(\overline{\Bbbk})$ since A is Zariski dense in $\overline{\mathrm{V}}(I)$. So g vanishes on $\overline{\varphi}(\overline{\mathrm{V}}(I)) \subset \mathbb{A}^m(\overline{\Bbbk})$ and, thus, on $\overline{\mathrm{V}}(J \cap \Bbbk[\boldsymbol{y}])$ by the preceeding corollary. In particular, g vanishes on $\mathrm{V}(J \cap \Bbbk[\boldsymbol{y}])$. We conclude that every algebraic subset of $\mathbb{A}^m(\Bbbk)$ containing $\varphi(A)$ must contain $\mathrm{V}(J \cap \Bbbk[\boldsymbol{y}])$ as well. □

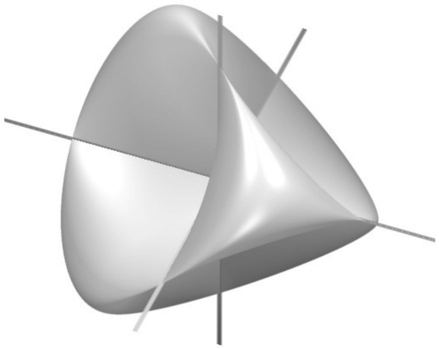
Remark 2.6.5. If $\Bbbk$ is infinite, then $\mathbb{A}^n(\Bbbk)$ is Zariski dense in $\mathbb{A}^n(\overline{\Bbbk})$. Indeed, the same argument as in Exercise 1.2.1 shows that if $f \in \overline{\Bbbk}[x_1, \dots, x_n]$ is a polynomial vanishing on $\mathbb{A}^n(\Bbbk)$, then f is zero. □

Exercise 2.6.6 (Steiner Roman Surface). Consider the real 2-sphere

$$S^2 = \mathrm{V}(x_1^2 + x_2^2 + x_3^2 - 1) \subset \mathbb{A}^3(\mathbb{R})$$

and the morphism

$$\varphi : S^2 \to \mathbb{A}^3(\mathbb{R}),\ (a_1, a_2, a_3) \mapsto (a_1a_2, a_1a_3, a_2a_3).$$



Show that the hypersurface defined by the polynomial

$$f = y_1^2y_2^2 + y_1^2y_3^2 + y_2^2y_3^2 - y_1y_2y_3$$

is the smallest algebraic subset of $\mathbb{A}^3(\mathbb{R})$ containing $\varphi(S^2)$. Show that $\varphi(S^2)$ is not Zariski closed. Precisely, what zeros of f are not contained in $\varphi(S^2)$?

In the analogous situation over $\mathbb{C}$, show that φ is onto. □

Definition 2.6.7. Let $B \subset \mathbb{A}^m(\Bbbk)$ be algebraic. A **polynomial parametrization** of B is a morphism

$$\varphi : \mathbb{A}^n(\Bbbk) \to \mathbb{A}^m(\Bbbk) \text{ such that } \overline{\varphi(\mathbb{A}^n(\Bbbk))} = B.$$

□

Example 2.6.8. Let $\Bbbk$ be infinite. As we already know, the map

$$\mathbb{A}^1(\Bbbk) \to C,\ a \mapsto (a, a^2, a^3),$$

is a polynomial parametrization of the twisted cubic curve C (in fact, it is an isomorphism onto C). This fits well with the fact that the polynomials $y^2 - xz$, $xy - z$, $x^2 - y$, $t - x$ form a Gröbner basis for the ideal

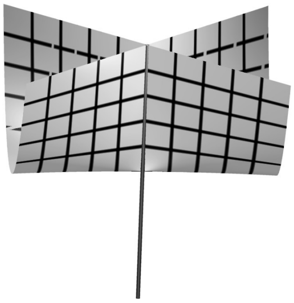
$$J = \langle x - t,\ y - t^2,\ z - t^3 \rangle \subset \Bbbk[t, x, y, z]$$

with respect to the product order $(>_1, >_2)$, where $>_2$ is the degree reverse lexicographic order on $\Bbbk[x, y, z]$ (and $>_1$ is the unique global monomial order on $\Bbbk[t]$). □

Exercise 2.6.9 (Whitney Umbrella). Show that the map

$$\varphi : \mathbb{A}^2(\mathbb{R}) \to \mathbb{A}^3(\mathbb{R}),\ (a, b) \mapsto (ab, b, a^2),$$

is a polynomial parametrization of the Whitney umbrella $\mathrm{V}(x^2 - y^2 z)$ which is not onto. Exactly, what points do not have a preimage?



In the analogous situation over $\mathbb{C}$, show that φ is onto. □

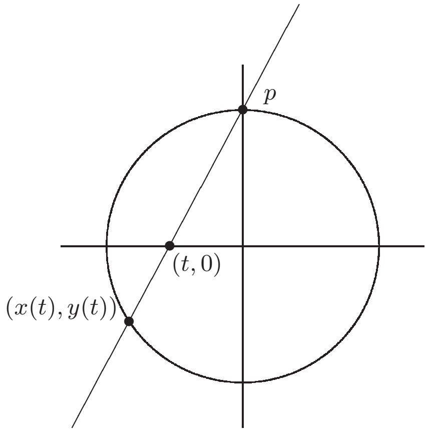
Exercise 2.6.10. If $\operatorname{char} \Bbbk \neq 2$, show that the circle

$$C = \mathrm{V}(x^2 + y^2 - 1) \subset \mathbb{A}^2(\Bbbk)$$

does not admit a polynomial parametrization. □

There is, however, a parametrization of the circle given by rational functions:

Example 2.6.11 (Stereographic Projection). If $\operatorname{char} \Bbbk \neq 2$, we construct a rational parametrization of the circle $C = \mathrm{V}(x^2+y^2-1) \subset \mathbb{A}^2(\Bbbk)$ by means of projecting C onto the x-axis, with $p=(0,1)$ as the projection center:



If $t \neq 0$, the line L through p and the point $(t,0)$ on the x-axis is given by the equation $y = -\frac{1}{t}x+1$. It intersects C in p and one further point $(x(t),y(t)) \in C$. The coordinate $x(t)$ is obtained as the nonzero solution of the equation $x^2+(-\frac{1}{t}x+1)^2-1 = x(\frac{t^2+1}{t^2}x-\frac{2}{t})=0$. Thus, the circle admits the rational parametrization

$$(x(t),y(t)) = (\frac{2t}{t^2+1}, \frac{t^2-1}{t^2+1}),\ t \in \mathbb{A}^1(\Bbbk) \setminus \mathrm{V}(t^2+1).$$

Observe that the line through p and the origin is the y-axis. It intersects C in p and the point $(0,-1)$ which is also in the image of the parametrization: $(x(0),y(0)) = (0,-1)$. The point p itself has no preimage (again, we would have to add some sort of "point at infinity" corresponding to the tangent line to C at p). □

In defining rational functions and rational parametrizations formally, we make use of the construction of the quotient field which we briefly recall, now (a more general construction will be considered in Section 4.2):

Remark-Definition 2.6.12. If R is an integral domain, the relation on $R \times (R \setminus \{0\})$ defined by

$$(r,u) \sim (r',u') \iff ru'-ur'=0$$

is an equivalence relation. We think of the equivalence class of $(r,u) \in R \times (R \setminus \{0\})$ as a fraction, and denote it by r/u. The set $\mathrm{Q}(R)$ of all equivalence classes becomes a field, with algebraic operations

$$r/u + r'/u' = (u'r+ur')/uu' \ \text{ and } \ r/u \cdot r'/u' = (rr')/(uu').$$

We consider R as a subring of $\mathrm{Q}(R)$ by means of the natural ring monomorphism

$$R \to \mathrm{Q}(R),\ r \mapsto r/1,$$

and call $\mathrm{Q}(R)$ the **quotient field** of R. □

Applying this construction to the polynomial ring $\Bbbk[x_1,\dots,x_n]$, we get the **field** $\Bbbk(x_1,\dots,x_n)$ **of rational functions** in $x_1,\dots,x_n$ with coeffients in $\Bbbk$. Applying it to the coordinate ring of an affine variety V, which is an integral domain by Proposition 1.7.2, we get the rational function field of V:

Definition 2.6.13. Let $V \subset \mathbb{A}^n(\Bbbk)$ be a variety. The **rational function field** of V, denoted $\Bbbk(V)$, is defined to be

$$\Bbbk(V) = \mathrm{Q}(\Bbbk[V]).$$

A **rational function** on V is an element $f \in \Bbbk(V)$. □

According to the definition, a rational function on V is a fraction $f = g/h$ of two polynomial functions $g, h \in \Bbbk[V]$, where $h \neq 0$. Viewing f itself as a function, however, has to be done with some care since the denominator h may have zeros.

Definition 2.6.14. Let $V \subset \mathbb{A}^n(\Bbbk)$ be a variety. A rational function f on V is **defined** at a point $p \in V$ (or **regular** at p) if there is a representation $f = g/h$ such that $g, h \in \Bbbk[V]$ and $h(p) \neq 0$. The set

$$\mathrm{dom}(f) := \{p \in V \mid f \text{ is defined at } p\}$$

is called the **domain of definition** of f. □

Proposition 2.6.15. *Let $V \subset \mathbb{A}^n(\Bbbk)$ be a variety, and let $f \in \Bbbk(V)$. Then:*

1. *The domain $\mathrm{dom}(f)$ is open and dense in the Zariski topology on V.*
2. *If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, then*

$$\mathrm{dom}(f) = V \iff f \in \Bbbk[V].$$

In other words, a rational function $f \in \Bbbk(V)$ is regular everywhere on V iff f is a polynomial function on V.

Proof. Considering the *ideal I_f of denominators* of f,

$$\begin{aligned} I_f &= \{h \in \Bbbk[V] \mid fh \in \Bbbk[V]\} \\ &= \{h \in \Bbbk[V] \mid \text{there is an expression } f = g/h \text{ with } g \in \Bbbk[V]\} \cup \{0\}, \end{aligned}$$

we find that

$$V \setminus \mathrm{dom}(f) = \mathrm{V}_V(I_f)$$

is an algebraic subset of V. Hence, $\mathrm{dom}(f)$ is Zariski open and, being nonempty, dense in the Zariski topology on V (see Proposition 1.11.8). Furthermore, if $\Bbbk = \overline{\Bbbk}$, then

$$\mathrm{dom}(f) = V \iff \mathrm{V}_V(\mathrm{I}_f) = \emptyset \iff 1 \in \mathrm{I}_f \iff f \in \Bbbk[V]$$

by the Nullstellensatz in $\Bbbk[V]$ (see Exercise 1.11.7). □

If $p \in \operatorname{dom}(f)$, the value $f(p) := g(p)/h(p) \in \Bbbk$ does not depend on the choice of representation $f = g/h$ with $h(p) \neq 0$. We, hence, have a well-defined map

$$f : \operatorname{dom}(f) \to \mathbb{A}^1(\Bbbk),\ p \mapsto f(p).$$

That the function f is not necessarily defined everywhere on V is usually indicated by writing

$$f : V \dashrightarrow \mathbb{A}^1(\Bbbk).$$

Remark 2.6.16. If R is a UFD, every element $f \in \mathrm{Q}(R)$ admits a representation $f = g/h$ such that $g, h \in R$ are coprime. In such a representation, g and h are uniquely determined up to common unit factors. □

Exercise 2.6.17. Show that

$$V = \mathrm{V}(x_1x_2 - x_3x_4) \subset \mathbb{A}^4(\Bbbk)$$

is a variety whose coordinate ring $\Bbbk[V]$ is not a UFD. Write $\overline{x}_i$ for the residue class of x_i in $\Bbbk[V]$, and observe that the fractions $\overline{x}_1/\overline{x}_3$ and $\overline{x}_4/\overline{x}_2$ represent the same rational function f on V. Show that there is no representation of f as a fraction g/h such that $h(p) \neq 0$ for *all* $p \in \operatorname{dom}(f)$. □

By definition, polynomial maps are maps whose components are polynomial functions. Similarly, we use rational functions to define rational maps:

Remark-Definition 2.6.18. Let $V \subset \mathbb{A}^n(\Bbbk)$ be a variety.

1. A **rational map**

$$\varphi : V \dashrightarrow \mathbb{A}^m(\Bbbk)$$

is a tuple $(f_1, \dots, f_m)$ of rational functions $f_i \in \Bbbk(V)$. The **domain of definition** of φ, written $\operatorname{dom}(\varphi)$, is the set

$$\operatorname{dom}(\varphi) = \bigcap_{i=1}^{m} \operatorname{dom}(f_i).$$

This set is open and dense in the Zariski topology on V. Furthermore, we have a well-defined map

$$\varphi : \operatorname{dom}(\varphi) \to \mathbb{A}^m(\Bbbk),\ p \mapsto \varphi(p) := (f_1(p), \dots, f_m(p)).$$

If $B \subset \mathbb{A}^m(\Bbbk)$ is any subset, its **preimage** under φ is the set

$$\varphi^{-1}(B) := \{p \in \operatorname{dom}(\varphi) \mid \varphi(p) \in B\}.$$

2. If $W \subset \mathbb{A}^m(\Bbbk)$ is another variety, a **rational map**

$$\varphi : V \dashrightarrow W$$

is a rational map $V \dashrightarrow \mathbb{A}^m(\Bbbk)$ such that $\varphi(\operatorname{dom}(\varphi)) \subset W$. □

Exercise* 2.6.19. Let $I \subset \Bbbk[x_1, \ldots, x_n]$ be a prime ideal, let $V = \mathrm{V}(I) \subset \mathbb{A}^n(\Bbbk)$ be the corresponding variety, and let $\varphi : V \dashrightarrow \mathbb{A}^m(\Bbbk)$ be a rational map given by rational functions $f_i = (g_i + I)/(h_i + I) \in \Bbbk(V)$, where the $g_i, h_i \in \Bbbk[x_1, \ldots, x_n]$. Supposing that V is Zariski dense in the locus of zeros of I in $\mathbb{A}^n(\overline{\Bbbk})$, design an algorithm which computes the Zariski closure of $\varphi(\operatorname{dom}(\varphi)) \subset \mathbb{A}^m(\Bbbk)$.
Hint. Consider the ideal

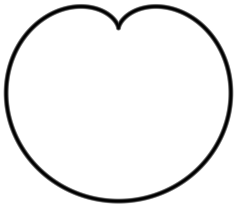
$$J = I\,\Bbbk[\boldsymbol{x}, \boldsymbol{y}] + \langle h_1 y_1 - g_1, \ldots, h_m y_m - g_m, zh - 1\rangle \subset \Bbbk[\boldsymbol{x}, \boldsymbol{y}, z],$$

where z is an extra variable, and $h = h_1 \cdots h_m$. □

Exercise 2.6.20. Consider the rational map $\varphi : \mathbb{A}^1(\mathbb{R}) \dashrightarrow \mathbb{A}^2(\mathbb{R})$ given by

$$x(t) = \frac{-1024t^3}{256t^4 + 32t^2 + 1} \quad \text{and} \quad y(t) = \frac{-2048t^4 + 128t^2}{256t^4 + 32t^2 + 1}.$$

Compute the smallest algebraic subset of $\mathbb{A}^2(\mathbb{R})$ containing $\varphi(\operatorname{dom}(\varphi))$:



□

By Theorem 1.11.13, the composite of two polynomial maps is again a polynomial map. The attempt of formulating an analogous result for rational maps reveals a difficulty which is caused by the fact that rational maps are not really maps: the composite $\psi \circ \varphi$ of two rational maps $\varphi : V \dashrightarrow W$ and $\psi : W \dashrightarrow X$ may not be defined. As a map in the usual sense, $\psi \circ \varphi$ should be defined on $\varphi^{-1}(\operatorname{dom}(\psi)) \cap \operatorname{dom}(\varphi)$. However, this set may well be empty. For instance, consider the morphism $\varphi : \mathbb{A}^1(\Bbbk) \to \mathbb{A}^2(\Bbbk)$, $a \mapsto (a, 0)$, and the rational function $\psi : \mathbb{A}^2(\Bbbk) \dashrightarrow \mathbb{A}^1(\Bbbk)$ given by x/y.

On the algebraic side, arguing as in the proof of Proposition 1.11.12, we obtain a well-defined $\Bbbk$-algebra homomorphism

$$\varphi^* : \Bbbk[W] \to \Bbbk(V).$$

Indeed, let φ be given by a tupel $(f_1, \ldots, f_m)$ of rational functions on V. If $g \in \Bbbk[W]$, then g is a polynomial expression in the coordinate functions $\overline{y}_i$ on W. Substituting the f_i for the $\overline{y}_i$, we get a rational function on V which we take to be the image $\varphi^*(g)$. The attempt of extending φ^* to a $\Bbbk$-algebra homomorphism $\Bbbk(W) \to \Bbbk(V)$ reveals our problem again: we would like to define the image of $g/h \in \Bbbk(W)$ as the fraction $\varphi^*(g)/\varphi^*(h)$; but this is not possible if h is in the kernel of φ^*.

Lemma-Definition 2.6.21. *Let $\varphi : V \dashrightarrow W$ be a rational map between affine varieties. Then the following are equivalent:*

1. *The image $\varphi(\mathrm{dom}(\varphi))$ is Zariski dense in W.*
2. *The $\Bbbk$-algebra homomorphism $\varphi^* : \Bbbk[W] \to \Bbbk(V)$ defined above is injective.*

If these conditions are satisfied, φ is called ***dominant****.*

Proof. If $g \in \Bbbk[W]$, then

$$g \in \ker \varphi^* \iff \varphi(\mathrm{dom}(\varphi)) \subset \mathrm{V}_W(g).$$

That is, φ^* is not injective iff $\varphi(\mathrm{dom}(\varphi))$ is contained in a proper algebraic subset of W. □

Now, we can formulate a result for rational maps which is analogous to Theorem 1.11.13 for polynomial maps:

Theorem 2.6.22. *Let $V \subset \mathbb{A}^n(\Bbbk)$ and $W \subset \mathbb{A}^m(\Bbbk)$ be subvarieties.*

1. *Every dominant rational map $\varphi : V \dashrightarrow W$ induces a $\Bbbk$-algebra homomorphism*
$$\varphi^* : \Bbbk(W) \to \Bbbk(V).$$
2. *Conversely, if $\phi : \Bbbk(W) \to \Bbbk(V)$ is a $\Bbbk$-algebra homomorphism, there exists a unique dominant rational map $\varphi : V \dashrightarrow W$ such that $\phi = \varphi^*$.*
3. *Let $\varphi : V \dashrightarrow W$ be a dominant rational map. If $X \subset \mathbb{A}^r(\Bbbk)$ is any variety, and $\psi : W \dashrightarrow X$ is any rational map, given by a tupel $(g_1, \dots, g_r)$ of rational functions on W, the* ***composition*** *$\psi \circ \varphi : V \dashrightarrow X$ is defined to be the rational map given by the tupel $(\varphi^*(g_1), \dots, \varphi^*(g_r))$. If, in addition, ψ is dominant, then $\psi \circ \varphi$ is dominant, and*
$$(\psi \circ \varphi)^* = \varphi^* \circ \psi^*.$$
□

Exercise* 2.6.23. Prove the theorem. □

Note that if $\psi \circ \varphi$ is defined, then $\mathrm{dom}(\psi \circ \varphi)$ contains $\varphi^{-1}(\mathrm{dom}(\psi))$, but may well be larger (see Exercise 2.6.26 below for examples).

According to our definition in Chapter 1, an isomorphism of algebraic sets is a morphism admitting an inverse morphism. In the same spirit, we define:

Definition 2.6.24. A rational map $\varphi : V \dashrightarrow W$ of affine varieties is called a **birational map** (or a **birational equivalence**) if it is dominant and admits a rational inverse. That is, there is a dominant rational map $\psi : W \dashrightarrow V$ such that $\psi \circ \varphi = \mathrm{id}_V$ and $\varphi \circ \psi = \mathrm{id}_W$. We say that V and W are **birationally equivalent** if there is a birational map $V \dashrightarrow W$. □

Theorem 2.6.22 implies:

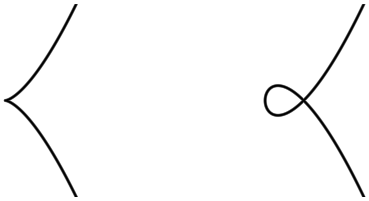
Corollary 2.6.25. *A dominant rational map $\varphi : V \to W$ of affine varieties is birational iff $\varphi^* : \Bbbk(W) \to \Bbbk(V)$ is an isomorphism of $\Bbbk$-algebras. Two affine varieties are birationally equivalent iff their function fields are isomorphic as $\Bbbk$-algebras.* □

Exercise 2.6.26. Consider the polynomial parametrizations

$$\mathbb{A}^1(\Bbbk) \to \mathrm{V}(y^2 - x^3) \subset \mathbb{A}^2(\Bbbk),\ a \mapsto (a^2, a^3),$$

and

$$\mathbb{A}^1(\Bbbk) \to \mathrm{V}(y^2 - x^3 - x^2) \subset \mathbb{A}^2(\Bbbk),\ a \mapsto (a^2 - 1, a^3 - a).$$



Show that each of the parametrizations admits a rational inverse. Use these examples to show that the domain of definition of the composite $\psi \circ \varphi$ of two rational maps may be strictly larger than $\varphi^{-1}(\mathrm{dom}(\psi))$. □

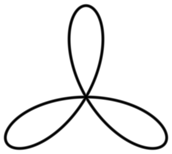
Now, finally, we come to the definition of a rational parametrization:

Definition 2.6.27. Let $W \subset \mathbb{A}^m(\Bbbk)$ be a variety. A **rational parametrization** of W is a dominant rational map

$$\varphi : \mathbb{A}^n(\Bbbk) \dashrightarrow W.$$

□

Exercise 2.6.28. If $\operatorname{char} \Bbbk \neq 2, 3$, find a rational parametrization of the affine plane curve with equation $y^3 - 3x^2y = (x^2 + y^2)^2$:



□

Systematic ways of computing rational parametrizations of curves will be discussed in Theorem 5.4.13, in Section 7.2, and in Chapter 8.

As already mentioned in Remark 1.2.6, most curves do not admit a rational parametrization. Here is a first example:

Example 2.6.29. Suppose that $\operatorname{char} \Bbbk \neq 2, 3$. The affine plane curve with equation

$$y^2 = x^3 + ax + b,$$

where $a, b \in \Bbbk$, has a rational parametrization iff the **discriminant** $D := 4a^2 + 27b^3$ is zero. This will follow from the general theory of curves developed in Chapters 7 and 8 of this book. For an elementary proof based on Fermat's method of infinite descent, see Reid (1988). □

Remark 2.6.30. Suppose that $\Bbbk = \overline{\Bbbk}$ is algebraically closed. In this case, an affine variety is called **unirational** if it admits a rational parametrization. It is called **rational** if it is birationally equivalent to some affine space $\mathbb{A}^d(\Bbbk)$. We will show in Corollary 8.4.10 that every unirational *curve* is rational. This result is also true for *surfaces* (see Barth et al (2004)), but fails to hold in higher dimension (see Iskovskikh and Manin (1971) and Clemens and Griffiths (1972)). □

2.7 The Role of the Ground Field

In the preceeding section, in proving results on the image of a morphism which hold over an arbitrary field, we made use of a strategy which allows one to benefit from Hilbert's Nullstellensatz though this requires that the ground field is algebraically closed. Namely, to study the set of solutions of a system of polynomial equations with coefficients in $\Bbbk$, one first investigates the locus of zeros in $\mathbb{A}^n(\mathbb{K})$, where $\mathbb{K}$ is an algebraically closed extension field of $\Bbbk$. Then, in a second step, one studies the solutions in $\mathbb{A}^n(\Bbbk)$ as a subset of those in $\mathbb{A}^n(\mathbb{K})$. In this book, we are mainly concerned with the first step. The second step involves methods from number theory (if $\Bbbk$ is a number field) and real algebraic geometry (if $\Bbbk = \mathbb{R}$).

On the other hand, to compute examples with exact computer algebra methods, one typically works over a finite field, the field of rational numbers, or a number field. Due to the behavior of Buchberger's algorithm, this fits nicely with the strategy outlined above:

Remark 2.7.1 (Buchberger's Algorithm and Field Extensions). Let $\mathbb{K} \supset \Bbbk$ be a field extension. If $I \subset \Bbbk[x_1, \dots, x_n]$ is an ideal, any Gröbner basis $f_1, \dots, f_r$ for I is also a Gröbner basis for the extended ideal $I\,\mathbb{K}[x_1, \dots, x_n]$. Indeed, all computations in Buchberger's test are carried through over $\Bbbk$.

This shows, in particular, that if a property of ideals can be checked using Gröbner bases, then I has this property iff the extended ideal has this property. To give an example, we know that elimination ideals can be computed using Gröbner bases. It follows that if I_1 is the first elimination ideal of I, then $I_1\,\mathbb{K}[x_2, \dots, x_n]$ is the first elimination ideal of $I\,\mathbb{K}[x_1, \dots, x_n]$. □

For almost every application of Buchberger's algorithm to geometry, the remark allows one to study the vanishing locus of I in $\mathbb{A}^n(\mathbb{K})$ by computations over $\Bbbk$. The exceptions are those discussed in Remark 2.4.13: for radical computations and for primary decomposition, algorithms for square-free decomposition and polynomial factorization are needed in addition to Buchberger's algorithm. These algorithms are sensitive to the ground field. From a theoretical point of view, the behavior of ideals under extensions of the ground field is discussed in Zariski and Samuel (1975–1976), Vol II, Chapter VII, §11. For the interested reader, we summarize some of this discussion, now.

We begin by pointing out that if $\mathfrak{q}$ is a primary ideal of $\Bbbk[x_1,\dots,x_n]$ with radical $\mathfrak{p}$, then the associated primes of $\mathfrak{q}\,\mathbb{K}[x_1,\dots,x_n]$ are precisely the prime ideals of $\mathbb{K}[x_1,\dots,x_n]$ which intersect $\Bbbk[x_1,\dots,x_n]$ in $\mathfrak{p}$ and have the same dimension as $\mathfrak{p}$ (the dimension of ideals will be treated in Chapter 3).

Note, however, that the extension $\mathfrak{p}\,\mathbb{K}[x_1,\dots,x_n]$ of a prime ideal $\mathfrak{p}$ of $\Bbbk[x_1,\dots,x_n]$ cannot always be written as an intersection of prime ideals (that is, the extended ideal may not be a radical ideal). The situation is different if $\mathbb{K} \supset \Bbbk$ is a separable field extension. In particular, if $\Bbbk$ is a perfect field, and $I \subset \Bbbk[x_1,\dots,x_n]$ is any radical ideal, then also $I\,\mathbb{K}[x_1,\dots,x_n]$ is a radical ideal. Recall that finite fields, fields of characteristic zero, and algebraically closed fields are perfect.

If the extended ideal $\mathfrak{p}\,\overline{\Bbbk}[x_1,\dots,x_n]$ of a prime ideal $\mathfrak{p}$ of $\Bbbk[x_1,\dots,x_n]$ is again prime, then $\mathfrak{p}\,\mathbb{K}[x_1,\dots,x_n]$ is prime for any extension field $\mathbb{K}$ of $\Bbbk$. In this case, we say that $\mathfrak{p}$ is **absolutely prime**.

Taking these remarks into account, we will ease our notation further on:

Convention 2.7.2. *From now on,* $\mathbb{K}$ *will denote an algebraically closed extension field of* $\Bbbk$*. We will write* $\mathbb{A}^n := \mathbb{A}^n(\mathbb{K})$*. If* $I \subset \Bbbk[x_1,\dots,x_n]$ *is any subset, then* $A = \mathrm{V}(I)$ *will be its locus of zeros in* $\mathbb{A}^n$*. Furthermore,* $\mathrm{I}(A)$ *will be the vanishing ideal of* A *in* $\mathbb{K}[x_1,\dots,x_n]$*, and* $\mathbb{K}[A] = \mathbb{K}[x_1,\dots,x_n]/\mathrm{I}(A)$ *will be the coordinate ring of* A*.* □

Remark-Definition 2.7.3. 1. With notation as above, we say that $\Bbbk$ is a **field of definition** of A, or that A is **defined over** $\Bbbk$. Moreover, we refer to

$$A(\Bbbk) := A \cap \mathbb{A}^n(\Bbbk)$$

as the set of **$\Bbbk$-rational points** of A.

2. If $\mathfrak{p} \subset \Bbbk[x_1,\dots,x_n]$ is absolutely prime, then $V = \mathrm{V}(\mathfrak{p}) \subset \mathbb{A}^n$ is a variety with rational function field

$$\mathbb{K}(V) = \mathrm{Q}(\mathbb{K}[x_1,\dots,x_n]/\mathfrak{p}\,\mathbb{K}[x_1,\dots,x_n]).$$

Furthermore,

$$\Bbbk(V) := \mathrm{Q}(\Bbbk[x_1,\dots,x_n]/\mathfrak{p})$$

is contained in $\mathbb{K}(V)$ as a subfield. We refer to the elements of $\Bbbk(V)$ as the **rational functions on V defined over** $\Bbbk$.

3. A **rational parametrization defined over** $\Bbbk$ is a rational parametrization

$$\varphi : \mathbb{A}^n \dashrightarrow W \subset \mathbb{A}^m,\ p \mapsto (f_1(p), \dots, f_m(p)),$$

such that $f_1, \dots, f_m$ are defined over $\Bbbk$. □

Parametrizations defined over number fields are useful for answering questions in arithmetic. Here is an example:

Exercise 2.7.4. Find all **Pythagorean tripels**, that is, triples (a, b, c) of integers such that $a^2 + b^2 = c^2$.
Hint. Use the parametrization of the circle given in Example 2.6.11 by means of the stereographic projection. □

Exercise 2.7.5. Let $n \geq 2$, and suppose that $\operatorname{char} \Bbbk \neq 2$. Let $Q \subset \mathbb{A}^n$ be a **nondegenerate quadric**. That is, Q is defined by an equation of type

$$(1, x_1, \dots, x_n) \begin{pmatrix} a_{00} & a_{01} & \dots & a_{0n} \\ a_{10} & a_{11} & & a_{1n} \\ \vdots & & \ddots & \vdots \\ a_{n0} & a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = 0,$$

where (a_{ij}) is a symmetric $(n+1) \times (n+1)$ matrix of scalars $a_{ij} \in \Bbbk$ which has maximal rank $n + 1$. Prove that Q admits a rational parametrization defined over $\Bbbk$ iff $Q(\Bbbk) \neq \emptyset$ (that is, Q has a $\Bbbk$-rational point). □

Exercise 2.7.6. Let $C \subset \mathbb{A}^2$ be an irreducible conic.

1. If C is defined over $\Bbbk$, show that the following are equivalent:
   a) There exists a $\Bbbk$-rational parametrization of C whose components are fractions of polynomials of degree ≤ 2.
   b) There exists a $\Bbbk$-rational point on C.
2. Let $D = \mathrm{V}(f) \subset \mathbb{A}^2$ be another curve. If $C \not\subset D$, show that C and D can have at most $2 \cdot \deg f$ intersection points. □

2.8 Hilbert's Syzygy Theorem

As a final application of Gröbner bases in this Chapter, we give an elementary proof of Hilbert‘s syzygy theorem. Hilbert's own proof (1890) is based on elimination and is, as Hilbert remarked, “nicht ganz ohne Mühe”[1]. The syzygy theorem is the starting point of homological algebra, a mathematical discipline of its own which is crosslinked to many other areas of mathematics (see, for instance, Eisenbud (1995)). We use the following terminology:

Definition 2.8.1. Let R be a ring. A **complex** of R-modules is a finite or infinite sequence of R-modules and homomorphisms of R-modules

[1] not without difficulty

$$\dots \longrightarrow M_{i+1} \xrightarrow{\phi_{i+1}} M_i \xrightarrow{\phi_i} M_{i-1} \longrightarrow \dots$$

such that $\phi_i \circ \phi_{i+1} = 0$ for all i. The **homology** of the complex at M_i is defined to be $\ker \phi_i / \operatorname{im} \phi_{i+1}$. We say that the complex is **exact at $\boldsymbol{M_i}$** if the homology at M_i is zero. It is **exact** if it is exact at every M_i. □

For instance, a finite sequence of type

$$M_r \to M_{r-1} \to \cdots \to M_{s+1} \to M_s$$

is exact iff it is exact at every M_i, $r-1 \leq i \leq s+1$.

Example 2.8.2. Let R be a ring, and let $\phi : M \to N$ be a homomorphism of R-modules. Write 0 for the trivial R-module, and let $0 \to M$ and $N \to 0$ be the zero homomorphisms. Then:

1. ϕ is injective $\iff$ the sequence $0 \to M \to N$ is exact.
2. ϕ is surjective $\iff$ the sequence $M \to N \to 0$ is exact.
3. ϕ is bijective $\iff$ the sequence $0 \to M \to N \to 0$ is exact. □

Example 2.8.3. Let R be a ring. A **short exact sequence** is an exact sequence of R-modules of type

$$0 \longrightarrow M' \xrightarrow{\phi} M \xrightarrow{\psi} M'' \longrightarrow 0.$$

That is, ϕ is injective, ψ is surjective, and $\operatorname{im} \phi = \ker \psi$. For instance, if M is an R-module and $N \subset M$ is a submodule, we have a canonical short exact sequence

$$0 \longrightarrow N \longrightarrow M \longrightarrow M/N \longrightarrow 0.$$

A sequence

$$\dots \longrightarrow M_{i+1} \xrightarrow{\phi_{i+1}} M_i \xrightarrow{\phi_i} M_{i-1} \longrightarrow \dots$$

as in Definition 2.8.1 is exact iff each induced sequence

$$0 \longrightarrow M_{i+1}/\ker \phi_{i+1} \longrightarrow M_i \longrightarrow \operatorname{im} \phi_i \longrightarrow 0$$

is exact. □

Exercise* **2.8.4.** Show that if

$$0 \to M_r \to M_{r-1} \to \dots \to M_s \to 0$$

is an exact sequence of finite dimensional $\Bbbk$-vector spaces, then

$$\sum_{i=r}^{s} (-1)^i \dim_{\Bbbk} M_i = 0.$$

□

Exercise* 2.8.5. Let M be an R-module, and let

$$0 \longrightarrow N' \longrightarrow N \longrightarrow N'' \longrightarrow 0$$

be a short exact sequence. Prove that the induced sequence

$$0 \longrightarrow \mathrm{Hom}_R(M, N') \longrightarrow \mathrm{Hom}_R(M, N) \longrightarrow \mathrm{Hom}_R(M, N'')$$

is exact. Show by example that the map $\mathrm{Hom}_R(M, N) \longrightarrow \mathrm{Hom}_R(M, N'')$ is not necessarily surjective. □

Following Hilbert, we, now, consider exact sequences of a type which allows us to obtain information on arbitrary modules from information on free modules.

It is clear from our discussion on syzygies that every module M over a ring R is the epimorphic image of a free R-module. Indeed, choose generators $\{f_\lambda\}$ of M, a free R-module F_0 on a corresponding basis $\{\epsilon_\lambda\}$, and consider the homomorphism

$$F_0 \xrightarrow{\pi} M,\ \epsilon_\lambda \mapsto f_\lambda.$$

In a next step, applying the same argument to the kernel of π, we get a free R-module F_1 together with an epimorphism $F_1 \to \ker \pi$. If ϕ is the composite map $F_1 \to \ker \pi \to F_0$, then $M = \mathrm{coker}\, \phi$.

Definition 2.8.6. Let M be a module over a ring R. A **free presentation** of M is an exact sequence

$$F_1 \xrightarrow{\phi} F_0 \longrightarrow M \longrightarrow 0,$$

with free R-modules F_0, F_1. Given such a presentation, we also say that M **is given by generators and relations**. Moreover, if F_0 and F_1 are finitely generated, we often regard ϕ as a matrix, and call it a **presentation matrix** of M. □

Further repetitions in the process discussed before the definition yield a (possibly infinite) exact sequence

$$\ldots \longrightarrow F_{i+1} \xrightarrow{\phi_{i+1}} F_i \xrightarrow{\phi_i} F_{i-1} \longrightarrow \ldots \longrightarrow F_1 \xrightarrow{\phi_1} F_0 \longrightarrow M \longrightarrow 0,$$

with free R-modules F_i (and $\phi_1 = \phi$).

Definition 2.8.7. By abuse of notation, we refer to every sequence as above, as well as to its "free part"

$$\ldots \longrightarrow F_{i+1} \xrightarrow{\phi_{i+1}} F_i \xrightarrow{\phi_i} F_{i-1} \longrightarrow \ldots \longrightarrow F_1 \xrightarrow{\phi_1} F_0,$$

as a **free resolution** of M. We call $\mathrm{im}\, \phi_i$ an ***i*th syzygy module** and its elements ***i*th order syzygies** of M. We say that the resolution is **finite** if there is an integer c such that $F_i = 0$ for $i \geq c+1$. In this case, the least such c is the **length** of the resolution. □

Note that the syzygy modules depend on the choices made when constructing the resolution. It follows from the construction and Exercise 1.10.9 that every finitely generated module M over a Noetherian ring R admits a free resolution by finitely generated free R-modules. If we, then, think of φ_i as a matrix, we call it an ***i*th syzygy matrix** of M.

Example 2.8.8. If R is a PID, and M is a finitely generated R-module, the structure theorem for such modules tells us that M is of type

$$M \cong R^s \oplus R/\langle d_1 \rangle \oplus \cdots \oplus R/\langle d_t \rangle,$$

where $s \geq 0$ is a uniquely determined integer, and the d_i are nonzero nonunits in R such that d_i divides d_{i+1}, for $i = 1, \ldots, t-1$ (see, for instance Dummit and Foote (2003), Section 12.1). The d_i, which are uniquely determined up to unit multiples, are known as the **elementary divisors** of M. In terms of syzygies, the structure theorem means that M has a free resolution of type

$$0 \longrightarrow F_1 \xrightarrow{\phi} F_0 \longrightarrow M \longrightarrow 0,$$

where

$$\phi = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \ldots & d_t \\ 0 & & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & & \ldots & 0 \end{pmatrix}.$$

□

Definition 2.8.9. If $d_1, \ldots, d_r$ are elements of a PID such that d_i divides d_{i+1}, for $i = 1, \ldots, t-1$, then a diagonal matrix of block form

$$\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \text{ where } D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_t \end{pmatrix},$$

is said to be in **Smith normal form**. □

Exercise 2.8.10. 1. In the case where $R = \Bbbk[x]$, prove the structure theorem by showing that every matrix A with entries in R can be put into Smith normal form. More precisely, design an algorithm which, given A, finds invertible matrices P and Q with entries in R such that PAQ is in Smith normal form (use Euclid's extended GCD algorithm in conjunction with elementary row and column operations).

2. Applying the algorithm from part 1, show that

$$\phi = \begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2x & 3 \\ x^2-x & 0 & 3x^2-3x & 2x \\ 0 & x^2-x & 0 & 3x^2-3x \end{pmatrix} \underset{\Bbbk[x]}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x^2(x-1) \end{pmatrix},$$

where $\underset{\Bbbk[x]}{\sim}$ refers to taking the Smith normal form over $\Bbbk[x]$. Conclude that $M := \operatorname{coker}\phi$ has the free resolution

$$0 \longrightarrow R^2 \xrightarrow{\begin{pmatrix} x & 0 \\ 0 & x^2(x-1) \end{pmatrix}} R^2 \longrightarrow 0\,.$$

From this, conclude that $M \cong R/\langle x\rangle \oplus R/\langle x^2\rangle \oplus R/\langle x-1\rangle$. □

The structure theorem for PID's says, in particular, that every finitely generated module over the polynomial ring $\Bbbk[x]$ in one variable has a free resolution of length one, by finitely generated free R-modules. Hilbert's syzygy theorem treats the case of several variables:

Theorem 2.8.11 (Hilbert's Syzygy Theorem). *If $R = \Bbbk[x_1,\dots,x_n]$, every finitely generated R-module M has a finite free resolution of length at most n, by finitely generated free R-modules.*

Proof. We give a constructive proof. If M is free, there is nothing to do. So suppose the contrary, and let M be given by generators and relations:

$$R^r \xrightarrow{\phi} R^{s_0} \longrightarrow M \longrightarrow 0$$

Regard ϕ as a matrix and, thus, its columns as a set of generators for $\operatorname{im}\phi$. Starting from these generators, compute a minimal Gröbner basis $f_1,\dots,f_{s_1}$ for $\operatorname{im}\phi$ with respect to some global monomial order on R^{s_0}. Consider the syzygies $G^{(i,\alpha)}$ obtained by applying Buchberger's test to $f_1,\dots,f_{s_1}$. With respect to the induced order on R^{s_1}, the $G^{(i,\alpha)}$ form a minimal Gröbner basis for the kernel of the composite map $\phi_1 : R^{s_1} \to \operatorname{im}\phi \to R^{s_0}$ which sends the ith canonical basis vector ϵ_i of R^{s_1} to f_i.

Computing, now, the syzygies on the $G^{(i,\alpha)}$ and so forth, we successively get minimal Gröbner bases which generate syzygy modules of M of higher order. At each stage, the new Gröbner basis depends, in particular, on how we arrange the elements of the Gröbner basis computed in the previous step. We show that if this arrangement is done properly, then the process just described will terminate after finitely many steps.

To begin with, fix an integer $1 \le k \le n$ such that none of the leading terms $\mathbf{L}(f_i)$ involves the variables $x_{k+1},\dots,x_n$ (choose $k = n$ if one of the $\mathbf{L}(f_i)$ involves x_n). In Buchberger's test, let the f_i be arranged such that, for $i > j$, the exponent of x_k in $\mathbf{L}(f_j)$ is strictly smaller than that of x_k in $\mathbf{L}(f_i)$ whenever these leading terms involve the same basis element of R^{s_0}. Then none of the resulting leading terms $\mathbf{L}(G^{(i,\alpha)}) = c^{(i,\alpha)}x^\alpha\epsilon_i$ involves $x_k,\dots,x_n$.

Arranging the Gröbner basis elements at each stage of our process accordingly, we obtain after, say, $\ell \leq k$ steps an exact sequence

$$R^{s_\ell} \xrightarrow{\phi_\ell} R^{s_{\ell-1}} \longrightarrow \ldots \longrightarrow R^{s_1} \xrightarrow{\phi_1} R^{s_0} \longrightarrow M \longrightarrow 0$$

together with a Gröbner basis $\mathcal{G}$ for $\ker \phi_\ell$ such that none of the leading terms $\mathbf{L}(g)$, $g \in \mathcal{G}$, involves $x_1, \ldots, x_n$. Having chosen a minimal Gröbner basis in the previous step, this implies that $\mathcal{G} = \{0\}$. Thus, $\ker \phi_\ell = 0$ and

$$0 \longrightarrow R^{s_\ell} \xrightarrow{\phi_\ell} R^{s_{\ell-1}} \longrightarrow \ldots \longrightarrow R^{s_1} \xrightarrow{\phi_1} R^{s_0} \longrightarrow M \longrightarrow 0$$

is a finite free resolution as desired. □

Example 2.8.12. Our computations in Exercise 2.3.21 show that the affine ring R/I, where $R = \Bbbk[x_1, \ldots, z_5]$ and I is generated by the 3×3 minors of the matrix

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ y_1 & y_2 & y_3 & y_4 & y_5 \\ z_1 & z_2 & z_3 & z_4 & z_5 \end{pmatrix},$$

has a free resolution of type

$$0 \longrightarrow R^6 \longrightarrow R^{15} \longrightarrow R^{10} \longrightarrow R \longrightarrow R/I \longrightarrow 0.$$

□

Example 2.8.13. Consider the ideal

$$I = \langle f_1, \ldots, f_5 \rangle \subset R = \Bbbk[w, x, y, z]$$

generated by the polynomials

$$f_1 = w^2 - xz,\ f_2 = wx - yz,\ f_3 = x^2 - wy,\ f_4 = xy - z^2,\ f_5 = y^2 - wz.$$

We compute a finite free resolution of $M = R/I$, starting with the degree reverse lexicographic order on $\Bbbk[w, x, y, z]$. We successively obtain three syzygy matrices ϕ_1, ϕ_2, and ϕ_3 which we present in a compact way as follows:

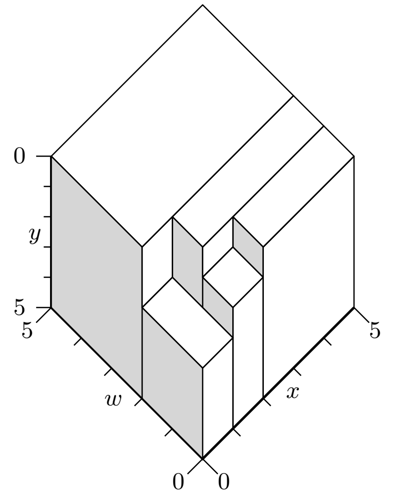
$$\begin{array}{r|cccccc}
\mathbf{w^2} - xz & -x & y & 0 & -z & 0 & -y^2 + wz \\
\mathbf{wx} - yz & \mathbf{w} & -x & -y & 0 & z & z^2 \\
\mathbf{x^2} - wy & -z & \mathbf{w} & 0 & -y & 0 & 0 \\
\mathbf{xy} - z^2 & 0 & 0 & \mathbf{w} & \mathbf{x} & -y & yz \\
\mathbf{y^2} - wz & 0 & 0 & -z & -w & \mathbf{x} & \mathbf{w^2} \\
\hline
 & 0 & y & -x & \mathbf{w} & -z & 1 \\
 & -y^2 + wz & z^2 & -wy & yz & -w^2 & \mathbf{x}
\end{array}$$

All initial terms are printed in bold. The first column of our table is the transposed of the matrix ϕ_1. It contains the original generators for I which, as Buchberger's test shows, form already a Gröbner basis for I. The syzygy matrix ϕ_2 resulting from the test is the 5×6 matrix in the middle of our table. Note that, for instance, $M_4 = \langle w, x\rangle$ can be read from the 4th row of ϕ_2. At this point, we already know that the columns of ϕ_2 form a Gröbner basis for Syz $(f_1, \dots, f_5)$ with respect to the induced monomial order on R^5. Buchberger's test applied to these Gröbner basis elements yields a 6×2 syzygy matrix ϕ_3 whose transposed is printed in the two bottom rows of our table. The map defined by ϕ_3 is injective since the initial terms involve different basis vectors. Thus, we obtain a free resolution of type

$$0 \longrightarrow R^2 \xrightarrow{\phi_3} R^6 \xrightarrow{\phi_2} R^5 \xrightarrow{\phi_1} R \longrightarrow R/I \longrightarrow 0.$$

Observe that once we have the initial terms of the Gröbner basis for I, we can easily compute the initial terms of the Gröbner bases for all syzygy modules, that is, all bold face entries of our table. This gives us an an early idea on the amount of computation lying ahead.

We visualize the monomials in $\mathbf{L}(I)$:



Note that the ranks of the free modules in the resolution are visible in this picture. □

Exercise 2.8.14. Compute a finite free resolution of the ideal generated by the 2×2 minors of the matrix

$$\begin{pmatrix} x_0 & x_1 & x_2 & x_3 \\ x_1 & x_2 & x_3 & x_4 \end{pmatrix}.$$

□

Remark 2.8.15. 1. If R is an arbitrary Noetherian ring, it is not necessarily true that every finitely generated R-module has a finite free resolution. For instance, if $R = \Bbbk[x, y]/\langle xy\rangle$, the ideal generated by the residue class $\overline{x} = x + \langle xy\rangle$ has the infinite periodic resolution

$$\ldots \longrightarrow R \xrightarrow{\overline{y}} R \xrightarrow{\overline{x}} R \xrightarrow{\overline{y}} R \longrightarrow \langle\overline{x}\rangle \longrightarrow 0.$$

By a result of Auslander-Buchsbaum and Serre (see, for instance, Eisenbud (1995), Theorem 19.12), the following conditions on a *local* Noetherian ring R are equivalent:

a) There exists a number s such that every finitely generated R-module has a finite free resolution of length at most s.

b) R is regular.

We will introduce regular rings in Chapter 4.

2. Hilbert's original application of the syzygy theorem, the proof of the polynomial nature of the Hilbert function, will be treated in Section 6.4.
3. Some references for further reading on free resolutions are Serre (1965), Eisenbud (1995), and Avramov (1989). □

Exercise 2.8.16 (Syzygies Over Affine Rings). Design an algorithm which computes syzygy modules over an affine ring $\Bbbk[x_1, \ldots, x_n]/I$ using Gröbner bases in $\Bbbk[x_1, \ldots, x_n]$. □

Exercise 2.8.17. Let $0 \leftarrow M \leftarrow F_0 \leftarrow F_1 \leftarrow \ldots$ be a free complex, i.e the F_i are free modules, and let $0 \leftarrow N \leftarrow G_0 \leftarrow G_1 \leftarrow \ldots$ be an exact complex. Then every morphism $\psi \in \operatorname{Hom}(M, N)$ extends to a morphism ψ_* of complexes, i.e. there is a commutative diagram

$$\begin{tikzcd}
0 & M \arrow[l] \arrow[d, "\psi"'] & F_0 \arrow[l] \arrow[d, "\psi_0"'] & F_1 \arrow[l] \arrow[d, "\psi_1"'] & F_2 \arrow[l] \arrow[d, "\psi_2"'] & \cdots \arrow[l] \\
0 & N \arrow[l] & G_0 \arrow[l] & G_1 \arrow[l] & G_2 \arrow[l] & \cdots \arrow[l]
\end{tikzcd}$$

The morphisms ψ_i are uniquely determined up to a homotopy, which means, that given another extension ψ'_*, there exist maps $h_i : F_i \to G_i + 1$, such that

$$\psi_i - \psi'_i = h_{i-1} \circ \varphi_i^F + \varphi_{i+1}^G \circ h_i .$$

Conclude, that in any extension ψ_* of id_M between two minimal free resolutions each ψ_i is an isomorphism. □

Appendix: Computing Ext and Tor

For the benefit of those readers, which are already familiar with the usage of the functors Ext and Tor of Cartan-Eilenberg [1956], we explain how to compute these modules over affine rings $R = \Bbbk[x_1, \ldots, x_n]/I$ with computer algebra.

The main purpose of the functors

$$\operatorname{Ext}^i_R(M, -) : N \mapsto \operatorname{Ext}^i_R(M, N),$$

is to measure the failure of the left exactness of the functor $\operatorname{Hom}_R(M, -)$ in the following sense: Given a short exact sequence

$$0 \to N' \to N \to N'' \to 0,$$

of R-modules, there is a long exact sequence

$$\begin{tikzcd}
0 \arrow[r] & \operatorname{Hom}_R(M, N') \arrow[r] & \operatorname{Hom}_R(M, N) \arrow[r] & \operatorname{Hom}_R(M, N'') \arrow[dll] \\
& \operatorname{Ext}^1_R(M, N') \arrow[r] & \operatorname{Ext}^1_R(M, N) \arrow[r] & \operatorname{Ext}^1_R(M, N'') \arrow[dll] \\
& \operatorname{Ext}^2_R(M, N') \arrow[r] & \cdots &
\end{tikzcd}$$

Thus $\operatorname{Ext}^1_R(M, N') = 0$ is a sufficient condition for the exactness of the short sequence

$$\operatorname{Hom}_R(M, 0 \to N' \to N \to N'' \to 0).$$

Similarly there are long exact sequences

$$\begin{tikzcd}
0 \arrow[r] & \operatorname{Hom}_R(N'', M) \arrow[r] & \operatorname{Hom}_R(N, M) \arrow[r] & \operatorname{Hom}_R(N', M) \arrow[dll] \\
& \operatorname{Ext}^1_R(N'', M) \arrow[r] & \operatorname{Ext}^1_R(N, M) \arrow[r] & \operatorname{Ext}^1_R(N', M) \arrow[dll] \\
& \operatorname{Ext}^2_R(N'', M) \arrow[r] & \cdots &
\end{tikzcd}$$

and

$$\begin{tikzcd}
M \otimes_R N' \arrow[r] & M \otimes_R N \arrow[r] & M \otimes_R N'' \arrow[r] & 0 \\
\operatorname{Tor}^R_1(M, N') \arrow[r] & \operatorname{Tor}^R_1(M, N) \arrow[r] & \operatorname{Tor}^R_1(M, N'') \arrow[ull] & \\
& \cdots \arrow[r] & \operatorname{Tor}^R_2(M, N'') \arrow[ull] &
\end{tikzcd}$$

which measure the exactness of

$$\mathrm{Hom}_R(0 \to N' \to N \to N'' \to 0, M)$$

and

$$M \otimes_R (0 \to N' \to N \to N'' \to 0)$$

respectively.

To compute $\mathrm{Ext}^i_R(M,N)$ one can use can either an injective resolution

$$0 \to N \to I^0 \to I^1 \to \dots$$

of N, or an projective resolution

$$\dots \to F_1 \to F_0 \to M \to 0$$

of M and apply the formula

$$\mathrm{Ext}^i_R(M,N) \cong \mathrm{H}^i \,\mathrm{Hom}(M, I^*) \cong \mathrm{H}^i \,\mathrm{Hom}(F_*, N).$$

Similarly,

$$\mathrm{Tor}^R_i(M,N) \cong \mathrm{H}_i(F_* \otimes N).$$

Since injective modules are rarely finitely presented we work with a projective or even simpler free resolution. We proceed in several steps.

Remark 2.8.18 (Presentation of homomorphism). Any morphism $\varphi \in \mathrm{Hom}_R(M,N)$ can be represented by a commutative diagram involving the free presentations E_* and F_* of M and N respectively.

$$\begin{tikzcd}
E_1 \arrow[r, "a_1"] \arrow[d, "\varphi_1"] & E_0 \arrow[r] \arrow[d, "\varphi_0"] & M \arrow[r] \arrow[d, "\varphi"] & 0 \\
F_1 \arrow[r, "b_1"] & F_0 \arrow[r] & N \arrow[r] & 0
\end{tikzcd}$$

Conversely any φ_0, which can be completed with some φ_1 to a commutative diagram, represents an homomorhism φ. An φ_0 represents the zero homomorphism, if it factors over F_1, that is $\varphi_0 = b_1 h$ for a map $h : E_0 \to F_1$.

Algorithm 2.8.19 (Hom). . **Input:** *Two R-modules specified via free presentations*

$$\begin{tikzcd}
E_1 \arrow[r, "a_1"] & E_0 \arrow[r] & M \arrow[r] & 0
\end{tikzcd}$$

and

$$\begin{tikzcd}
F_1 \arrow[r, "b_1"] & F_0 \arrow[r] & N \arrow[r] & 0 .
\end{tikzcd}$$

Output: *a presentation of* $\mathrm{Hom}_R(M,N)$.

1. *Compute generators* $F_2 \xrightarrow{b_2} F_1$ *of* $\ker(b_1 \colon F_1 \to F_0)$.

2. Compute the homology of the sequence

$$\mathrm{Hom}(E_0,F_1)\oplus\mathrm{Hom}(E_1,F_2)\to\mathrm{Hom}(E_0,F_0)\oplus\mathrm{Hom}(E_1,F_1)\to\mathrm{Hom}(E_1,F_0)$$

defined by

$$(h_0,h_1)\mapsto(b_1h_0,h_0a_1-b_2h_1)\ \textit{and}\ (\varphi_0,\varphi_1)\mapsto\varphi_0a_1-\varphi_1b_1.$$

Note that for free modules F and G the module $\mathrm{Hom}(F,G)$ is free of rank $\mathrm{Hom}(F,G)=\mathrm{rank}\,F\,\mathrm{rank}\,G$. So we have to compute in the Algorithm above the homology of a complex of free modules, which is simpler than the general case.

Algorithm 2.8.20 (Homology). . **Input:** *A complex*

$$\begin{tikzcd} M \arrow[r, "\varphi"] & N \arrow[r, "\psi"] & L \end{tikzcd}$$

specified via presentations

$$\begin{tikzcd} E_1 \arrow[r, "a_1"] \arrow[d, "\varphi_1"] & E_0 \arrow[r] \arrow[d, "\varphi_0"] & M \arrow[r] \arrow[d, "\varphi"] & 0\ . \\ F_1 \arrow[r, "b_1"] \arrow[d, "\psi_1"] & F_0 \arrow[r] \arrow[d, "\psi_0"] & N \arrow[r] \arrow[d, "\psi"] & 0 \\ G_1 \arrow[r, "c_1"] & G_0 \arrow[r] & L \arrow[r] & 0 \end{tikzcd}$$

Output: *A presentation of the homology*

$$\mathrm{H}=\mathrm{H}(M\to N\to L)=\frac{\ker(\psi)}{\mathrm{im}(\varphi)}.$$

1. Compute the syzygies matrix $(h_0\ g_0)^t$ of $\big(\psi_0\ c_1\big)$:

$$H_0 \overset{\begin{pmatrix}h_0\\ g_0\end{pmatrix}}{\longrightarrow} F_0\oplus G_1 \overset{\big(\psi_0\ c_1\big)}{\longrightarrow} G_0.$$

2. Compute the syzygy matrix $(h_1\ f_1\ e_0)^t$ in

$$H_1 \overset{\begin{pmatrix}h_1\\ f_1\\ e_0\end{pmatrix}}{\longrightarrow} H_0\oplus F_1\oplus E_0 \overset{\big(h_0\ b_1\ \varphi_0\big)}{\longrightarrow} F_0.$$

3. $\mathrm{H}=\mathrm{Coker}(H_1\xrightarrow{h_1}H_0)$.

Exercise 2.8.21 ($\ker, \operatorname{coker}$**).** Give an algorithm to compute $\operatorname{Ker}(\psi : N \to L)$ and $\operatorname{Coker}(\varphi : M \to N)$ by simplifying the computation of homology in these cases.

Exercise 2.8.22. Complete the Algorithm 2.8.19 for the computation of $\operatorname{Hom}(M, N)$ by including a simplified version of homology in this cases.

Exercise 2.8.23. Given a homorphism $N' \xrightarrow{f} N$ and a module M specified by presentations, design an algorithm which computes the presentations of

$$\operatorname{Hom}(N, M) \quad \overset{\operatorname{Hom}(f,M)}{\longrightarrow} \quad \operatorname{Hom}(N', M)$$

and

$$\operatorname{Hom}(M, N') \quad \overset{\operatorname{Hom}(M,f)}{\longrightarrow} \quad \operatorname{Hom}(M, N).$$

Algorithm 2.8.24 (Ext). . Input: *An integer i and two R-modules specified via free presentations*

$$\begin{tikzcd} F_1 \arrow[r, "a_1"] & F_0 \arrow[r] & M \arrow[r] & 0 \end{tikzcd}$$

and

$$\begin{tikzcd} G_1 \arrow[r, "b_1"] & G_0 \arrow[r] & N \arrow[r] & 0\,. \end{tikzcd}$$

Output: *a presentation of* $\operatorname{Ext}^i_R(M, N)$.

1. *Compute $i+1$ steps of a free resolution of M:*

$$F_{i+1} \overset{a_{i+1}}{\longrightarrow} F_i \overset{a_i}{\longrightarrow} F_{i-1} \overset{a_{i-1}}{\longrightarrow} \dots \overset{a_2}{\longrightarrow} F_1 \overset{a_1}{\longrightarrow} F_0$$

2. *Make a presentation of the complex* $\operatorname{Hom}(F_*, N)$*:*

$$\begin{tikzcd} \operatorname{Hom}(F_{i-1}, G_1) \arrow[r, "\operatorname{Hom}(F_{i-1},b_1)"] \arrow[d, "\operatorname{Hom}(a_i,G_1)"] & \operatorname{Hom}(F_{i-1}, G_0) \arrow[r] \arrow[d, "\operatorname{Hom}(a_i,G_0)"] & \operatorname{Hom}(F_{i-1}, N) \arrow[r] \arrow[d, "\operatorname{Hom}(a_i,N)"] & 0 \\ \operatorname{Hom}(F_i, G_1) \arrow[r, "\operatorname{Hom}(F_i,b_1)"] \arrow[d, "\operatorname{Hom}(a_{i+1},G_1)"] & \operatorname{Hom}(F_i, G_0) \arrow[r] \arrow[d, "\operatorname{Hom}(a_{i+1},G_0)"] & \operatorname{Hom}(F_i, N) \arrow[r] \arrow[d, "\operatorname{Hom}(a_{i+1},N)"] & 0 \\ \operatorname{Hom}(F_{i+1}, G_1) \arrow[r, "\operatorname{Hom}(F_{i+1},b_1)"] & \operatorname{Hom}(F_{i+1}, G_0) \arrow[r] & \operatorname{Hom}(F_{i+1}, N) \arrow[r] & 0 \end{tikzcd}$$

3. *Compute a presentation of the homology*

$$\operatorname{Ext}^i_R(M, N) = \operatorname{H}^i(\operatorname{Hom}(F_*, N)) = \frac{\ker \operatorname{Hom}(a_{i+1}, N)}{\operatorname{im} \operatorname{Hom}(a_i, N)}$$

Finally to compute Tor recall the definition of $M \otimes N$. For free modules F and G with basis f_i and g_j, the tensor product $F \otimes G$ is free on the basis $f_i \otimes g_j$. In general, given presentations

$$\begin{tikzcd} F_1 \arrow[r, "a_1"] & F_0 \arrow[r] & M \arrow[r] & 0 \end{tikzcd}$$

and

$$\begin{tikzcd} G_1 \arrow[r, "b_1"] & G_0 \arrow[r] & N \arrow[r] & 0 \end{tikzcd} ,$$

then

$$F_1 \otimes G_0 \oplus F_0 \otimes G_1 \overset{a_1 \otimes id_{G_0} + id_{F_0} \otimes b_1}{\longrightarrow} F_0 \otimes G_0 \to M \otimes N \to 0$$

is a presentation of the tensor product.

Algorithm 2.8.25 (Tor). . **Input:** *An integer i and two R-modules specified via free presentations*

$$\begin{tikzcd} F_1 \arrow[r, "a_1"] & F_0 \arrow[r] & M \arrow[r] & 0 \end{tikzcd}$$

and

$$\begin{tikzcd} G_1 \arrow[r, "b_1"] & G_0 \arrow[r] & N \arrow[r] & 0 \end{tikzcd} .$$

Output: *a presentation of* $\operatorname{Tor}_i^R(M, N)$*.*

1. *Compute $i+1$ steps of a free resolution of M:*

$$F_{i+1} \overset{a_{i+1}}{\longrightarrow} F_i \overset{a_i}{\longrightarrow} F_{i-1} \overset{a_{i-1}}{\longrightarrow} \ldots \overset{a_2}{\longrightarrow} F_1 \overset{a_1}{\longrightarrow} F_0$$

2. *Make a presentation of the complex $F_* \otimes N$:*

$$\begin{tikzcd} F_{i+1} \otimes G_1 \arrow[r, "id_{F_{i+1}} \otimes b_1"] \arrow[d, "a_{i+1} \otimes \mathrm{id}_{G_1}"] & F_{i+1} \otimes G_0 \arrow[r] \arrow[d, "a_{i+1} \otimes id_{G_0}"] & F_{i+1} \otimes N \arrow[r] \arrow[d, "a_{i+1} \otimes id_N"] & 0 \\ F_i \otimes G_1 \arrow[r, "id_{F_i} \otimes b_1"] \arrow[d, "a_i \otimes id_{G_1}"] & F_i \otimes G_0) \arrow[r] \arrow[d, "a_i \otimes id_{G_0}"] & F_i \otimes N \arrow[r] \arrow[d, "a_i \otimes id_N"] & 0 \\ F_{i-1} \otimes G_1 \arrow[r, "id_{F_{i-1}} \otimes b_1"] & F_{i-1} \otimes G_0 \arrow[r] & F_{i-1} \otimes N \arrow[r] & 0 \end{tikzcd}$$

3. *Compute a presentation of the homology*

$$\operatorname{Tor}_i^R(M, N) = \mathrm{H}_i(F_* \otimes N) = \frac{\ker(a_i \otimes id_N)}{\mathrm{im}(a_{i+1} \otimes id_N)}$$

Exercise 2.8.26. Given a module M and a short exact sequence

$$0 \to N' \to N \to N'' \to 0$$

specified via presentations, design an algorithm which computes the connecting homomorphisms

$$\mathrm{Ext}^i_R(M, N'') \to \mathrm{Ext}^{i+1}_R(M, N')$$

$$\mathrm{Ext}^i_R(N', M) \to \mathrm{Ext}^{i+1}_R(N'', M)$$

and

$$\mathrm{Tor}^R_i(N', M) \to \mathrm{Tor}^R_{i-1}(N'', M)$$

of the long exact sequences.

Exercise 2.8.27. Consider $R = \Bbbk[x_1, x_2, x_3, x_4]$ and $M = R/\langle x_1x_2, x_2x_3, x_1x_4, x_3x_4\rangle$ and $N = R/\langle x_1 - x_2, x_3 - x_4\rangle$. Compute all $\mathrm{Tor}^R_i(M, N)$.

Besides measuring exactness $\mathrm{Ext}^1_R(M, N)$ is used to describe extensions.

Definition 2.8.28. Let A and C be R-modules. An extension of C by A is an R-module B, together with a short exact sequence

$$0 \to A \to E \to C \to 0.$$

Two extension E, E' are isomorphic, if there exists a commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r] \arrow[d] & E \arrow[r] \arrow[d] & C \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & A \arrow[r] & E' \arrow[r] & C \arrow[r] & 0 \end{tikzcd}$$

with id_A and id_C as outer arrows.

$$0 \to A \to A \oplus C \to C \to 0$$

is called the trivial (or split) extension. □

Exercise 2.8.29. Prove:

1. An extension
$$0 \to A \to E \to C \to 0$$
is isomorphic to the split extention iff $id_C \in \mathrm{Hom}_R(C, C)$ maps to $0 \in \mathrm{Ext}^1_R(C, A)$ under the connecting homomorphism δ.
2. For every class $e \in \mathrm{Ext}^1(C, A)$ there exists an extension
$$0 \to A \to E \to C \to 0$$
with $\delta(id_C) = e$.

2.9 Additional exercises

Exercise 2.9.1. Let $\mathbb{F}_q$ be the field with q elements of characteristic p. Prove that

$$f_1 = x_1^q - x_q, \ldots, f_n = x_n^q - x_n \in \mathbb{F}_p[x_1 \ldots, x_n]$$

form a Gröbner basis for the ideal of the finitely many $\mathbb{F}_q$-rational points $\mathrm{I}(\mathbb{A}^n(\mathbb{F}_q))$ (with respect to any global monomial order on $\mathbb{F}_p[x_1 \ldots, x_n]$).

Exercise 2.9.2. Let $f_1, \ldots, f_r \in F = R$ be nonzero polynomials, and let $\mathrm{GCD}(\mathbf{L}(f_i), \mathbf{L}(f_j)) = 1$ for some pair (i, j). Show that a standard expression for $\mathrm{S}(f_i, f_j)$ with remainder zero is obtained by rewriting the syzygy $f_j f_i - f_i f_j = 0$.

Exercise 2.9.3. Integer programming, applications of binomial ideals.

Exercise 2.9.4. Show that the key properties of $>_{\mathrm{lex}}$ respectively $>_{\mathrm{drlex}}$ characterize these orders among all global monomial orders.
Hint. If $>$ satisfies the key property of $>_{\mathrm{drlex}}$, we have, for instance, $x_2^2 >_{\mathrm{drlex}} x_1 x_3$. Since $>$ is compatible with multiplication, also $x_2^2 x_4 >_{\mathrm{drlex}} x_1 x_3 x_4$. □

Exercise 2.9.5. Let $V = \mathrm{V}(I)$ be an absolutely irreducible variety defined by a binomial ideal. Show that V is rational.

Exercise 2.9.6. toric varieties

Exercise 2.9.7. Systems of polynomial equations of type

$$f_1 = x_1 - h_1(x_2, \ldots, x_n), \ldots, f_{n-1} = x_{n-1} - h_{n-1}(x_n), f_n = h_n(x_n)$$

are called **triangular** (the Gröbner basis in Exercise 2.4.6 gives an example). Such a system has at most $\deg f_n$ solutions in $\mathbb{A}^n(\overline{\Bbbk})$. It has precisely $\deg f_n$ solutions iff f_n is square-free.
... □

Exercise 2.9.8. module quotients, annihilators

Exercise 2.9.9 (5-Lemma). Let R be a ring, and let

$$\begin{tikzcd}
M_1 \arrow[r] \arrow[d, "\alpha_1"'] & M_2 \arrow[r] \arrow[d, "\beta_1"'] & M_3 \arrow[r] \arrow[d, "\gamma"'] & M_4 \arrow[r] \arrow[d, "\beta_2"'] & M_5 \arrow[d, "\alpha_2"'] \\
N_1 \arrow[r] & N_2 \arrow[r] & N_3 \arrow[r] & N_4 \arrow[r] & N_5
\end{tikzcd}$$

be a commutative diagram of R-modules with exact rows. Show that if β_1 and β_2 are isomorphisms, α_1 is an epimorphism, and α_2 is a monomorphism, then γ is an isomorphism.

Exercise 2.9.10. system solving

Chapter 3

Noether Normalization

We know from Chapter 1 that the strong version of Hilbert's Nullstellensatz follows from its weak version. Now, in the first section of this chapter, we will establish the weak version. A key ingredient of our proof is the projection theorem, which is interesting in its own right. In fact, interpreting the projection theorem from an algebraic point of view, we will be lead to the concept of integral ring extensions. Preparing, thus, the grounds for dimension theory, we will show three major results on prime ideals in integral ring extensions: the lying over theorem, the going up theorem, and the going down theorem. Dimension theory itself will take center stage in Sections 3.3 and 3.4. Motivated by our proof of the Nullstellensatz, and formulated in terms of affine rings, our definition of the dimension of an algebraic set relies on the concept of Noether normalization. There are several equivalent ways of characterizing dimension. A characterization in terms of leading ideals is the key to computing dimension via Gröbner bases. The notion of Krull dimension will allow us to assign a dimension to every ring.

In the final section of this chapter, starting from a field theoretic version of Noether normalization, we will show how to reduce problems concerning the birational geometry of varieties to problems concerning hypersurfaces.

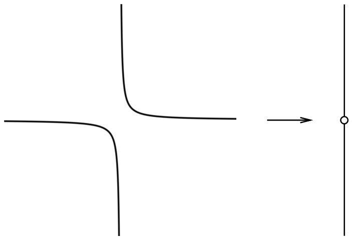
3.1 Proof of the Nullstellensatz

With notation as in Convention 2.7.2, our goal is to show that if $I \subset \mathbb{k}[x_1, \ldots, x_n]$ is an ideal, then $A = \mathrm{V}(I) \subset \mathbb{A}^n$ is empty iff $1 \in I$. As pointed out in Section 1.3, this is clearly true in the case of one variable. To prove the general case, we do induction on the number of variables, projecting A to $\mathbb{A}^{n-1}$ in order to connect to the induction hypothesis. Here, as already remarked in Section 2.6, we face the problem that the projected set may not be Zariski closed. The key point of our proof is to show that this problem may be overcome by choosing a sufficiently general projection map.

For this, we proceed in two steps. First, in the projection theorem, we specify an extra hypothesis which guarantees that the image of A under projection onto the last $n-1$ components is Zariski closed. Then, in Lemma 3.1.3, we show how to achieve the extra hypothesis by means of a triangular change of coordinates (which can be taken linear if $\Bbbk$ is infinite).

As some sort of motivation for the extra hypothesis, we discuss a simplified version of the example given in Exercise 2.6.1:

Example 3.1.1. Let $\pi : \mathbb{A}^2 \to \mathbb{A}^1$, $(a, b) \mapsto b$, be projection of the xy-plane onto the y-axis. Then π maps the hyperbola $C = \mathrm{V}(xy-1)$ onto the punctured line $\pi(C) = \mathbb{A}^1 \setminus \{0\}$ which is not an algebraic subset of $\mathbb{A}^1$:



If $\mathbb{K} = \mathbb{C}$, a reason for this failure can be seen in the fact that the function $y \mapsto 1/y$ is unbounded on C near $\mathbb{A}^1(\mathbb{C}) \times \{0\}$ in the Euclidean topology.

In contrast, suppose that $A \subset \mathbb{A}^2(\mathbb{C})$ is an algebraic set on which a monic equation in x of type

$$x^d + c_1(y)x^{d-1} + \ldots + c_d(y) = 0$$

is satisfied for some $d \geq 1$. Then, since the roots of this equation in x vary continously with y in the Euclidean topology, the preimage $(\pi|A)^{-1}(U)$ of any bounded domain $U \subset \mathbb{A}^1(\mathbb{C})$ is bounded as well. □

Taking our cue from this observation, we show:

Theorem 3.1.2 (Projection Theorem). *Let $I \subset \Bbbk[x_1, \ldots, x_n]$ be an ideal, and let $I_1 = I \cap \Bbbk[x_2, \ldots, x_n]$ be its first elimination ideal. Suppose that I contains a polynomial f which is monic in x_1 of some degree $d \geq 1$:*

$$f = x_1^d + c_1(x_2, \ldots, x_n)x_1^{d-1} + \ldots + c_d(x_2, \ldots, x_n),$$

with coefficients $c_i \in \Bbbk[x_2, \ldots, x_n]$. Let

$$\pi : \mathbb{A}^n \to \mathbb{A}^{n-1},\ (a_1, \ldots, a_n) \mapsto (a_2, \ldots, a_n),$$

be projection onto the last $n-1$ components, and let $A = \mathrm{V}(I) \subset \mathbb{A}^n$. Then

$$\pi(A) = \mathrm{V}(I_1) \subset \mathbb{A}^{n-1}.$$

In particular, $\pi(A)$ is Zariski closed.

Proof. As we already know, the inclusion $\pi(A) \subset \mathrm{V}(I_1)$ holds since $I_1 \subset I$. For the opposite inclusion, let $p' = (a_2, \ldots, a_n) \in \mathbb{A}^{n-1} \setminus \pi(A)$ be any point. To prove that $p' \in \mathbb{A}^{n-1} \setminus \mathrm{V}(I_1)$, we need to show that there is a polynomial $h \in I_1$ such that $h(p') \neq 0$. For this, we first suppose that $\Bbbk = \mathbb{K}$ is algebraically closed. In this case, we find the desired h in two steps:

Step 1. For each polynomial $g \in \Bbbk[x_1, \ldots, x_n]$, there is a polynomial $\widetilde{g} \in \Bbbk[x_1, \ldots, x_n]$ of degree $< d$ in x_1 such that $\widetilde{g}(x_1, p') = 0$ and $g \equiv \widetilde{g} \mod I$.

Indeed, consider the homomorphism

$$\phi\colon \Bbbk[x_1, \ldots, x_n] \to \Bbbk[x_1],\ g \mapsto g(x_1, p').$$

The image $\phi(I) \subset \Bbbk[x_1]$ is an ideal whose locus of zeros in $\mathbb{A}^1$ is empty by the assumption on p'. The Nullstellensatz in one variable implies that $\phi(I) = \Bbbk[x_1]$. In particular, if $g \in \Bbbk[x_1, \ldots, x_n]$ is any polynomial, we can find a polynomial $g_1 \in I$ such that $g(x_1, p') - g_1(x_1, p') = 0 \in \Bbbk[x_1]$. Set $g_2 = g - g_1$. Euclidean division with remainder in $\Bbbk[x_2, \ldots, x_n][x_1]$ yields an expression $g_2 = qf + \widetilde{g}$ such that the degree of $\widetilde{g}$ in x_1 is $< d$ (here, we make use of the assumption that f is monic in x_1 of degree d). Plugging in p', we see that $\widetilde{g}(x_1, p')$ is the unique remainder of degree $< d$ on Euclidean division of $0 = g_2(x_1, p')$ by $f(x_1, p')$ in $\Bbbk[x_1]$. Thus, $\widetilde{g}(x_1, p') = 0$. Moreover, $g - \widetilde{g} = qf + g_1 \in I$. That is, $g \equiv \widetilde{g} \mod I$.

Step 2. Applying the above to each of the polynomials $1, x_1, \ldots, x_1^{d-1}$, we get expressions

$$\begin{array}{rcll} 1 & \equiv & g_{00} + \ldots + g_{0,d-1}x_1^{d-1} & \mod I, \\ x_1 & \equiv & g_{10} + \ldots + g_{1,d-1}x_1^{d-1} & \mod I, \\ \vdots & & & \vdots \\ x_1^{d-1} & \equiv & g_{d-1,0} + \ldots + g_{d-1,d-1}x_1^{d-1} & \mod I, \end{array}$$

with $g_{ij} \in \Bbbk[x_2, \ldots, x_n]$ and $g_{ij}(p') = 0$ for all i, j. In matrix notation,

$$(E_d - B) \begin{pmatrix} 1 \\ \vdots \\ x_1^{d-1} \end{pmatrix} \equiv 0 \mod I,$$

where $B = (g_{ij})$ and E_d is the $d \times d$ identity matrix. Multiplying by the matrix of cofactors of $(E_d - B)$, we get

$$\det(E_d - B) \begin{pmatrix} 1 \\ \vdots \\ x_1^{d-1} \end{pmatrix} \equiv 0 \mod I.$$

In particular, $h := \det(E_d - B) \cdot 1 \in I \cap \Bbbk[x_2, \ldots, x_n] = I_1$. Moreover, $h(p') = 1 \neq 0$ since all the $g_{ij}(p')$ are zero.

This settles the case where $\Bbbk = \mathbb{K}$. For the general case, recall from Remark 2.7.1 on Buchberger's algorithm and field extensions that $I_1^e := I_1\, \mathbb{K}[x_2, \dots, , x_n]$ is the first elimination ideal of $I\, \mathbb{K}[x_1, \dots, , x_n]$. According to what we just proved, there is a polynomial in I_1^e which does not vanish at the point p'. Since I_1^e is generated by the polynomials in I_1, at least one of these polynomials does not vanish at p'. □

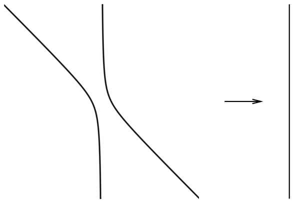
Lemma 3.1.3. *Let $f \in \Bbbk[x_1, \dots, x_n]$ be a nonconstant polynomial.*

1. *If $\Bbbk$ is infinite, let $a_2, \dots, a_n \in \Bbbk$ be sufficiently general. Substituting*
$$x_i = \widetilde{x}_i + a_i x_1$$
in f, $i = 2, \dots, n$, we get a polynomial of type
$$a x_1^d + c_1(\widetilde{x}_2, \dots, \widetilde{x}_n) x_1^{d-1} + \ldots + c_d(\widetilde{x}_2, \dots, \widetilde{x}_n),$$
where $a \in \Bbbk$ is a nonzero scalar, $d \geq 1$, and each $c_i \in \Bbbk[\widetilde{x}_2, \dots, \widetilde{x}_n]$.
2. *If $\Bbbk$ is arbitrary, let $r \in \mathbb{N}$ be sufficiently large. Substituting*
$$x_i = \widetilde{x}_i + x_1^{r^{i-1}}$$
in f, $i = 2, \dots, n$, we get a polynomial as in 1.

Proof. 1. Let $f = f_d + f_{d-1} + \ldots + f_0$, $f_d \neq 0$, be the decomposition of f into its homogeneous components. After substituting $\widetilde{x}_i + a_i x_1$ for x_i in f, $i = 2, \dots, n$, the coefficient of x_1^d is $f_d(1, a_2, \dots, a_n)$. Since f_d is homogeneous and nonzero, also $f_d(1, x_2, \dots, x_n)$ is nonzero. Thus, since $\Bbbk$ is infinite, $f_d(1, a_2, \dots, a_n)$ is nonzero for sufficiently general $a_2, \dots, a_n \in \Bbbk$ by Exercise 1.2.1. The result follows.

2. Write f as the finite sum of its terms,
$$f = \sum c_{\alpha_1 \ldots \alpha_n} x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n},$$
and let $r \in \mathbb{N}$. After substituting $\widetilde{x}_i + x_1^{r^{i-1}}$ for x_i in f, $i = 2, \dots, n$, the terms depending only on x_1 are of type $c_{\alpha_1 \ldots \alpha_n} x_1^{\alpha_1 + \alpha_2 r + \cdots + \alpha_n r^{n-1}}$. If r is strictly larger than all exponents α_i appearing in a term of f, the numbers $\alpha_1 + \alpha_2 r + \ldots + \alpha_n r^{n-1}$ are distinct for different $(\alpha_1, \dots, \alpha_n)$, and the terms depending only on x_1 cannot cancel with each other. The result follows. □

Example 3.1.4. Substituting $y = \widetilde{y} + x$ in $xy - 1$, we get the polynomial $x^2 + x\widetilde{y} - 1$ which is monic in x. Accordingly, the hyperbola $C = \mathrm{V}(xy - 1)$ projects *onto* $\mathbb{A}^1$ via $(a, b) \mapsto (a, b - a) \mapsto b - a$:



□

Exercise 3.1.5. Consider the ideal

$$I = \langle xy(x+y)+1 \rangle \subset \mathbb{F}_2[x,y].$$

Determine coordinates in which I satisfies the extra hypothesis of the projection theorem. Show that the extra hypothesis cannot be achieved by means of a *linear* change of coordinates. □

Proof of the Nullstellensatz, Weak Version. If $I \subset \Bbbk[x_1,\dots,x_n]$ is an ideal containing 1, its locus of zeros in $\mathbb{A}^n$ is clearly empty.

For the converse, suppose that the result is true for polynomials in $n-1$ variables, and let $I \subset \Bbbk[x_1,\dots,x_n]$ be an ideal such that $1 \notin I$. We have to show that $\mathrm{V}(I) \subset \mathbb{A}^n$ is nonempty. This is clear if $I = \langle 0 \rangle$. If I is nonzero, pick a nonconstant polynomial $f \in I$. In suitable coordinates $x_1, \widetilde{x}_2, \dots \widetilde{x}_n$, chosen as in Lemma 3.1.3, f becomes a monic polynomial in x_1 as required by the extra hypothesis of the projection theorem (adjust the constant leading term in x_1, if necessary). Since $1 \notin I$, we have $1 \notin I \cap \Bbbk[\widetilde{x}_2,\dots,\widetilde{x}_n]$ as well. It follows from the induction hypothesis that $\mathrm{V}(I \cap \Bbbk[\widetilde{x}_2,\dots,\widetilde{x}_n]) \subset \mathbb{A}^{n-1}$ contains a point. By the projection theorem, this point is the image of a point in $\mathrm{V}(I)$ under the projection which maps $(a_1,a_2,\dots,a_n)$ to $(\widetilde{a}_2,\dots,\widetilde{a}_n)$. In particular, $\mathrm{V}(I)$ is nonempty, and we are done by induction. □

Remark 3.1.6. Let $I \subset \Bbbk[x_1,\dots,x_n]$ be an ideal such that $1 \notin I$.

1. Successively carrying out the induction step in the proof above, applying Lemma 3.1.3 at each stage, we may suppose that the coordinates are chosen such that *each* nonzero elimination ideal $I_{k-1} = I \cap \Bbbk[x_k, x_{k+1},\dots,x_n]$ contains a monic polynomial of type

$$\begin{aligned} f_k = {} & x_k^{d_k} + c_1^{(k)}(x_{k+1},\dots,x_n)x_k^{d_k-1} + \dots + c_{d_k}^{(k)}(x_{k+1},\dots,x_n) \\ & \in \Bbbk[x_{k+1},\dots,x_n][x_k]. \end{aligned} \tag{3.1}$$

Then, if $0 \le c \le n$ is minimal with $I_c = \langle 0 \rangle$, each projection map

$$\pi_k : \mathrm{V}(I_{k-1}) \to \mathrm{V}(I_k),\ (a_k, a_{k+1},\dots,a_n) \mapsto (a_{k+1},\dots,a_n),$$

$1 \le k \le c$, is surjective. Hence, the composite map

$$\pi = \pi_c \circ \dots \circ \pi_1 : \mathrm{V}(I) \to \mathbb{A}^{n-c}.$$

is surjective as well. Furthermore, the π_k and, thus, π have finite fibers: if a point $(a_{k+1},\dots,a_n) \in \mathrm{V}(I_k)$ can be extended to a point $(a_k,a_{k+1},\dots,a_n) \in \mathrm{V}(I_{k-1})$, then a_k must be among the finitely many roots of the univariate polynomial $f_k(x_k,a_{k+1},\dots,a_n) \in \mathbb{K}[x_k]$.

2. In practical terms, combining the above with univariate root finding, we get the following recipe for finding explicit points of $\mathrm{V}(I)$.
Compute a lexicographic Gröbner basis $\mathcal{G}$ for I. Then $\mathcal{G}$ contains lexicographic Gröbner bases for the whole flag of elimination ideals I_{k-1}, $k = 1, \dots, n$. Moreover, the extra hypothesis of the projection theorem is fulfilled for *each* $I_{k-1} \neq \langle 0 \rangle$ iff polynomials f_k of type (3.1) are among the Gröbner basis elements (up to nonzero scalar factors).

In this case, every point $(a_{c+1}, \dots, a_n) \in \mathbb{A}^{n-c}$ can be extended to a point $(a_1, \dots, a_c, a_{c+1}, \dots, a_n) \in \mathrm{V}(I)$ by building up one coordinate at a time: If $(a_{k+1}, \dots, a_{c+1}, \dots, a_n) \in \mathrm{V}(I_k) \subset \mathbb{A}^{n-k}$ has already been chosen, consider the map

$$\varPhi_k : \Bbbk[x_k, x_{k+1}, \dots, x_n] \to \mathbb{K}[x_k],\ x_{k+1} \mapsto a_{k+1}, \dots, x_n \mapsto a_n.$$

The image $\varPhi_k(I_{k-1})$ is a principal ideal generated by the greatest common divisor of the images of the elements of $\mathcal{G} \cap \Bbbk[x_k, x_{k+1}, \dots, x_n]$. Pick a_k to be a root of that generator.

If one monic polynomial is missing, start over again in new coordinates. □

We will explore the full strength and the algebraic background of the observations made in the remark above in Sections 3.2 and 3.3.

Example 3.1.7. Consider the curve $C = \mathrm{V}(f_1, f_2) \subset \mathbb{A}^3(\mathbb{C})$, where

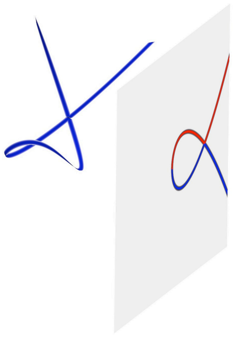
$$\begin{aligned} f_1 &= y^3z - 2y^2z - z^3 + x^2 + z, \\ f_2 &= xy^3z - 2xy^2z - xz^3 + x^3 + y^3 - 2y^2 + xz - z^2 + y. \end{aligned}$$

Computing the reduced lexicographic Gröbner basis for the ideal $\langle f_1, f_2 \rangle$, with variables ordered as $x > y > z$, we get the two polynomials below:

$$x^2 - yz + z, \quad y^3 - 2y^2 + y - z^2.$$

The first Gröbner basis element is monic in x of degree 2. Thus, projection of C to the yz-plane is $2 : 1$ and *onto* the curve C_1 defined by the second Gröbner basis element. In turn, C_1 is projected $3 : 1$ *onto* the z-axis. In sum, C is projected $6 : 1$ onto the z-axis.

The real picture below shows both curves C and C_1. Only the blue part of C_1 has real preimage points on C. The red part has complex preimage points.

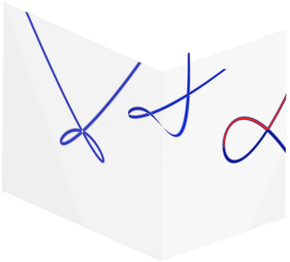


If we reorder the variables as $y > z > x$, the reduced lexicographic Gröbner basis for $\langle f_1, f_2\rangle$ consists of five polynomials:

$$\begin{array}{lll} y^3 - 2y^2 + y - z^2, & y^2x^2 - yx^2 - z^3, & yz - z - x^2, \\ yx^4 - z^4, & z^5 - zx^4 - x^6. & \end{array}$$

The image C_2 of C under projection to the zx-plane is defined by the last Gröbner basis element. Inspecting the other Gröbner basis elements, we see that every point $p \in C_2$ except the origin has a unique preimage point on C which is real iff p is real.

The following picture simultaneously shows C_2 and C_1:



□

Typically, in the situation of Remark 3.1.6, each of the successive projections except the last one is one-to-one over a Zariski dense part of the image (see **??** in Chapter 6 for a precise statement). In this sense, the projection to the zx-plane in our example above is more typical.

Exercise 3.1.8. Check that the polynomials

$$f_1 = x^3 - xz, \quad f_2 = yx^2 - yz \in \Bbbk[x, y, z]$$

form a lexicographic Gröbner basis. Conclude that $\mathrm{V}(f_1, f_2) \subset \mathbb{A}^3$ projects *onto* the yz-plane. Determine the points of the yz-plane with 1,2, and 3 preimage points, respectively. □

Exercise 3.1.9. Consider the ideal

$$I = \langle yz + 1, x(y + z) - 1\rangle \subset \mathbb{R}[x, y, z].$$

Determine coordinates in which *all* nonzero elimination ideals of I satisfy the extra hypothesis of the projection theorem. Compare the pictures of the corresponding algebraic sets in the given and new coordinates. □

3.2 Integral Ring Extensions

In the situation of the projection theorem, if $\pi_1 : \mathrm{V}(I) \to \mathrm{V}(I_1)$ is projection onto the last $n-1$ components, the extra hypothesis of the theorem guarantees that π_1 is surjective with finite fibers. This fact has a ring theoretic analogue, the lying over theorem, which is the first major result presented in this section. We begin by establishing the relevant terminology.

If R is a subring of a ring S, we say that $R \subset S$ is a **ring extension**. More generally, if $R \hookrightarrow S$ is any ring monomorphism, we identify R with its image in S and consider, thus, $R \subset S$ as a ring extension. With this notation, the algebraic counterpart of the map π_1 is the ring extension

$$R = \mathbb{k}[x_2, \dots, x_n]/I_1 \subset S = \mathbb{k}[x_1, \dots, x_n]/I$$

which is induced by the inclusion $\mathbb{k}[x_2, , \dots, x_n] \subset \mathbb{k}[x_1, \dots, x_n]$. We may, then, rephrase the extra hypothesis of the projection theorem by saying that the element $\overline{x}_1 = x_1 + I \in S$ is integral over R in the following sense:

Definition 3.2.1. Let $R \subset S$ be a ring extension. An element $s \in S$ is said to be **integral over *R*** if it satisfies a monic polynomial equation

$$s^d + r_1 s^{d-1} + \ldots + r_d = 0, \text{ with all } r_i \in R.$$

The equation is, then, called an **integral equation** for s over R. If, in addition, all the coefficients r_i are contained in some ideal I of R, we say that s is **integral over *I***, and call the equation an integral equation for s over I. If every element $s \in S$ is integral over R, we say that S is **integral over *R***, or that $R \subset S$ is an **integral extension**. □

Integral extensions are for rings what algebraic extensions are for fields. As in the special case of fields, we have two different notions of finiteness.

Definition 3.2.2. *Let $R \subset S$ be a ring extension.*

1. *We say that the extension is* **finite***, or that S is* **finite over *R***, *if S is finitely generated as an R-module. That is, the R-module S is the epimorphic image of a free R-module R^k.*
2. *We say that the extension is* **of finite type***, or that S is* **of finite type over *R***, *if S is finitely generated as an R-algebra. That is, the R-algebra S is the epimorphic image of a polynomial algebra $R[y_1, \dots, y_m]$.* □

Clearly, every finite extension is of finite type. Our next result shows that actually

finite type + integral = finite:

Proposition 3.2.3. *Let $R \subset S$ be a ring extension, let $s \in S$ (and let $I \subset R$ be an ideal). Then the following are equivalent:*

1. *s is integral over R (over I).*
2. *$R[s]$ is finite over R (and $s \in \operatorname{rad}(IR[s])$).*
3. *$R[s]$ is contained in a subring S' of S which is finite over R (and $s \in \operatorname{rad}(IS')$).*

In particular, if $s_1, \dots, s_m \in S$ are integral over R, then $R[s_1, \dots, s_m]$ is finite over R.

Proof. $1 \implies 2$: Let $f \in R[x]$ be a monic polynomial of degree d such that $f(s) = 0$. Division with remainder by f in $R[x]$ yields for every polynomial $g \in R[x]$ a representation $g = qf + r$ such that $\deg r < d$. Plugging in s, we get $g(s) = r(s)$. Hence, $1, s, \dots, s^{d-1}$ generate $R[s]$ as an R-module. If all coefficients of f are contained in I, it follows from the monic equation $f(s) = 0$ that $s^d \in IR[s]$ and, thus, that $s \in \text{rad}\,(IR[s])$.

$2 \implies 3$: Take $S' = R[s]$.

$3 \implies 1$: We argue as in Step 2 of the proof of the projection theorem. Let $m_1, \dots, m_l \in S'$ be a finite set of generators for S' as an R-module. If $s \in \text{rad}\,(IS')$, then $s^k \in IS'$ for some k. We use this to show that s is integral over I (if no ideal I is distinguished, take $I = R$ and $k = 1$ in what follows). For each i, we write $s^k m_i$ as an R-linear combination of the m_j:

$$s^k m_i = \sum_j r_{ij} m_j, \text{ with all } r_{ij} \in I.$$

In matrix notation,

$$(s^k E_l - B) \begin{pmatrix} m_1 \\ \vdots \\ m_l \end{pmatrix} = 0,$$

where $B = (r_{ij})$ and E_l is the $l \times l$ identity matrix. Multiplying by the matrix of cofactors of $(s^k E_l - B)$, we get $\det(s^k E_l - B) \cdot m_i = 0$ for every i. Since, in particular, $1 \in S'$ can be written as an R-linear combination of the m_i, the determinant must be zero. Expanding it, we get the desired integral equation for s over I. □

Corollary 3.2.4 (Transitivity of Integral Extensions). *If $R \subset S \subset T$ is a chain of ring extensions, and if T is integral over S, and S is integral over R, then T is integral over R.*

Proof. We apply Proposition 3.2.3. Let $t \in T$, and let $t^d + s_1 t^{d-1} + \cdots + s_d = 0$ be an integral equation for t over S. Then $R[s_1, \dots, s_d]$ and, thus, also $R[s_1, \dots, s_d, t] = \sum_{i=1}^{d-1} R[s_1, \dots, s_d] t^i$ are finite over R since the s_i are integral over R. In particular, t is integral over R. □

Corollary-Definition 3.2.5. *If $R \subset S$ is a ring extension, the set*

$$\{s \in S \mid s \text{ is integral over } R\}$$

is a subring of S containing R. It is called the ***integral closure*** *of R in S.*

Proof. We use, again, Proposition 3.2.3. If $s_1, s_2 \in S$ are integral over R, then $R[s_1, s_2]$ is finite over R. In particular, $s_1 \pm s_2$ and $s_1 s_2$ are integral over R. □

In the situation of the projection theorem,

$$R = \mathbb{k}[x_2, \ldots, x_n]/I_1 \subset S = \mathbb{k}[x_1, \ldots, x_n]/I = R[\overline{x}_1]$$

is a finite ring extension. Note that any extension of affine rings is of finite type. Hence, in this case, the notions of integral and finite extensions coincide.

Exercise* 3.2.6 (Integrality Criterion for Affine Rings). Let I be an ideal of $\mathbb{k}[x_1, \ldots, x_n]$, and let $\overline{f}_1 = f_1 + I, \ldots, \overline{f}_m = f_m + I \in \mathbb{k}[x_1, \ldots, x_n]/I$. Consider a polynomial ring $\mathbb{k}[y_1, \ldots, y_m]$, the homomorphism

$$\phi : \mathbb{k}[y_1, \ldots, y_m] \to S = \mathbb{k}[x_1, \ldots, x_n]/I,\ y_i \mapsto \overline{f}_i,$$

and the ideal

$$J = I\,\mathbb{k}[\boldsymbol{x}, \boldsymbol{y}] + \langle f_1 - y_1, \ldots, f_m - y_m \rangle \subset \mathbb{k}[\boldsymbol{x}, \boldsymbol{y}].$$

Let $>$ be an elimination order on $\mathbb{k}[\boldsymbol{x}, \boldsymbol{y}]$ with respect to $x_1, \ldots, x_n$, and let $\mathcal{G}$ be a Gröbner basis for J with respect to $>$. By Proposition 2.5.12, the elements of $\mathcal{G} \cap \mathbb{k}[\boldsymbol{y}]$ generate $\ker \phi$. View $R := \mathbb{k}[y_1, \ldots, y_m]/\ker \phi$ as a subring of S by means of ϕ. Show that $R \subset S$ is integral iff for each i, $1 \leq i \leq n$, there is an element of $\mathcal{G}$ whose leading monomial is of type $x_i^{\alpha_i}$ for some $\alpha_i \geq 1$. □

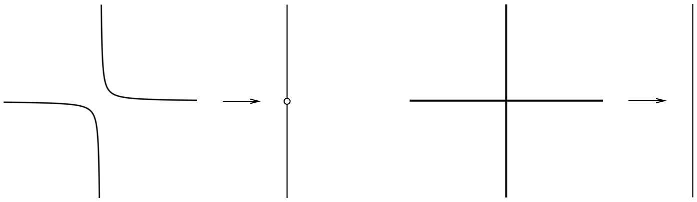
Example 3.2.7. Both ring extensions

$$\mathbb{k}[y] \subset \mathbb{k}[x, y]/\langle xy - 1 \rangle,\ y \mapsto \overline{y},$$

and

$$\mathbb{k}[y] \subset \mathbb{k}[x, y]/\langle xy \rangle,\ y \mapsto \overline{y},$$

are not integral (apply, for instance, the criterion given in Exercise 3.2.6).



Geometrically, in contrast to the situation of the projection theorem, projection of $\mathrm{V}(xy-1)$ to the y-axis is not onto (there is no point lying over $0 \in \mathbb{A}^1$), whereas projection of $\mathrm{V}(xy)$ to the y-axis is onto, but the fiber over 0 is not finite (there are infinitely many points lying over 0). □

In algebraic terms, lying over refers to maximal ideals instead of points. More generally, the lying over theorem stated below is a result concerning prime ideals. In this context, if $R \subset S$ is a ring extension, and $\mathfrak{P}$ is a prime ideal of S, then $\mathfrak{p} := \mathfrak{P} \cap R$ is necessarily a prime ideal of R, and we say that $\mathfrak{P}$ **lies over** $\mathfrak{p}$.

Theorem 3.2.8 (Lying Over). *Let $R \subset S$ be an integral ring extension, and let $\mathfrak{p}$ be a prime ideal of R. Then:*

1. *There exists a prime ideal $\mathfrak{P}$ of S lying over $\mathfrak{p}$:*

$$\begin{tikzcd}[row sep=small, column sep=small] {}^{\exists}\,\mathfrak{P} \arrow[d, dotted, no head] \arrow[r, phantom, "\subset"] & S \arrow[d, no head] \\ \mathfrak{p} \arrow[r, phantom, "\subset"] & R \end{tikzcd}$$

2. *There are no strict inclusions between prime ideals of S lying over $\mathfrak{p}$.*
3. *If $\mathfrak{P}$ is a prime ideal of S lying over $\mathfrak{p}$, then $\mathfrak{P}$ is maximal iff $\mathfrak{p}$ is.*
4. *If S is Noetherian, only finitely many prime ideals of S lie over $\mathfrak{p}$.* □

The proof of the theorem is based on the prime existence lemma of Krull which we show next. We need the following notation:

Definition 3.2.9. A subset U of a ring R is **multiplicatively closed** if $1 \in U$ and the product of any two elements of U is in U. □

Typical examples of multiplicatively closed sets are the subsets of type $U = \{f^k \mid k \geq 0\}$, where $f \in R$, and the subsets of type $U = R \setminus \mathfrak{p}$, where $\mathfrak{p} \subset R$ is a prime ideal.

Lemma 3.2.10 (Krull's Prime Existence Lemma). *Let R be a ring, let $I \subset R$ be an ideal, and let U be a multiplicatively closed subset of R such that $I \cap U = \emptyset$. Then there is a prime ideal $\mathfrak{p}$ of R containing I, and such that $\mathfrak{p} \cap U = \emptyset$.*

Proof. If R is Noetherian, the proof is yet another application of Noetherian induction. In the general case, we use Zorn's lemma, considering the set

$$\Gamma = \{J \subset R \text{ ideal} \mid I \subset J \text{ and } J \cap U = \emptyset\}.$$

This set is partially ordered by inclusion and nonempty since $I \in \Gamma$. Furthermore, if $\{J_\lambda\}$ is a totally ordered subset of Γ, then $J = \bigcup_\lambda J_\lambda \in \Gamma$ is an upper bound for this subset. By Zorn's lemma, there is a maximal element $\mathfrak{p}$ of Γ.

We show that $\mathfrak{p}$ is a prime ideal. First of all, $\mathfrak{p}$ is a proper ideal of R since otherwise $1 \in \mathfrak{p} \cap U = \emptyset$, absurd. Let, now, r_1, r_2 be elements of $R \setminus \mathfrak{p}$. Then, for $j = 1, 2$, the ideal $\mathfrak{p} + \langle r_j \rangle$ is not contained in Γ due to our choice of $\mathfrak{p}$. Hence, $(\mathfrak{p} + \langle r_j \rangle) \cap U \neq \emptyset$, which means that we can find elements $p_j \in \mathfrak{p}$ and $a_j \in R$ such that $p_j + a_j r_j \in U$, $j = 1, 2$. Then $(p_1 + a_1 r_1)(p_2 + a_2 r_2) \in U \subset R \setminus \mathfrak{p}$, so that $a_1 a_2 r_1 r_2 \notin \mathfrak{p}$. In particular, $r_1 r_2 \notin \mathfrak{p}$, as desired. □

At this point, we include two exercises with results needed in Chapter 4:

Exercise* 3.2.11. If R is a ring, show that its nilradical is the intersection of all the prime ideals of R:

$$\operatorname{rad}\langle 0 \rangle = \bigcap_{\mathfrak{p} \subset R \text{ prime}} \mathfrak{p}.$$

□

Exercise* **3.2.12.** If R is a ring containing only finitely minimal primes, show that these ideals contain zerodivisors only. □

Remark 3.2.13. Let $R \subset S$ be a ring extension, and let I be an ideal of S. Regard $R/(I \cap R)$ as a subring of S/I in the natural way, and suppose that S is integral over R. Then S/I is integral over $R/(I \cap R)$ as well. Indeed, if $\overline{s} = s + I \in S/I$, an integral equation for $\overline{s}$ over $R/(I \cap R)$ is obtained from an integral equation for s over R in the obvious way. □

Proof of the Lying Over Theorem. 1. Consider the ideal $\mathfrak{p}S$ generated by $\mathfrak{p}$ in S and the multiplicatively closed subset $U = R \setminus \mathfrak{p} \subset S$. Using the assumption that $R \subset S$ is integral, we will verify that $\mathfrak{p}S \cap U = \emptyset$. This will allow us, then, to apply Krull's prime existence lemma.

If $s \in \mathfrak{p}S$ is any element, there is an expression $s = \sum_{i=1}^{m} r_i s_i$, with all $r_i \in \mathfrak{p}$ and $s_i \in S$. Then $s \in \mathfrak{p}R[s_1, \dots, s_m]$, so that s is integral over $\mathfrak{p}$ by Proposition 3.2.3. Consider an integral equation

$$s^d + r_1 s^{d-1} + \ldots + r_d = 0$$

such that all $r_i \in \mathfrak{p}$. We have to show that $s \notin U$. Suppose the contrary. Then, in particular, $s \in R$, so that $s^d = -r_1 s^{d-1} - \cdots - r_d \in \mathfrak{p}$. Since $\mathfrak{p}$ is a prime ideal, it follows that $s \in \mathfrak{p}$, a contradiction to $s \in U = R \setminus \mathfrak{p}$.

This shows that $\mathfrak{p}S \cap U = \emptyset$. The prime existence lemma yields a prime ideal $\mathfrak{P}$ of S such that $\mathfrak{p} \subset \mathfrak{p}S \subset \mathfrak{P}$ and $\mathfrak{P} \cap R \subset R \setminus U = \mathfrak{p}$. Hence, $\mathfrak{P}$ is a prime ideal of S lying over $\mathfrak{p}$.

2. If $\mathfrak{P}_1 \subset \mathfrak{P}_2$ are two prime ideals of S lying over $\mathfrak{p}$, then $\overline{R} = R/\mathfrak{p}_1 \subset \overline{S} = S/\mathfrak{P}_1$ is an integral extension of integral domains such that $(\mathfrak{P}_2/\mathfrak{P}_1) \cap \overline{R} = \langle 0 \rangle$. We have to show that $\mathfrak{P}_1$ is not properly contained in $\mathfrak{P}_2$. Suppose the contrary. Then there is a nonzero element $\overline{s} \in \mathfrak{P}_2/\mathfrak{P}_1$, and we obtain a contradiction by considering an integral equation $\overline{s}^d + \overline{r}_1 \overline{s}^{d-1} + \ldots + \overline{r}_d = 0$ for $\overline{s}$ over $\overline{R}$ of smallest possible degree d. Indeed, since $\overline{r}_d \in (\mathfrak{P}_2/\mathfrak{P}_1) \cap \overline{R} = \langle 0 \rangle$ is zero and $\overline{S}$ is an integral domain, we may divide the equation by $\overline{s}$ to obtain an integral equation of smaller degree.

3. If $\mathfrak{p}$ is maximal, then $\mathfrak{P}$ is maximal as well by part 2. For the converse, consider the integral extension $R/\mathfrak{p} \subset S/\mathfrak{P}$. If $S/\mathfrak{P}$ is a field, its only maximal ideal is $\langle 0 \rangle$. Then, in turn, $\langle 0 \rangle$ is the only maximal ideal of $R/\mathfrak{p}$ by part 1, so that $R/\mathfrak{p}$ is a field, too.

4. If $\mathfrak{P}$ is a prime ideal of S lying over $\mathfrak{p}$, then $\mathfrak{p}S \subset \mathfrak{P}$. By part 2, $\mathfrak{P}$ is a minimal prime of $\mathfrak{p}S$. Since, by assumption, S is Noetherian, we may, then, apply Proposition 1.8.11 to conclude that $\mathfrak{P}$ is one of the finitely many minimal associated primes of $\mathfrak{p}S$. □

The following examples illustrate the lying over theorem and its proof.

Example 3.2.14. The ring extension

$$R = \mathbb{Z} \subset S = \mathbb{Z}[\sqrt{-5}] \cong \mathbb{Z}[x]/\langle x^2 + 5 \rangle$$

is integral, and the ideal $\mathfrak{p}$ generated by 2 in $\mathbb{Z}$ is maximal. The ideal generated by 2 in $\mathbb{Z}[\sqrt{-5}]$, however, is not even prime. Indeed, $(1+\sqrt{-5})\cdot(1-\sqrt{-5}) = 3\cdot 2 \in \mathfrak{p}$. Using that $\mathbb{Z}[\sqrt{-5}]/\langle 2\rangle \cong \mathbb{F}_2[x]/\langle x^2+1\rangle = \mathbb{F}_2[x]/\langle (x+1)^2\rangle$, we see that $\mathfrak{P} = \langle 2, 1+\sqrt{-5}\rangle \subset \mathbb{Z}[\sqrt{-5}]$ is the unique maximal ideal lying over $\mathfrak{p}$. □

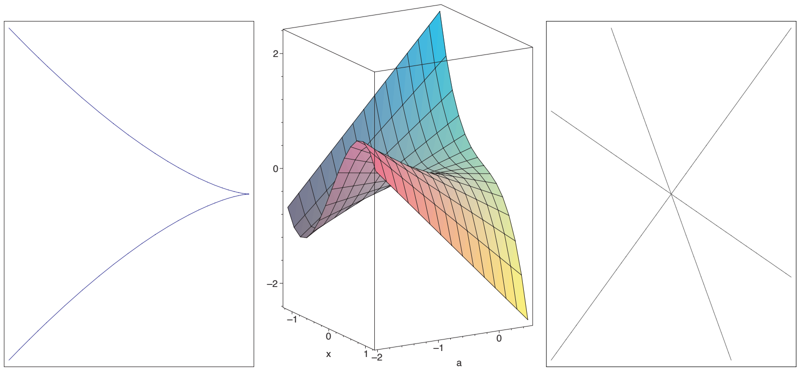
Example 3.2.15. The extension of polynomial rings

$$R = \mathbb{R}[e_2, e_3] \subset T = \mathbb{R}[t_1, t_2]$$

defined by $e_2 = t_1t_2 + t_1(-t_1-t_2) + t_2(-t_1-t_2)$ and $e_3 = t_1t_2(t_1+t_2)$ is integral by Proposition 3.2.3 since both t_1 and t_2 are roots of the equation $x^3 + e_2x + e_3 = 0$ (the third root is $-t_1-t_2$). Viewing

$$S = \mathbb{R}[t_1, e_2, e_3] \cong \mathbb{R}[x, e_2, e_3]/\langle x^3 + e_2x + e_3\rangle$$

as an intermediate ring in the natural way, we get a chain of integral ring extensions $R \subset S \subset T$.



R, S and T, and branch loci.

Let $\mathfrak{p} = \langle e_2 - a_2, e_3 - a_3\rangle \subset R$ be the maximal ideal corresponding to a point $(a_2, a_3) \in \mathbb{A}^2(\mathbb{R})$. The proof of part 4 of the lying over theorem shows that the maximal ideals of S and T lying over $\mathfrak{p}$ arise from primary decompositions of $\mathfrak{p}S$ and $\mathfrak{p}T$. On the other hand, the polynomial $t_1^3 + a_2t_1 + a_3$ has at least one real root, say b_1. Then $\mathfrak{p}_1 = \langle t_1 - b_1, e_2 - a_2, e_3 - a_3\rangle \subset S$ is a prime ideal lying over $\mathfrak{p}$, and with residue field $S/\mathfrak{p}_1 \cong \mathbb{R}$. If the other two roots of $t_1^3 + a_2t_1 + a_3$ are nonreal (they are, then, conjugate complex roots), the polynomial $t_1^2 + b_1t_1 - a_3/b_1$ is an irreducible factor of $t_1^3 + a_2t_1 + a_3$ over $\mathbb{R}$, so that $\mathfrak{p}_2 = \langle t_1^2 + b_1t_1 - a_3/b_1, e_2 - a_2, e_3 - a_3\rangle \subset S$ is a prime ideal lying over $\mathfrak{p}$, and with residue field $S/\mathfrak{p}_2 \cong \mathbb{C}$. It turns out that the number of real roots of $t_1^3 + a_2t_1 + a_3$ depends on the sign of the discriminant $D = -4e_2^3 - 27e_3^2$ evaluated at (a_2, a_3). If $D(a_2, a_3) < 0$, then $\mathfrak{p}S = \mathfrak{p}_1 \cap \mathfrak{p}_2$ decomposes into two maximal ideals such that, say, $S/\mathfrak{p}_1 \cong \mathbb{R}$ and $S/\mathfrak{p}_2 \cong \mathbb{C}$. Furthermore, $\mathfrak{p}T$ decomposes into three maximal ideals, all with residue field $\mathbb{C}$. If $D(a_2, a_3) > 0$, then $\mathfrak{p}S = \mathfrak{p}_1 \cap \mathfrak{p}_2 \cap \mathfrak{p}_3$ decomposes into 3 maximal ideals with residue fields

$S/\mathfrak{p}_i \cong \mathbb{R}$. Moreover, $\mathfrak{p}T$ decomposes into six maximal ideals, all with residue field $\mathbb{R}$. □

Exercise 3.2.16. Check all the statements made in Example 3.2.15. □

An important property of an integral ring extension $R \subset S$ is that nested pairs of prime ideals of R and of S are closely related. This is the content of two major results of Cohen-Seidenberg whose treatment is next. In Section 3.4, it will turn out that these results are fundamental to dimension theory.

Corollary 3.2.17 (Going Up Theorem of Cohen-Seidenberg). *Let $R \subset S$ be an integral ring extension. If $\mathfrak{p}_1 \subset \mathfrak{p}_2$ are prime ideals of R, and $\mathfrak{P}_1$ is a prime ideal of S lying over $\mathfrak{p}_1$, there exists a prime ideal $\mathfrak{P}_2$ of S lying over $\mathfrak{p}_2$ such that $\mathfrak{P}_1 \subset \mathfrak{P}_2$:*

$$\begin{tikzcd}
\mathfrak{P}_1 \arrow[r, phantom, "\subset\ \exists"] \arrow[d, no head] & \mathfrak{P}_2 \arrow[d, no head, dotted] \\
\mathfrak{p}_1 \arrow[r, phantom, "\subset"] & \mathfrak{p}_2
\end{tikzcd}$$

Proof. Applying the lying over theorem to the integral extension $\overline{R} = R/\mathfrak{p}_1 \subset \overline{S} = S/\mathfrak{P}_1$, we get a prime ideal $\overline{\mathfrak{P}}_2$ of $\overline{S}$ lying over $\mathfrak{p}_2/\mathfrak{p}_1$. The preimage $\mathfrak{P}_2$ of $\overline{\mathfrak{P}}_2$ in S has the desired properties. □

Though we arrived at the algebraic results presented so far in this section by revisiting the projection theorem and its proof, there is, as we show next, no need to restrict ourselves to projections if we want to view the results in the geometric context again. We use the following terminology:

Remark-Definition 3.2.18. Let $A \subset \mathbb{A}^n$ and $B \subset \mathbb{A}^m$ be algebraic sets, and let $\varphi : A \to B$ be a morphism. If $\varphi(A)$ is Zariski dense in B, the induced homomorphism $\varphi^* : R = \mathbb{K}[B] \to S = \mathbb{K}[A]$ is injective (see Lemma 2.6.21 and its proof). We regard, then, R as a subring of S by means of φ^*, and call φ a **finite morphism** if $R \subset S$ is an integral (hence finite) ring extension. □

Recall that a map between topological spaces is said to be **closed** if it sends closed sets to closed sets.

Corollary 3.2.19 (Properties of Finite Morphisms). *Let $\varphi : A \to B$ be a finite morphism of affine algebraic sets. Then:*

1. *If W is a subvariety of B, there is a subvariety V of A such that $\varphi(V) = W$. There are at most finitely many such varieties V. In particular, φ is surjective and has finite fibers.*
2. *The image of every algebraic subset of A under φ is an algebraic subset of B. That is, φ is closed with regard to the respective Zariski topologies.*
3. *If $W_1 \supset W_2$ is a nested pair of subvarieties of B, and V_1 is a subvariety of A such that $\varphi(V_1) = W_1$, there is a subvariety V_2 of V_1 such that $\varphi(V_2) = W_2$:*

$$\begin{tikzcd} V_1 \arrow[r, phantom, "\supset"] \arrow[d, no head] & {}^{\exists}\, V_2 \arrow[d, dotted, no head] \\ W_1 \arrow[r, phantom, "\supset"] & W_2 \end{tikzcd}$$

Proof. The assumption on φ means that φ^* constitutes an integral ring extension

$$R = \mathbb{K}[B] \subset S = \mathbb{K}[A].$$

1. Let $\mathfrak{p} = \mathrm{I}_B(W)$ be the vanishing ideal of W in R. By lying over, there is a prime ideal $\mathfrak{P} \subset S$ such that $\mathfrak{P} \cap R = \mathfrak{p}$. Then $V = \mathrm{V}_A(\mathfrak{P})$ is a subvariety of A such that $\varphi(V) \subset W$.

To show equality, let p be a point of W, and let $\mathfrak{m}$ be its vanishing ideal in R. Then $\mathfrak{p} \subset \mathfrak{m}$. Going up yields a prime ideal $\mathfrak{M}$ of S lying over $\mathfrak{m}$ and containing $\mathfrak{P}$:

$$\begin{tikzcd} \mathfrak{P} \arrow[r, phantom, "\subset"] \arrow[d, no head] & {}^{\exists}\, \mathfrak{M} \arrow[d, dotted, no head] \\ \mathfrak{p} \arrow[r, phantom, "\subset"] & \mathfrak{m} \end{tikzcd}$$

In fact, $\mathfrak{M}$ is a maximal ideal by part 3 of the lying over theorem. The Nullstellensatz implies that $\mathrm{V}_A(\mathfrak{M})$ consists of a single point $q \in V$. Necessarily, $p = \varphi(q) \in \varphi(V)$, so that $\varphi(V) = W$.

That only finitely many subvarieties of A are mapped onto W is clear since only finitely many prime ideals of S are lying over $\mathfrak{p}$.

2. Decomposing into irreducible components, we reduce to the case of a subvariety V of A. Then $V = \mathrm{V}_A(\mathfrak{P})$ for some prime ideal $\mathfrak{P}$ of S, and $W = \mathrm{V}_B(\mathfrak{P} \cap R)$ is a subvariety of B such that $\varphi(V) \subset W$. As in the proof of part 1, going up yields equality.

3. Again, apply the going up theorem as in the proof of part 1, replacing $\mathfrak{p} \subset \mathfrak{m}$ by $I_A(W_1) \subset I_A(W_2)$ and $\mathfrak{P}$ by $\mathrm{I}(V_1)$. □

Example 3.2.20. The algebraic subset $\mathrm{V}(xy^2 - y)$ of the xy-plane is the union of a hyperbola and a line. Projecting it to the x-axis, we get a morphism which is surjective and has finite fibers. However, this morphism is not finite. In fact, it is not even closed. □

Note that "going up" refers to the algebraic version of the theorem which gives a prime ideal $\mathfrak{P}_2$ larger than $\mathfrak{P}_1$. Remarkably enough, there is also a going down theorem. We need, however, a stronger hypothesis.

Example 3.2.21. Consider the homomorphism of polynomial rings

$$\phi : \Bbbk[x, y, z] \to \Bbbk[s, t],\ x \mapsto s,\ y \mapsto t^2 - 1,\ z \mapsto t(t^2 - 1).$$

Computing the reduced lexicographic Gröbner basis for the ideal

$$J = \langle s - x,\ t^2 - 1 - y,\ t(t^2 - 1) - z \rangle,$$

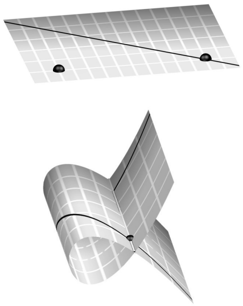
we get the polynomials

$$\begin{array}{lll} y^3+y^2-z^2, & tz-y^2-y, & ty-z, \\ t^2-y-1, & s-x. & \end{array}$$

Inspecting the Gröbner basis elements, we find: The kernel of ϕ is the principal ideal generated by the first Gröbner basis element $z^2-y^2(y+1)$, and the induced ring extension

$$R=\mathbb{k}[x,y,z]/\langle z^2-y^2(y+1)\rangle \subset S=\mathbb{k}[s,t]$$

is integral (apply the criterion given in Exercise 3.2.6). Geometrically, the map $\mathbb{A}^2 \to \mathbb{A}^3$ corresponding to the ring extension parametrizes $\mathrm{V}(z^2-y^2(y+1))$:



The ideal $\mathfrak{P}_1=\langle s-t\rangle$ is the unique prime ideal of S lying over the prime ideal $\mathfrak{p}_1=\mathfrak{P}_1\cap R=\langle \overline{x}^2-1-\overline{y}, \overline{x}(\overline{x}^2-1)-\overline{z}\rangle$ of R. The ideal $\mathfrak{p}_2=\langle \overline{x}-1,\overline{y},\overline{z}\rangle$ is a maximal ideal of R containing $\mathfrak{p}_1$. There are precisely two maximal ideals of S lying over $\mathfrak{p}_2$, namely $\langle s-1,t+1\rangle$ and $\langle s-1,t-1\rangle$. Their geometric counterparts are the two points in the plane which are distinguished in the picture. If $\mathfrak{P}_2$ is chosen to be $\mathfrak{P}_2=\langle s-1,t+1\rangle$, then $\mathfrak{P}_2$ does not contain $\mathfrak{P}_1$. Geometrically, the point $(1,-1)$ does not lie on the line $s=t$. Thus, "going down" does not hold in this example. □

Exercise* 3.2.22. Prove all the statements made in Example 3.2.21. □

Definition 3.2.23. Let R be an integral domain. The integral closure of R in its quotient field $\mathrm{Q}(R)$,

$$\widetilde{R} := \{s\in \mathrm{Q}(R) \mid s \text{ is integral over } R\},$$

is called the **normalization** of R. If $R=\widetilde{R}$, then R is said to be **normal**. □

Proposition 3.2.24. *If R is a UFD, then R is normal.*

Proof. Let $s\in\mathrm{Q}(R)$. Since R is a UFD, we may write s as a fraction $s=p/q$ such that p and q are coprime. Let

$$s^d+r_1s^{d-1}+\ldots+r_d=0$$

be an integral equation for s over R. Multiplying by q^d, the equation becomes

$$p^d = -q(r_1 p^{d-1} + \ldots + r_d q^{d-1}) \in R.$$

So p is divisible by q since R is a UFD. Since p and q are coprime, we conclude that q is a unit, and that $s = pq^{-1} \in R$. □

Example 3.2.25. 1. The polynomial ring $\Bbbk[x_1, \ldots, x_n]$ is factorial and, thus, normal.

2. The ring $R = \Bbbk[x, y, z]/\langle z^2 - y^2(y+1)\rangle$ in Example 3.2.21 is not normal since $t = z/y \in Q(R) \setminus R$ is integral over R. □

Exercise 3.2.26. Show that the following rings are integral domains, and find their normalizations:

1. The coordinate ring of the plane curve $V(y^2 - x^{2k+1}) \subset \mathbb{A}^2$, where $k \geq 1$.
2. The coordinate ring of the Whitney umbrella $V(x^2 - y^2 z) \subset \mathbb{A}^3$. □

In the proof of the Going Down Theorem 3.2.28 presented below, we will use the following result:

Lemma 3.2.27. *Let R be a normal ring, let $K = Q(R)$ be its quotient field, let $L \supset K$ be an extension field, and let $\mathfrak{p}$ be a prime ideal of R. If $s \in L$ is integral over $\mathfrak{p}$, then s is algebraic over K, and if $p_s = x^d + c_1 x^{d-1} + \cdots + c_d$ is the minimal polynomial of s over K, all coefficients c_i lie in $\mathfrak{p}$.*

Proof. Clearly, s is algebraic over K. Let $\overline{K}$ be the algebraic closure of K, and let $s = s_1, \ldots, s_d \in \overline{K}$ be the roots of p_s. Then, for each j, there is an automorphism of $\overline{K}$ fixing K and mapping s to s_j. Thus, if $f(s) = 0$ is an integral equation for s, where $f \in R[x]$ has coefficients in $\mathfrak{p}$, then also $f(s_j) = 0$ for each j. We conclude that the s_j are integral over $\mathfrak{p}$. Since the coefficients c_i of p_s are polynomial expressions in the s_j, it follows from Proposition 3.2.3 that the c_i must lie in rad $(\mathfrak{p}\widetilde{R})$, where $\widetilde{R} \subset K$ is the normalization of R. Since $R = \widetilde{R}$ and rad $\mathfrak{p} = \mathfrak{p}$ by our assumptions on R and $\mathfrak{p}$, the c_i actually lie in $\mathfrak{p}$, as desired. □

Theorem 3.2.28 (Going Down Theorem of Cohen-Seidenberg). *Let $R \subset S$ be an integral extension of integral domains, with R normal. If $\mathfrak{p}_1 \subset \mathfrak{p}_2$ are prime ideals of R, and $\mathfrak{P}_2$ is a prime ideal of S lying over $\mathfrak{p}_2$, there exists a prime ideal $\mathfrak{P}_1$ of S lying over $\mathfrak{p}_1$ such that $\mathfrak{P}_1 \subset \mathfrak{P}_2$:*

$$\begin{tikzcd}
{}^{\exists}\,\mathfrak{P}_1 \arrow[r, phantom, "\subset"] \arrow[d, dotted, no head] & \mathfrak{P}_2 \arrow[d, no head] \\
\mathfrak{p}_1 \arrow[r, phantom, "\subset"] & \mathfrak{p}_2
\end{tikzcd}$$

Proof. We consider three multiplicatively closed subsets of S:

$$U_1 := R \setminus \mathfrak{p}_1,\ U_2 := S \setminus \mathfrak{P}_2,\ \text{and}\ U := U_1 \cdot U_2 = \{r \cdot s \mid r \in U_1, s \in U_2\}.$$

As a first step of the proof, we show that $\mathfrak{p}_1 S \cap U = \emptyset$. Then we apply Krull's prime existence lemma to obtain the result.

Step 1. Suppose there is an element $s \in \mathfrak{p}_1 S \cap U$.

Then s is is integral over $\mathfrak{p}_1$ since $s \in \mathfrak{p}_1 S$ (argue as in the proof of the first part of the lying over theorem). Applying Lemma 3.2.27, we see that the minimal polynomial of $s \in L = \mathrm{Q}(S)$ over $K = \mathrm{Q}(R)$ is of type $p_s = x^d + c_1 x^{d-1} + \cdots + c_d$, with coefficients $c_i \in \mathfrak{p}_1 \subset R$.

Since $s \in U$, we may write s as a product $s = r \cdot \widetilde{s}$, with $r \in U_1$ and $\widetilde{s} \in U_2$. Then

$$p_{\widetilde{s}} = x^d + \frac{c_1}{r} x^{d-1} + \cdots + \frac{c_d}{r^d}$$

is the minimal polynomial of $\widetilde{s}$ over K. Applying, again, Lemma 3.2.27, we see that the coefficients c_i/r^i of $p_{\widetilde{s}}$ must be contained in R since $\widetilde{s}$ is integral over R. In fact, the c_i/r^i are contained in $\mathfrak{p}_1$ since $c_i \in \mathfrak{p}_1$ and $r^i \notin \mathfrak{p}_1$ for each i. It follows that $\widetilde{s}$ is even integral over $\mathfrak{p}_1$. So $\widetilde{s} \in \mathrm{rad}\,(\mathfrak{p}_1 S) \subset \mathfrak{P}_2$ by Proposition 3.2.3, a contradiction to $\widetilde{s} \in U_2$.

Step 2. Krull's prime existence lemma yields a prime ideal $\mathfrak{P}_1$ of S such that $\mathfrak{p}_1 S \subset \mathfrak{P}_1$ and $\mathfrak{P}_1 \cap U = \emptyset$. In particular, $\mathfrak{P}_1 \cap U_1 = \emptyset$, so that $\mathfrak{P}_1$ is lying over $\mathfrak{p}_1$, and $\mathfrak{P}_1 \cap U_2 = \emptyset$, so that $\mathfrak{P}_1 \subset \mathfrak{P}_2$. □

Remark 3.2.29. Even if R is a Noetherian integral domain, its normalization $\widetilde{R}$ need not be Noetherian. In particular, the extension $R \subset \widetilde{R}$ need not be finite (see Nagata (1962), Appendix A1. Examples of bad Noetherian rings). It is finite, however, if R is an affine domain. In this case, $\widetilde{R}$ is again an affine domain. The proof of this important finiteness result of Emmy Noether makes use of Noether normalization and Galois theory (see, for instance, Eisenbud (1995), Corollary 13.13). We refer to de Jong (1998) for an algorithm which computes the normalization of affine domains. □

3.3 Noether Normalization

In the previous section, we proved results which reflect the projection theorem from an algebraic point of view. In this section, we will revisit our original application of the projection theorem, namely the proof of the Nullstellensatz. As pointed out in Remark 3.1.6, this proof yields a composition of projections

$$\pi = \pi_c \circ \cdots \circ \pi_1 : A = \mathrm{V}(I) \to \mathbb{A}^{n-c}$$

which is surjective and has finite fibers (see Figure 3.1 for an illustration). Intuitively, the number $d = n - c$ should be the dimension of A. To make this a formal definition, it is convenient to work on the level of rings. We will use:

Theorem-Definition 3.3.1. *If S is an affine $\Bbbk$-algebra, there are elements $y_1, \dots, y_d \in S$ such that:*

1. *$y_1, \dots, y_d$ are algebraically independent over $\Bbbk$.*
2. *$\Bbbk[y_1, \dots, y_d] \subset S$ is an integral (hence finite) ring extension.*

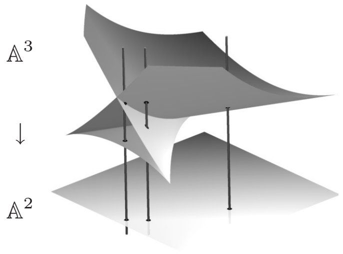


Fig. 3.1. *We project a surface which is called the* **swallowtail**. *Its equation is* $-4z^3y^2 - 27y^4 + 16z^4x - 128x^2z^2 + 144xzy^2 + 256x^3{=}0$.

If $y_1, \dots, y_d$ *satisfy conditions 1 and 2, the inclusion*

$$\Bbbk[y_1, \dots, y_d] \subset S$$

is called a **Noether normalization** *for* S.

Proof. We rewrite the proof of the Nullstellensatz in algebraic terms. Let $S = \Bbbk[x_1, \dots, x_n]/I$ for some ideal I of some polynomial ring $\Bbbk[x_1, \dots, x_n]$. As in Remark 3.1.6, we suppose that the coordinates are chosen such that each nonzero elimination ideal $I_{k-1} = I \cap \Bbbk[x_k, \dots, x_n]$ contains a polynomial which is monic in x_k. Then, if c is the smallest integer such that $I_c = \langle 0 \rangle$, we have a sequence of integral ring extensions

$$\Bbbk[x_{c+1}, \dots, x_n] \subset \Bbbk[x_c, \dots, x_n]/I_{c-1} \subset \dots \subset S$$

whose composite is a Noether normalization for S, with $d = n - c$. □

Remark 3.3.2. If $\Bbbk$ is infinite, and finitely many generators for S over $\Bbbk$ are given, the y_i may be chosen to be $\Bbbk$-linear combinations of the generators. □

In practical terms, Remark 3.1.6 shows one way of finding a Noether normalization for $\Bbbk[x_1, \dots, x_n]/I$. To begin with, compute a lexicographic Gröbner basis $\mathcal{G}$ for I. Let c be defined as in the proof above. For each $0 \le k \le c-1$, check whether $\mathcal{G}$ contains a polynomial in $x_k, \dots, x_n$ which is monic in x_k (up to a nonzero scalar factor). If so, the composition

$$R = \Bbbk[x_{c+1}, \dots, x_n] \subset \Bbbk[x_1, \dots, x_n] \to S = \Bbbk[x_1, \dots, x_n]/I$$

is a Noether normalization. If one of the monic polynomials is missing, start over again in new coordinates.

Since $>_{\text{lex}}$ is an expensive monomial order, a Gröbner basis computation with respect to $>_{\text{drlex}}$ may detect a Noether normalization faster – provided the sufficient conditions given below are satisfied:

Proposition 3.3.3. *Let $I \subsetneq \Bbbk[x_1, \dots, x_n]$ be an ideal, and let $>$ be a global monomial order on $\Bbbk[x_1, \dots, x_n]$. Suppose that, for some c, the following two conditions hold:*

1. $\mathbf{L}(I) \cap \Bbbk[x_{c+1}, \dots, x_n] = \langle 0 \rangle$.
2. $\mathbf{L}(I) \supset \langle x_1, \dots, x_c \rangle^m$ *for some* m.

Then the composition

$$R = \Bbbk[x_{c+1}, \dots, x_n] \subset \Bbbk[x_1, \dots, x_n] \to S = \Bbbk[x_1, \dots, x_n]/I$$

is a Noether normalization. For the lexicographic order, the conditions are also necessary.

Proof. The residue classes $\overline{x}_{c+1}, \dots, \overline{x}_n \in S$ are algebraically independent over $\Bbbk$ iff the map $R \to S$ is a ring inclusion iff $I \cap \Bbbk[x_{c+1}, \dots, x_n] = \langle 0 \rangle$. This condition is obviously satisfied if condition 1 holds. For $>_{\text{lex}}$, also the converse is true due to the key property of $>_{\text{lex}}$ (see Section 2.5).

On the other hand, by Macaulay's Theorem 2.3.5, the R-module S is finitely generated iff there are only finitely many monomials in $\Bbbk[x_1, \dots, x_c]$ which are not contained in $\mathbf{L}(I)$. This, in turn, is equivalent to condition 2. □

Example 3.3.4. Let $C \subset \mathbb{A}^3$ be the twisted cubic curve. By Exercise 2.5.6, the reduced Gröbner basis for $\mathrm{I}(C)$ with respect to $>_{\text{drlex}}$ consists of the three polynomials

$$f_1 = x^2 - y, \quad f_2 = xy - z, \quad f_3 = y^2 - xz.$$

Hence, $\mathbf{L}(\mathrm{I}(C)) = \langle x^2, xy, y^2 \rangle = \langle x, y \rangle^2$, and it follows from Proposition 3.3.3 that

$$\Bbbk[z] \subset \Bbbk[x, y, z]/\mathrm{I}(C)$$

is a Noether normalization. □

If the conditions of Proposition 3.3.3 are not satisfied, start over again in new coordinates, and hope for the best. In some cases, the conditions can be achieved by just permuting the variables. In contrast to a general change of coordinates, a permutation of variables does not destroy sparseness.

Now, we come to the definition of dimension:

Definition 3.3.5. Let $\emptyset \neq A \subset \mathbb{A}^n$ be an algebraic set. If $\Bbbk$ is a field of definition of A, if $I \subset \Bbbk[x_1, \dots, x_n]$ is an ideal such that $A = \mathrm{V}(I)$, and if

$$\Bbbk[y_1, \dots, y_d] \subset \Bbbk[x_1, \dots, x_n]/I$$

is a Noether normalization, we define d to be the **dimension** of A, written

$$\dim A = d.$$

By convention, the dimension of the empty subset of $\mathbb{A}^n$ is -1. □

To show that $\dim A$ is well defined, we characterize the number d above in field theoretic terms:

Theorem 3.3.6 (Dimension Theorem). *Definition 3.3.5 is independent of the choices made. Furthermore, we have:*

1. *The dimension of an algebraic subset of $\mathbb{A}^n$ is the maximum dimension of its irreducible components.*
2. *If $V \subset \mathbb{A}^n$ is a variety, and $\mathbb{K}(V)$ is its rational function field, then*

$$\dim V = \operatorname{trdeg}_{\mathbb{K}} \mathbb{K}(V).$$

Proof. Using the notation of Definition 3.3.5, we proceed in four steps. In Steps 1 and 2, we show that it is enough to consider the case where $\Bbbk[x_1, \dots, x_n]/I$ is the coordinate ring of A. In Step 3, we reduce to the case of a variety which, in turn, is dealt with in Step 4. The last two steps show at the same time that dimension can be characterized as in statements 1 and 2.

Step 1. Whether elements $y_1, \dots, y_d \in \Bbbk[x_1, \dots, x_n]/I$ satisfy the two conditions in Theorem 3.3.1 can be checked using Gröbner bases (see Proposition 2.5.12 and Exercise 3.2.6). Taking Remark 2.7.1 on Buchberger's algorithm and field extensions into account, we find that $\Bbbk[y_1, \dots, y_d] \subset \Bbbk[x_1, \dots, x_n]/I$ is a Noether normalization iff

$$\mathbb{K}[y_1, \dots, y_d] \subset \mathbb{K}[x_1, \dots, x_n]/I\,\mathbb{K}[x_1, \dots, x_n]$$

is a Noether normalization.

Step 2. Let $\Bbbk[y_1, \dots, y_d] \subset \Bbbk[x_1, \dots, x_n]/I$ be a Noether normalization. Then the composition

$$\phi : \Bbbk[y_1, \dots, y_d] \subset \Bbbk[x_1, \dots, x_n]/I \to \Bbbk[x_1, \dots, x_n]/(\operatorname{rad} I)$$

is injective and, thus, a Noether normalization as well. Indeed, otherwise, we could find a nonzero element $f \in \ker \phi$. According to the definition of the radical, a suitable power of f would, then, define a nontrivial $\Bbbk$-algebra relation on $y_1, \dots, y_d \in \Bbbk[x_1, \dots, x_n]/I$.

Step 3. Let

$$\mathbb{K}[y_1, \dots, y_d] \subset \mathbb{K}[A]$$

be a Noether normalization, and let $A = V_1 \cup \dots \cup V_s$ be the decomposition of A into its irreducible components. For each i, consider the composition

$$\phi_i : \mathbb{K}[y_1, \dots, y_d] \subset \mathbb{K}[A] \to \mathbb{K}[V_i].$$

There are two possiblities. Either, ϕ_i is injective and, thus, a Noether normalization. Or, the composition of the induced map $\mathbb{K}[y_1, \dots, y_d]/\ker \phi_i \to \mathbb{K}[V_i]$ with a Noether normalization $\mathbb{K}[z_1, \dots, z_e] \to \mathbb{K}[y_1, \dots, y_d]/\ker \phi$ is a Noether normalization such that $e < d$. But ϕ_i is injective for at least one i. Indeed, otherwise, we could find a nonzero element $f_i \in \ker \phi_i$ for each i, and the product of the f_i would define a nontrivial $\mathbb{K}$-algebra relation on $y_1, \dots, y_d \in \mathbb{K}[A]$.

Step 4. If $V \subset \mathbb{A}^n$ is a variety, and $\mathbb{K}[y_1, \ldots, y_d] \subset \mathbb{K}[V]$ is a Noether normalization of its coordinate ring, then $\mathbb{K}(y_1, \ldots, y_d) \subset \mathbb{K}(V)$ is an algebraic field extension. Hence,

$$d = \operatorname{trdeg}_{\mathbb{K}} \mathbb{K}(y_1, \ldots, y_d) = \operatorname{trdeg}_{\mathbb{K}} \mathbb{K}(V). \qquad \square$$

With notation as in Definition 3.3.5, let V be an irreducible component of $\mathrm{V}(I)$ of maximal dimension $d = \operatorname{trdeg}_{\mathbb{K}} \mathbb{K}(V) = \dim \mathrm{V}(I)$. Since $\mathbb{K}(V) = \mathbb{K}(\overline{x}_1, \ldots, \overline{x}_n)$ is generated by the coordinate functions on V, there is an algebraically independent set of these of cardinality d. In other words, there is a subset of variables $\boldsymbol{u} \subset \{x_1, \ldots, x_n\}$ of cardinality d such that $\mathrm{I}(V) \cap \mathbb{K}[\boldsymbol{u}] = \langle 0 \rangle$. In particular, $I \cap \Bbbk[\boldsymbol{u}] = \langle 0 \rangle$. Together with the argument given in the proof of Theorem 3.3.8 below, this shows that d is the maximum cardinality of a subset $\boldsymbol{u} \subset \{x_1, \ldots, x_n\}$ such that $I \cap \Bbbk[\boldsymbol{u}] = \langle 0 \rangle$.

Example 3.3.7. The monomial ideal $I = \langle xz, yz \rangle$ defines the union of the xy-plane and the z-axis in $\mathbb{A}^2$. Since $I \cap \Bbbk[x, y] = \langle 0 \rangle$, we must have $\dim \mathrm{V}(I) = 2$. On the other hand, $I \cap \Bbbk[z]$ is zero, too, but $\{z\}$ cannot be enlarged to a set of variables $\boldsymbol{u}$ of cardinality 2 such that $I \cap \Bbbk[\boldsymbol{u}] = \langle 0 \rangle$. □

One way of finding the dimension of an algebraic set is to compute a Noether normalization for its coordinate ring as discussed earlier in this section. This may require that we apply a sufficiently general change of coordinates which usually makes subsequent computations expensive. The characterization of dimension in terms of elimination ideals given above is, at least for arbitrary ideals, even less practical since it requires the computation of quite a number of Gröbner bases with respect to different elimination orders. In the case of a monomial ideal, however, the computation of the elimination ideals is comparatively cheap. Thus, the following result is the key to computing dimension in the case of arbitrary ideals:

Theorem 3.3.8. *Let $I \subsetneq \Bbbk[x_1, \ldots, x_n]$ be an ideal, let $\mathrm{V}(I)$ be its locus of zeros in $\mathbb{A}^n$, and let $>$ be a global monomial order on $\Bbbk[x_1, \ldots, x_n]$. Then*

$$\dim \mathrm{V}(I) = d,$$

where d is the maximum cardinality of a subset of variables $\boldsymbol{u} \subset \{x_1, \ldots, x_n\}$ such that

$$\mathbf{L}(I) \cap \Bbbk[\boldsymbol{u}] = \langle 0 \rangle.$$

Proof. Applying, again, Remark 2.7.1 on Buchberger's algorithm and field extensions, we see that any set of monomial generators for $\mathbf{L}(I)$ also generates $\mathbf{L}(I\,\mathbb{K}[x_1, \ldots, x_n])$. We may, hence, suppose that $\Bbbk = \mathbb{K}$ is algebraically closed.

To show that $\dim \mathrm{V}(I) \geq d$, consider an integer $k > \dim \mathrm{V}(I)$, and let $\mathrm{V}(I) = V_1 \cup \cdots \cup \mathrm{V}_s$ be the decomposition of $\mathrm{V}(I)$ into its irreducible components. Then, for every set of variables $\boldsymbol{u} = \{x_{i_1}, \ldots, x_{i_k}\}$ and every component V_j, the coordinate functions $\overline{x}_{i_1}, \ldots, \overline{x}_{i_k} \in \Bbbk(V_j)$ are algebraically dependent over $\Bbbk$ by Theorem 3.3.6. This means that, for each j, there is a nonzero polynomial $f_j \in \Bbbk[\boldsymbol{u}]$ vanishing on V_j. By Hilbert's Nullstellensatz, a suitable power of the product $f_1 \cdots f_s$ lies in I. In particular, $I \cap \Bbbk[\boldsymbol{u}] \neq \langle 0 \rangle$, so that also $\mathbf{L}(I) \cap \Bbbk[\boldsymbol{u}] \neq \langle 0 \rangle$.

To show that $\dim \mathrm{V}(I)$ is exactly d, we need Hilbert functions of algebraic sets in weighted projective spaces. These will be introduced in Chapter 6. *We will complete the proof of the theorem in Exercise 6.4.52.* □

Example 3.3.9. We already know from Exercise 3.3.4 that the dimension of the twisted cubic curve is 1. Applying Theorem 3.3.8, this can be seen as follows: Considering, again, the reduced Gröbner basis

$$f_1 = x^2 - y, \quad f_2 = xy - z, \quad f_3 = y^2 - xz$$

for $\mathrm{I}(C)$ with respect to $>_{\mathrm{drlex}}$, we find that $\boldsymbol{u} = \{z\}$ is a set of variables of maximal cardinality such that

$$\langle x^2, xy, y^2 \rangle \cap \Bbbk[\boldsymbol{u}] = \langle 0 \rangle.$$

□

By Theorem 3.3.6, the dimension of an algebraic set A is the maximum dimension of its irreducible components. If all the components have the same dimension d, we say that A is **equidimensional** of dimension d. The words **curve**, **surface**, and **volume** (or **threefold**) refer to an equidimensional algebraic set of dimension 1,2, and 3, respectively.

Exercise* 3.3.10. Let $A \subset \mathbb{A}^n$ be an algebraic set. Show that A is a hypersurface iff it is equidimensional of dimension $n-1$. □

In arbitrary dimension, we get sufficient conditions for equidimensionality by strengthening condition 1 in Proposition 3.3.3. This is the content of the following two results.

Proposition 3.3.11. *Let I be a proper ideal of $\Bbbk[x_1, \dots, x_n]$, and let $>$ be a global monomial order on $\Bbbk[x_1, \dots, x_n]$. Suppose that, for some c, the following two conditions hold:*

1'. $\mathbf{L}(I)$ is generated by monomials in $\Bbbk[x_1, \dots, x_c]$.
2. $\mathbf{L}(I) \supset \langle x_1, \dots, x_c \rangle^m$ for some m.

Then the composition

$$R = \Bbbk[x_{c+1}, \dots, x_n] \subset \Bbbk[x_1, \dots, x_n] \to S = \Bbbk[x_1, \dots, x_n]/I$$

is a Noether normalization such that S is a free R-module (of finite rank).

Proof. Condition 1' implies condition 1 of Proposition 3.3.3. Thus, if conditions 1' and 2 hold, it is clear from Proposition 3.3.3 and its proof that there are only finitely many monomials $m_1, \dots, m_k$ in $\Bbbk[x_1, \dots, x_c]$ which are not contained in $\mathbf{L}(I)$, that the $\overline{m}_i = m_i + I$ generate S as an R-module, and that $R \to S$ is a Noether normalization.

We show that the $\overline{m}_i$ are R-linearly independent. For this, let $\sum_{i=1}^k f_i \overline{m}_i = 0 \in S$ be an R-relation which is zero. Then $f := \sum_{i=1}^k f_i m_i \in I$, so that $\mathbf{L}(f) \in \mathbf{L}(I)$. But $\mathbf{L}(f)$ is of type $\mathbf{L}(f) = m m_j$, for some term $m \in R = \Bbbk[x_{c+1}, \dots, x_n]$ and some j. Since $m_j \notin \mathbf{L}(I)$, condition 1' implies that $\mathbf{L}(f) = 0$ and, thus, that $f = 0$. Then all the f_i must be zero, as desired. □

Theorem-Definition 3.3.12 (Unmixedness Theorem). *Let I be a proper ideal of $\Bbbk[x_1, \dots, x_n]$. Suppose that, for some c, the composition*

$$R = \Bbbk[x_{c+1}, \dots, x_n] \subset \Bbbk[x_1, \dots, x_n] \to S = \Bbbk[x_1, \dots, x_n]/I$$

is a Noether normalization such that S is a free R-module (of finite rank). Then, for every associated prime $\mathfrak{p}$ of I, the dimension of $\mathrm{V}(\mathfrak{p}) \subset \mathbb{A}^n$ is $n - c$. In particular:

1. *I is **unmixed**, that is, I has no embedded components.*
2. *$\mathrm{V}(I) \subset \mathbb{A}^n$ is equidimensional of dimension $n - c$.*

Proof. Let $\mathfrak{p}$ be an associated prime of I. Then, by composing the natural map $S = \Bbbk[x_1, \dots, x_n]/I \to \Bbbk[x_1, \dots, x_n]/\mathfrak{p}$ with the Noether normalization $R \to S$, we get a homomorphism

$$\phi : R = \Bbbk[x_{c+1}, \dots, x_n] \to T = \Bbbk[x_1, \dots, x_n]/\mathfrak{p}$$

which exhibits T as a finitely generated R-module. To show that ϕ is injective (and, thus, that ϕ constitutes a Noether normalization), suppose the contrary. Then there is a nonzero polynomial $g \in \mathfrak{p} \cap \Bbbk[x_{c+1}, \dots, x_n]$. Since $\mathfrak{p} = I : f$ for some polynomial $f \in \Bbbk[x_1, \dots, x_n] \setminus I$ by the 1st Uniqueness Theorem 1.8.7 for primary decomposition, it follows that $gf \equiv 0 \mod I$, contradicting the fact that $S = \Bbbk[x_1, \dots, x_n]/I$ is free over R. We conclude that $\dim \mathrm{V}(\mathfrak{p}) = n - c$.

For statement 1, let $\mathfrak{p}_1 \subset \mathfrak{p}_2$ be two associated primes of I, and let $\overline{\mathfrak{p}}_1$ and $\overline{\mathfrak{p}}_2$ be their images in $S = \Bbbk[x_1, \dots, x_n]/I$. Then, by the argument above, $\overline{\mathfrak{p}}_1$ and $\overline{\mathfrak{p}}_2$ are both lying over the zero ideal of $R = \Bbbk[x_{c+1}, \dots, x_n]$. By part 2 of the lying over theorem, $\mathfrak{p}_1$ cannot be strictly contained in $\mathfrak{p}_2$. Hence, I has no embedded components.

Statement 2 is clear in the case where $\Bbbk = \mathbb{K}$ is algebraically closed since, then, the irreducible components of $\mathrm{V}(I) \subset \mathbb{A}^n$ are precisely the vanishing loci of the associated primes of I (there are only isolated components by statement 1). The result in the general case follows, once more, from Remark 2.7.1 on Buchberger's algorithm and field extensions: First, we may use Proposition 3.3.3 to check whether $R \to S$ is a Noether normalization (fix $>_{\mathrm{lex}}$ on $\Bbbk[x_1, \dots, x_n]$). If this is true, it is clear from Proposition 3.3.3 and its proof that there are only finitely many monomials $m_1, \dots, m_k$ in $\Bbbk[x_1, \dots, x_c]$ which are not contained in $\mathbf{L}(I)$, and that the $\overline{m}_i = m_i + I$ generate S as an R-module. The check whether the $\overline{m}_i$ are R-linearly independent amounts to computing that the elimination ideal $\langle m_1, \dots, m_k \rangle \cap I \cap \mathbb{k}[x_{c+1}, \dots, x_n]$ is zero, a task which can be dealt with using Gröbner bases. □

Remark 3.3.13. The importance of the unmixedness theorem is usually emphasized by calling an affine $\mathbb{k}$-algebra S **Cohen-Macaulay** if it admits a Noether normalization

$$R = \mathbb{k}[y_1, \dots, y_d] \subset S$$

such that S is a free R-module. We should point out that if S is free over R for one Noether normalization $R \subset S$, then the same is true for every Noether normalization of S. Moreover, the general definition of a Cohen-Macaulay ring given in other textbooks coincides in the case of affine rings with the definition given here. The key ingredient of the proof of these nontrivial facts is the theorem of Quillen and Suslin which, settling a conjecture of Serre (see Kunz (1985)), asserts that all finitely generated projective modules over $\mathbb{k}[x_1, \dots, x_n]$ are free.

We refer to Bruns and Herzog (1993), Matsumura (1986), and Eisenbud (1995) for some historical remarks on the name Cohen-Macaulay and for further reading on the topic of Cohen-Macaulay rings. In our book, the general definition of a Macaulay ring will be given in Definition 4.6.23, but we will not discuss this topic any further. □

Example 3.3.14. If I is the monomial ideal

$$I = \langle x_1^2, x_2^2, x_1x_2x_3 \rangle \subset \mathbb{k}[x_1, x_2, x_3],$$

then $R = \mathbb{k}[x_3] \to S = \mathbb{k}[x_1, x_2, x_3]/I$ is a Noether normalization by Proposition 3.3.3. In fact, S is generated over R by the residue classes $1, \overline{x}_1, \overline{x}_2, \overline{x}_1\overline{x}_2$. Hence, S is not a free R-module since $\overline{x}_3 \cdot (\overline{x}_1\overline{x}_2) = 0 \in S$. Accordingly, condition 1' of Theorem 3.3.11 is not fulfilled. □

Example 3.3.15. Considering, once more, the twisted cubic curve $C \subset \mathbb{A}^3$ and the reduced Gröbner basis

$$f_1 = x^2 - y, \quad f_2 = xy - z, \quad f_3 = y^2 - xz$$

for $\mathrm{I}(C)$ with respect to $>_{\mathrm{drlex}}$, we see that

$$\mathbb{k}[z] \subset \mathbb{k}[x, y, z]/\mathrm{I}(C) = \mathbb{k}[C]$$

is a Noether normalization such that $\mathbb{k}[C]$ is a free $\mathbb{k}[z]$-module of rank 3 ($1, \overline{x}, \overline{y}$ form a basis). In particular, C is indeed a curve in the sense that it is equidimensional of dimension 1. □

3.4 Krull Dimension

If $V_1 \subsetneq V_2 \subset \mathbb{A}^n$ are varieties, that is, if $\mathrm{I}(V_2) \subsetneq \mathrm{I}(V_1)$ are prime ideals, then $\dim V_1 < \dim V_2$ by lying over (argue as in the proof of statement 1 of the

Unmixedness Theorem 3.3.12). Taking our cue from this observation, we will be lead to yet another characterization of dimension.

We use the following notation. If R is a ring, a sequence

$$\mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \ldots \subsetneq \mathfrak{p}_d$$

of prime ideals of R with strict inclusions is called a **chain of prime ideals** of R. The number d of inclusions is called the **length** of the chain. The chain is said to be **maximal** if it cannot be extended to a chain of greater length by inserting an extra prime ideal.

Definition 3.4.1. Let R be a ring. The **Krull dimension** (or simply the **dimension**) **of $\boldsymbol{R}$**, written $\dim R$, is the supremum of the lengths of chains of prime ideals of R. If I is a proper ideal of R, the **dimension of $\boldsymbol{I}$**, written $\dim I$, is defined to be the dimension of R/I. □

By lying over and going up, we get:

Proposition 3.4.2. *If $R \subset S$ is an integral ring extension, then*

$$\dim R = \dim S.$$

□

In what follows, we show that the dimension of an ideal $I \subsetneq \mathbb{k}[x_1, \ldots, x_n]$ coincides with the dimension of its locus of zeros $\mathrm{V}(I) \subset \mathbb{A}^n$. Considering a Noether normalization for $\mathbb{k}[x_1, \ldots, x_n]/I$, and taking Proposition 3.4.2 above into account, this amounts to showing that the Krull dimension of a polynomial ring over $\mathbb{k}$ equals the number of its variables.

That the dimension of $\mathbb{k}[x_1, \ldots, x_n]$ is at least n is clear since

$$\langle 0 \rangle \subsetneq \langle x_1 \rangle \subsetneq \langle x_1, x_2 \rangle \subsetneq \ldots \subsetneq \langle x_1, \ldots, x_n \rangle$$

is a chain of prime ideals of length n. To show that there is no chain of greater length, we will proceed by induction on the number of variables, relying on the following result:

Theorem 3.4.3 (Noether Normalization, Refined Version). *Let S be an affine $\mathbb{k}$-algebra, and let $I \subsetneq S$ be an ideal. There exist integers $\delta \leq d$ and a Noether normalization*

$$\mathbb{k}[y_1, \ldots, y_d] \subset S$$

such that

$$I \cap \mathbb{k}[y_1, \ldots, y_d] = \langle y_1, \ldots, y_\delta \rangle.$$

Proof. Let $\mathbb{k}[x_1, \ldots, x_d] \subset S$ be any Noether normalization. Since the composition of two finite ring extensions is again finite, it is enough to find a Noether normalization $\mathbb{k}[y_1, \ldots, y_d] \subset \mathbb{k}[x_1, \ldots, x_d]$ such that $I \cap \mathbb{k}[y_1, \ldots, y_d] = \langle y_1, \ldots, y_\delta \rangle$ for some $\delta \leq d$. We may, thus, suppose that $S = \mathbb{k}[x_1, \ldots, x_d]$ is a polynomial ring.

In this case, if $I = \langle 0 \rangle$, there is nothing to show. If I is nonzero, by Lemma 3.1.3, we may choose the coordinates such that I contains a monic polynomial $f = x_1^k + c_1 x_1^{k-1} + \ldots + c_k$, with all $c_i \in \mathbb{k}[x_2, \ldots, x_d]$. Let $y_1 := f$. Then

$$\mathbb{k}[y_1, x_2, \ldots, x_d] \subset \mathbb{k}[x_1, \ldots, x_d]$$

is a finite ring extension since

$$x_1^k + c_1 x_1^{k-1} + \ldots + c_k - y_1 = 0$$

is an integral equation for x_1 over $\mathbb{k}[y_1, x_2, \ldots, x_d]$. On the other hand, by induction on d, we may suppose that there is a Noether normalization $\mathbb{k}[y_2, \ldots, y_d] \subset \mathbb{k}[x_2, \ldots, x_d]$ such that $I \cap \mathbb{k}[y_2, \ldots, y_d] = \langle y_2, \ldots, y_\delta \rangle$ for some $\delta \leq d$. Then the composition

$$\mathbb{k}[y_1, \ldots, y_d] \subset \mathbb{k}[y_1, x_2, \ldots, x_d] \subset \mathbb{k}[x_1, x_2, \ldots, x_d]$$

is a finite ring extension as well. Moreover, $y_1, \ldots, y_d$ are algebraically independent over $\mathbb{k}$ since, otherwise, the transcendence degree of $\mathbb{k}(x_1, \ldots, x_d)$ over $\mathbb{k}$ would be smaller than d, contradicting the algebraic independence of the x_i. Finally, since every polynomial $f \in I \cap \mathbb{k}[y_1, \ldots, y_d]$ can be written as a sum $f = gy_1 + h$, where $g \in \mathbb{k}[y_1, \ldots, y_d]$ and $h \in I \cap \mathbb{k}[y_2, \ldots, y_d] = \langle y_2, \ldots, y_\delta \rangle$, we conclude that $I \cap \mathbb{k}[y_1, \ldots, y_d] = \langle y_1, \ldots, y_\delta \rangle$. This shows that the desired Noether normalization exists. □

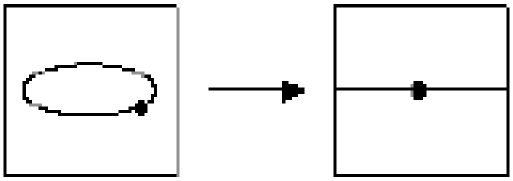
The geometric interpretation of the theorem is as follows: Given an algebraic set $A \subset \mathbb{A}^n$ together with a subvariety $B \subset A$, there is a surjective map $\pi : A \to \mathbb{A}^d$ with finite fibers which maps B onto a linear subspace of $\mathbb{A}^d$.

Exercise 3.4.4. Let $I \subset S = \mathbb{k}[x_1, \ldots, x_4]$ be the ideal which is generated by the 2×2 minors of the matrix

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ x_2 & x_3 & x_4 \end{pmatrix}.$$

Find a Noether normalization as in Theorem 3.4.3. □

Exercise 3.4.5. Formulate and prove a refined version of Noether normalization involving chains of ideals $I_1 \subset \cdots \subset I_m \subset S$.



□

Theorem 3.4.6. *The polynomial ring $\Bbbk[x_1, \dots, x_n]$ has Krull dimension n. In fact, every maximal chain of prime ideals of $\Bbbk[x_1, \dots, x_n]$ has length n.*

Proof. Since every chain of prime ideals of the Noetherian ring $\Bbbk[x_1, \dots, x_n]$ can be extended to a maximal chain of prime ideals, it suffices to prove the second assertion. That is, given a maximal chain

$$\mathfrak{P}_0 \subsetneq \mathfrak{P}_1 \subsetneq \dots \subsetneq \mathfrak{P}_m \tag{3.2}$$

of prime ideals of $\Bbbk[x_1, \dots, x_n]$, we must show that $m = n$. We proceed in three steps.

Step 1. To begin, $\mathfrak{P}_0 = \langle 0 \rangle$ since $\Bbbk[x_1, \dots, x_n]$ is an integral domain. Furthermore, $\mathfrak{P}_m$ is a maximal ideal. n particular, $m \geq 1$. Applying Theorem 3.4.3 to $\mathfrak{P}_1$, we get a Noether normalization $\Bbbk[y_1, \dots, y_n] \subset \Bbbk[x_1, \dots, x_n]$ such that $\mathfrak{P}_1 \cap \Bbbk[y_1, \dots, y_n] = \langle y_1, \dots, y_\delta \rangle$ for some $\delta \leq n$. Then $\delta = 1$ since, otherwise, going-down would yield a prime ideal $\mathfrak{Q} \subset \Bbbk[x_1, \dots, x_n]$ lying over $\langle y_1, \dots, y_{\delta-1} \rangle$, and such that $\mathfrak{P}_0 = \langle 0 \rangle \subsetneq \mathfrak{Q} \subsetneq \mathfrak{P}_1$.

Writing $\mathfrak{p}_i = \mathfrak{P}_i \cap \Bbbk[y_1, \dots, y_n]$ for all i, we get a chain

$$\langle 0 \rangle = \mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \cdots \subsetneq \mathfrak{p}_m \tag{3.3}$$

of prime ideals of $\Bbbk[y_1, \dots, y_n]$ (all inclusions are strict by part 2 of the lying over theorem). We show that this chain is maximal. Suppose, to the contrary, that there is a prime ideal $\mathfrak{q} \subset \Bbbk[y_1, \dots, y_n]$ with strict inclusions $\mathfrak{p}_i \subsetneq \mathfrak{q} \subsetneq \mathfrak{p}_{i+1}$ for some i. Then $1 \leq i \leq m-1$ since $\mathfrak{p}_0 = \langle 0 \rangle$, and since $\mathfrak{p}_m$ is maximal by part 3 of the lying over theorem. Applying Theorem 3.4.3 to $\mathfrak{p}_i$, we get a Noether normalization $\Bbbk[z_1, \dots, z_n] \subset \Bbbk[y_1, \dots, y_n]$ such that $\mathfrak{p}_i \cap \Bbbk[z_1, \dots, z_n] = \langle z_1, \dots, z_{\delta'} \rangle$ for some $\delta' \leq n$. The composition

$$\Bbbk[z_{\delta'+1}, \dots, z_n] \subset \Bbbk[y_1, \dots, y_n] \to \Bbbk[y_1, \dots, y_n]/\mathfrak{p}_i$$

is a Noether normalization as well, and we have strict inclusions

$$\langle 0 \rangle \subsetneq (\mathfrak{q}/\mathfrak{p}_i) \cap \Bbbk[z_{\delta'+1}, \dots, z_n] \subsetneq (\mathfrak{p}_{i+1}/\mathfrak{p}_i) \cap \Bbbk[z_{\delta'+1}, \dots, z_n].$$

Since $\Bbbk[z_{\delta'+1}, \dots, z_n] \subset \Bbbk[x_1, \dots, x_n]/\mathfrak{P}_i$ is also a Noether normalization, we see by going down that we may insert a prime ideal between $\langle 0 \rangle$ and $\mathfrak{P}_{i+1}/\mathfrak{P}_i$ in $\Bbbk[x_1, \dots, x_n]/\mathfrak{P}_i$ and, thus, also between $\mathfrak{P}_i$ and $\mathfrak{P}_{i+1}$ in $\Bbbk[x_1, \dots, x_n]$. This contradicts the maximality of (3.2). We conclude that (3.3) is maximal, too.

Step 3. The maximal chain (3.3) corresponds to a maximal chain of prime ideals of $\Bbbk[y_1, \dots, y_n]/\mathfrak{p}_1 = \Bbbk[y_1, \dots, y_n]/\langle y_1 \rangle \cong \Bbbk[y_2, \dots, y_n]$ of length $m-1$. Thus, we are done by induction on the number of variables. □

Corollary 3.4.7. *If R is an affine domain over $\Bbbk$, then*

$$\dim R = \operatorname{trdeg}_{\Bbbk} \mathrm{Q}(R).$$

This is the common length of all maximal chains of prime ideals of R.

Proof. Let $\Bbbk[y_1, \dots, y_d] \subset R$ be a Noether normalization. Then $\dim R = \dim \Bbbk[y_1, \dots, y_d] = d$ by Proposition 3.4.2 and Theorem 3.4.6. Since also $\operatorname{trdeg}_{\Bbbk} \mathrm{Q}(R) = \operatorname{trdeg}_{\Bbbk} \Bbbk(y_1, \dots, y_d) = d$, we must have $\dim R = \operatorname{trdeg}_{\Bbbk} \mathrm{Q}(R)$.

Write, now, R as the quotient of a polynomial ring $\Bbbk[x_1, \dots, x_n]$ by a prime ideal $\mathfrak{q}$, and fix a chain $\mathfrak{q}_0 = \langle 0 \rangle \subsetneq \cdots \subsetneq \mathfrak{q}_c = \mathfrak{q}$ of prime ideals which cannot be extended to a longer chain of prime ideals with largest ideal $\mathfrak{q}$. The fixed chain and the preimage of any given maximal chain of prime ideals of R fit together to a maximal chain of prime ideals of $\Bbbk[x_1, \dots, x_n]$ which necessarily has length n by Theorem 3.4.6. From this, the result follows. □

Definition 3.4.8. Let R be a ring, and let $I \subsetneq R$ be an ideal. The **codimension of I**, written $\operatorname{codim} I$, is defined as follows. If $I = \mathfrak{p}$ is a prime ideal, its codimension is the supremum of the lengths of all chains of prime ideals of R with largest prime ideal $\mathfrak{p}$. If I is arbitrary, its codimension is the minimum of the codimensions of the prime ideals containing I. □

Corollary 3.4.9. *If R is an affine domain over $\Bbbk$, and $I \subsetneq R$ is an ideal, then*

$$\dim I + \operatorname{codim} I = \dim R.$$

Proof. The assertion is a consequence of the preceeding corollary since $\dim I$ can be expressed in terms of a maximal chain of prime ideals of R which includes a prime ideal $\mathfrak{p} \supset I$ such that $\operatorname{codim} I = \operatorname{codim} \mathfrak{p}$. □

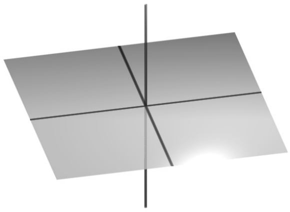
From the proof, we see that if I is a proper ideal of an arbitrary ring R, then

$$\dim I + \operatorname{codim} I \leq \dim R.$$

The following example shows, however, that in rings other than affine domains, equality does not necessarily hold:

Example 3.4.10. Let $R = \Bbbk[x, y, z]/\langle xz, yz \rangle$ be the coordinate ring of the union of the xy-plane and the z-axis, and let $\mathfrak{P} = \langle \overline{x}, \overline{y}, \overline{z} - 1 \rangle$ be the maximal ideal of R corresponding to the point $p = (0, 0, 1)$ on the z-axis. Then

$$\operatorname{codim} \mathfrak{P} + \dim \mathfrak{P} = 1 + 0 \neq 2 = \dim R.$$



Observe that R contains maximal chains of prime ideals of different length. □

The notion of codimension originates from the geometric setting. If $\emptyset \neq A \subset \mathbb{A}^n$ is an algebraic set, and $B \subset A$ is an algebraic subset, the **codimension of *B* in *A***, written $\operatorname{codim}_A B$, is defined as follows. If B is nonempty, rewrite Definition 3.4.8 in terms of subvarieties of A. Equivalently, $\operatorname{codim}_A B = \operatorname{codim} \mathrm{I}_A(B)$. If B is the empty subset of A, by convention, $\operatorname{codim}_A B = \infty$. The analogue of Corollary 3.4.9 holds accordingly. In the situation of Example 3.4.10, the codimension of a point $q \in \mathrm{V}(xz, yz)$ is 1 if q lies on the z-axis, and 2, otherwise.

Remark 3.4.11. The notion of Krull dimension extends the concept of dimension from affine algebraic sets, that is, from affine rings, to arbitrary rings (commutative, and with a multiplicative identity). For instance, we can, thus, assign a dimension to the ring of integers:

$$\dim \mathbb{Z} = 1.$$

Indeed, each nonzero prime ideal of $\mathbb{Z}$ is a principal ideal generated by a prime number and, thus, a maximal ideal. More generally, every principal ideal domain which is not a field has Krull dimension 1. □

In developing some intuitive understanding of Krull dimension, the beginner may face a couple of surprises. For example, it turns out that even Noetherian rings may have infinite dimension (see Nagata (1962), Appendix A1. Examples of bad Noetherian rings).

3.5 Reduction to Hypersurfaces

Our goal in this section is to show that every affine variety is birationally equivalent to a hypersurface in some affine space. In fact, we prove a somewhat stronger result which is based on a field theoretic version of Noether normalization.

Proposition 3.5.1 (Noether Normalization and Separability). *Let S be an affine domain over the algebraically closed field $\mathbb{K}$, and let L be the quotient field of S. Then there are $y_1, \dots, y_d \in S$ such that:*

1. *$\mathbb{K}[y_1, \dots, y_d] \subset S$ is a Noether normalization.*
2. *$\mathbb{K}(y_1, \dots, y_d) \subset L$ is a separable field extension.*

Proof. In characteristic zero, every field extension is separable. We suppose, therefore, that $\operatorname{char} \mathbb{K} = p > 0$.

Let $S = \mathbb{K}[x_1, \dots, x_n]/\mathfrak{p}$ for some prime ideal $\mathfrak{p}$ of some polynomial ring $\mathbb{K}[x_1, \dots, x_n]$. If $\mathfrak{p} = \langle 0 \rangle$, there is nothing to prove. If $\mathfrak{p}$ is nonzero, it contains an irreducible polynomial f. For each i, considering f as a polynomial in x_i, with coefficients in $\mathbb{K}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$, we either have that f is separable in x_i or that the formal derivative of f with respect to x_i is zero. In the latter case, $f \in \mathbb{K}[x_1, \dots, x_i^p, \dots, x_n]$ by Exercise 1.1.3.

Suppose that f is inseparable in each x_i. Then $f \in \mathbb{K}[x_1^p, \dots, x_n^p]$. Since $\mathbb{K}$ is algebraically closed, it contains a pth root of every coefficient of f. Using the characteristic p identity $(a+b)^p = a^p + b^p$, we see that f has a pth root in $\mathbb{K}[x_1, \dots, x_n]$. That is, there is a polynomial $g \in \Bbbk[x_1, \dots, x_n]$ such that $g^p = f$:

$$f = \sum a_\alpha x_1^{p\alpha_1} \cdots x_n^{p\alpha_n} = (\sum b_\alpha x^\alpha)^p, \quad \text{where } \; b_\alpha^p = a_\alpha.$$

This contradicts the irreducibility of f.

So f is separable in at least one of the x_i, say in x_1. Choosing $a_2, \dots, a_n \in \mathbb{K}$ sufficiently general as in the proof of Lemma 3.1.3, and expanding $f(x_1, \widetilde{x}_2 + a_2 x_1, \dots, \widetilde{x}_n + a_n x_1)$, we get a polynomial

$$a x_1^e + c_1(\widetilde{x}_2, \dots, \widetilde{x}_n) x_1^{e-1} + \ldots + c_e(\widetilde{x}_2, \dots, \widetilde{x}_n)$$

which provides both an integral equation for $\overline{x}_1 \in S$ over $\mathbb{K}[\widetilde{x}_2, \dots, \widetilde{x}_n]$ and a separable equation for $\overline{x}_1 \in L$ over $\mathbb{K}(\widetilde{x}_2, \dots, \widetilde{x}_n)$. The result follows as in the proof of Theorem 3.3.1 since the composition of any sequence of separable field extensions is separable. □

Theorem 3.5.2 (Reduction to Hypersurfaces). *For any affine variety V of dimension d, there exists a finite morphism $V \to W$ onto a hypersurface $W \subset \mathbb{A}^{d+1}$ which is a birational equivalence of V with W.*

Proof. Let $\overline{x}_1, \dots, \overline{x}_n$ be generators for $S = \mathbb{K}[V]$ as a $\mathbb{K}$-algebra, and choose $y_1, \dots, y_d \in \mathbb{K}[V]$ as in Proposition 3.5.1 above. Then $\mathbb{K}[V]$ is a finite $\mathbb{K}[y_1, \dots, y_d]$-algebra, and $L = \mathbb{K}(V)$ is a finite separable field extension of $\mathbb{K}(y_1, \dots, y_d)$ which is generated by $\overline{x}_1, \dots, \overline{x}_n$. By the primitive element theorem from Galois theory (see, for instance, Dummit and Foote (2003), Section 14.4), we can find a $\mathbb{K}(y_1, \dots, y_d)$-linear combination y_{d+1} of the $\overline{x}_i$ such that $\mathbb{K}(V)$ is generated by y_{d+1} over $\mathbb{K}(y_1, \dots, y_d)$. Clearing denominators, y_{d+1} can be taken as a $\mathbb{K}[y_1, \dots, y_d]$-linear combination of the $\overline{x}_i$ and, thus, as an element of $\mathbb{K}[V]$.

If $f(y_1, \dots, y_d, y_{d+1}) = 0$ is an integral equation for y_{d+1} over $\Bbbk[y_1, \dots, y_d]$ of minimal degree, then f is an irreducible polynomial in $d+1$ variables which, considered as a univariate polynomial with coefficients in $\mathbb{K}[y_1, \dots, y_d]$, is the minimal polynomial of $\mathbb{K}(V)$ over $\Bbbk(y_1, \dots, y_d)$. Hence, f defines an irreducible hypersurface $W \subset \mathbb{A}^{d+1}$, and the finite ring inclusion $\phi : \Bbbk[W] = \mathbb{K}[y_1, \dots . y_d, y_{d+1}] \to \mathbb{K}[V]$ extends to an isomorphism $\mathbb{K}(W) \to \Bbbk(V)$ of rational function fields. It follows, that the morphism $V \to W$ induced by ϕ is both finite and a birational equivalence of V with W. □

If $\Bbbk$ is a field of definition of V, the arguments given in the two proofs above actually show that in characteristic zero, W and the morphism $V \to W$ can be chosen to be defined over $\Bbbk$, too. In positive characteristic, we might need a finite field extension.

3.6 Additional exercises

Chapter 4

Local Properties

In the preceeding chapters, we developed the geometry-algebra dictionary from a global point of view, focusing on geometric questions which concern a given algebraic set A as a whole. Accordingly, we studied functions defined on all of A, the polynomial functions on A, and used the ring $\Bbbk[A]$ formed by these functions to express geometric properties of A in ring theoretic terms. Algorithmically, we computed Gröbner bases with respect to what we called global monomial orders.

In this chapter, we will be interested in geometric properties which are local in the sense that they reflect the behavior of A near a given point $p \in A$. In defining the basic local property, which is smoothness, we will rely on the concept of the tangent space. Intuitively, p is a smooth point of A if the tangent space T_pA approximates A near p (otherwise, we will say that p is a singular point of A). Here, we will define T_pA over any field in a purely algebraic way (no limiting process as in calculus is needed). We will show that the singular points form an algebraic subset of A, and we will prove the Jacobian criterion which, in many cases of interest, allows one to compute the equations of this subset, and to check whether the given polynomials defining A actually generate a radical ideal.

We will, then, describe the construction of the local ring $\mathcal{O}_{A,p}$ whose elements are germs of functions defined on Zariski open neighborhoods of p in A. It will turn out that A is smooth at p iff $\mathcal{O}_{A,p}$ is a regular local ring. Focusing on the general and purely algebraic nature of the construction of $\mathcal{O}_{A,p}$, we will be lead to the concept of localization which plays an important role in commutative algebra. In fact, localization often allows one to reduce problems concerning arbitrary rings to problems concerning local rings which are much easier. One reason why local rings are easier to handle than arbitrary rings is Nakayama's lemma. As a typical application of this lemma, we prove a special case of Krull's intersection theorem.

Returning to more geometric questions, we will use the local ring $\mathcal{O}_{\mathbb{A}^2,p}$ to define the intersection multiplicity of two plane curves at a point $p \in \mathbb{A}^2$.

Making, thus, preperations for the treatment of Bezout's theorem in Chapter 5, we will verify a number of properties of intersection multiplicities.

Algorithmically, the computation of the multiplicities is based on a version of Buchberger's algorithm for computing Gröbner bases with respect to what we will call local monomial orders.

Motivated by rationality problems which may arise in such computations, we will give an alternative definition of the multiplicities using the notion of modules of finite length. Discussing this notion, we will show that a ring R has finite length iff it is Artinian, that is, R satisfies the *descending* chain condition. Applying this fact in a localized situation (which will allow us to benefit from Nakayama'a lemma), we will prove Krull's principal ideal theorem.

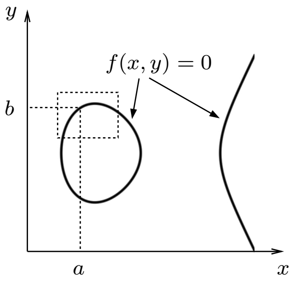
In the final section, we will treat the completion $\widehat{\mathcal{O}_{A,p}}$ of $\mathcal{O}_{A,p}$. This will help us to overcome a drawback of $\mathcal{O}_{A,p}$ which is due to the fact that Zariski open sets are rather large. Since $\mathcal{O}_{A,p}$ consists of (germs of) functions defined on such sets, it carries information on too much of A. In contrast, the larger ring $\widehat{\mathcal{O}_{A,p}}$ carries far more local information. Another topic, which we will treat briefly, is the tangent cone TC_pA which approximates A near p even if p is a singular point of A.

4.1 Smoothness

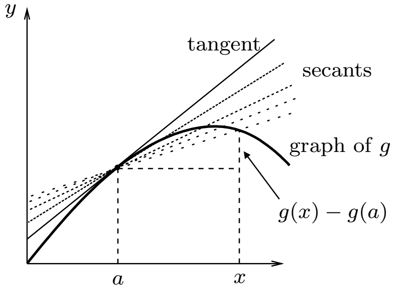
We will define smoothness such that in case $\mathbb{K} = \mathbb{C}$, an algebraic set $A \subset \mathbb{A}^n$ is smooth at a point $p \in A$ iff A is a complex submanifold of $\mathbb{A}^n$ in an Euclidean neighborhood of p. Equivalently, we will require that the hypothesis of the implicit function theorem is fulfilled. In making this precise, we will first study the hypersurface case, which is intuitively easy to understand, and where important consequences of the definition are easy to prove.

We fix our ideas by illustrating the special case of a plane curve. Let $f \in \mathbb{C}[x, y]$ be a nonconstant square-free polynomial, let $C = \mathrm{V}(f) \subset \mathbb{A}^2(\mathbb{C})$ be the corresponding curve, and let $p = (a, b) \in C$ be a point. In this situation, the complex variable version of the implicit function theorem asserts that if the gradient $\left(\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p)\right)$ is nonzero, then there is an Euclidean neighborhood of p in which C can be exhibited as the graph of a holomorphic function. Supposing, say, that $\frac{\partial f}{\partial y}(p) \neq 0$, the precise statement is that there are open neighbourhoods U_1 of a and U_2 of b in the Euclidean topology and a holomorphic function $g : U_1 \to U_2$ such that $g(a) = b$ and

$$C \cap (U_1 \times U_2) = \{(x, g(x)) \mid x \in U_1\}.$$



Reflecting this fact, we get a well defined tangent line to C at p (the linear approximation of C near p) by interpreting the existence of the differential quotient of g at $x=a$ geometrically – the tangent line is the limiting position of secant lines to C passing through p:



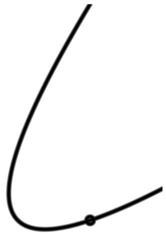
Since

$$g'(a) = -\frac{\partial f}{\partial x}(p) / \frac{\partial f}{\partial y}(p)$$

by the chain rule, we may rewrite the equation $y = b + g'(a)(x-a)$ of the tangent line in terms of f:

$$\frac{\partial f}{\partial x}(p)(x-a) + \frac{\partial f}{\partial y}(p)(y-b) = 0. \tag{4.1}$$

There is no algebraic geometry analogue of the implicit function theorem: Even though we are concerned with a *polynomial* f in our considerations, it is usually not possible to choose the U_i as neighborhoods in the Zariski topology and g as a polynomial function. From a topological point of view, as illustrated by the example in the following picture, the Zariski open sets are simply too big:



On the other hand, using formal partial derivatives, equation (4.1) makes sense even in case $\mathbb{K} \neq \mathbb{C}$. We, therefore, define:

Remark-Definition 4.1.1. 1. If $f \in \mathbb{K}[x_1, \dots, x_n]$ is a polynomial, and $p = (a_1, \dots, a_n) \in \mathbb{A}^n$ is a point, the **differential of *f* at *p***, written $d_p f$, is defined to be

$$d_p f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)(x_i - a_i) \in \mathbb{K}[x_1, \dots, x_n].$$

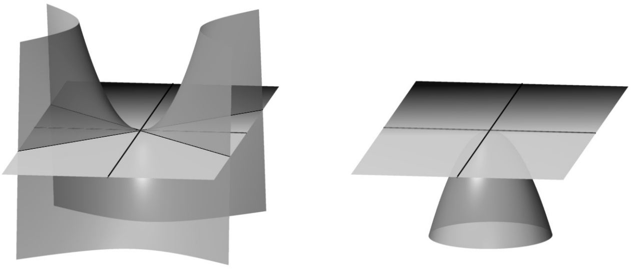
That is, $d_p f$ is the linear part of the Taylor expansion of f at p:

$$f = f(p) + d_p f + \text{terms of degree} \geq 2 \text{ in the } x_i - a_i.$$

2. Let $A \subset \mathbb{A}^n$ be a hypersurface, let $p \in A$ be a point, and let $f \in \mathbb{K}[x_1, \dots, x_n]$ be a generator for $\mathrm{I}(A)$. Then the **tangent space to *A* at *p***, denoted $T_p A$, is the linear subvariety

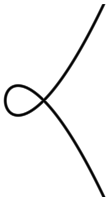
$$T_p A = \mathrm{V}(d_p f) \subset \mathbb{A}^n.$$

We say that p is a **smooth** (or a **nonsingular**) **point** of A if $T_p A$ is a hyperplane, that is, if $d_p f$ is nonzero.

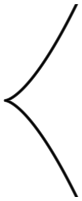


Otherwise, $T_p A = \mathbb{A}^n$, and we call p a **singular point** of A. □

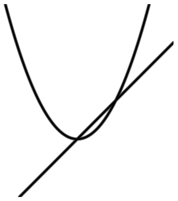
Example 4.1.2. The origin $o = (0, 0) \in \mathbb{A}^2(\mathbb{C})$ is a singular point of each cubic curve shown below:



$y^2 = x^3 + x^2$



$y^2 = x^3$



$y^2 = xy + x^2y - x^3$ □

The tangent space T_pA is the union of all lines meeting A with multiplicity at least 2 at p:

Proposition 4.1.3. *Let $A \subset \mathbb{A}^n$ be a hypersurface, and let $\mathrm{I}(A) = \langle f \rangle$.*

1. *Let $p = (a_1, \dots, a_n) \in A$ be a point, and let $L \subset \mathbb{A}^n$ be a line through p, given by the parametric equations $x_i = a_i + tv_i$, $i = 1, \dots, n$, where $v = (v_1, \dots, v_n) \in \mathbb{A}^n$ is a direction vector of L. Then $L \subset T_pA$ iff the polynomial $F(t) := f(p + tv) \in \mathbb{K}[t]$ vanishes with multiplicity ≥ 2 at 0.*
2. *The set A_{sing} of singular points of A is a proper algebraic subset of A:*

$$A_{\mathrm{sing}} = \mathrm{V}(f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}) \subsetneq A.$$

Proof. 1. The result follows from the chain rule: $\frac{\partial F}{\partial t}(0) = \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}(p)$.

2. That $A_{\mathrm{sing}} = \mathrm{V}(f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})$ is clear from our definitions. In particular, A_{sing} is an algebraic subset of A. To show that A_{sing} is properly contained in A, suppose the contrary. Then, for all i, the partial derivative $\frac{\partial f}{\partial x_i}$ is contained in $\langle f \rangle$, so that $\frac{\partial f}{\partial x_i} = 0$ by degree reasoning. If $\mathrm{char}\, \mathbb{K} = 0$, this implies that f is constant, contradicting our assumption that A is a hypersurface. If $\mathrm{char}\, \mathbb{K} = p > 0$, we must have $f \in \mathbb{K}[x_1^p, \dots, x_n^p]$ (see Exercise 1.1.3). As in the proof of Proposition 3.5.1, we conclude that f has a pth root in $\mathbb{K}[x_1, \dots, x_n]$. This contradicts the fact that $\mathrm{I}(A) = \langle f \rangle$ is a radical ideal. □

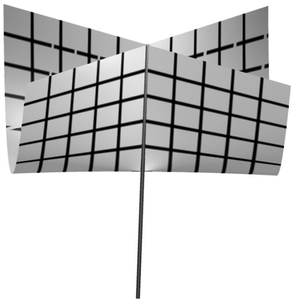
Example 4.1.4. The set of singular points of the Whitney umbrella

$$\mathrm{V}(x^2 - y^2 z) \subset \mathbb{A}^3(\mathbb{C})$$

is the z-axis

$$\mathrm{V}(x^2 - y^2 z, 2x, -2yz, -y^2) = \mathrm{V}(x, y).$$

We show a real picture:



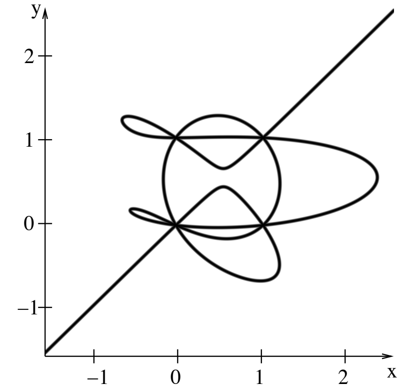
□

Exercise 4.1.5. 1. Find all singular points of the curve

$$\mathrm{V}(x^2 - 2x^3 + x^4 + y^2 - 2y^3 + y^4 - \frac{3}{2}x^2y^2) \subset \mathbb{A}^2(\mathbb{C}).$$

Draw a picture of the real points of this curve.

2. Find all singular points of the curve $\mathrm{V}(f) \subset \mathbb{A}^2(\mathbb{C})$, where f is the degree-7 polynomial considered in Example 1.2.4, part 3.



□

We, now, turn from hypersurfaces to arbitary algebraic sets:

Definition 4.1.6. Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p \in A$ be a point. The **tangent space** to A at p, denoted T_pA, is the linear subvariety

$$T_pA = \mathrm{V}(d_pf \mid f \in \mathrm{I}(A)) \subset \mathbb{A}^n.$$

□

As in Proposition 4.1.3, a line $L = \{p + tv \mid t \in \mathbb{K}\}$ is contained in T_pA iff all polynomials $f(p + tv) \in \mathbb{K}[t]$, $f \in \mathrm{I}(A)$, vanish with multiplicity ≥ 2 at 0.

Remark 4.1.7. 1. In defining the tangent space, it suffices to consider a set of generators for the vanishing ideal of A: if $\mathrm{I}(A) = \langle f_1, \ldots, f_r \rangle$, then

$$T_pA = \mathrm{V}(d_pf_i \mid i = 1, \ldots, r) \subset \mathbb{A}^n.$$

In particular,

$$\dim_{\mathbb{K}} T_pA = n - \operatorname{rank}\left(\frac{\partial f_i}{\partial x_j}(p)\right).$$

2. The function

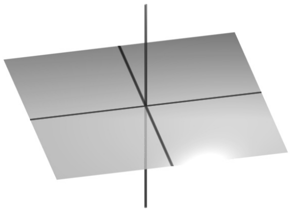
$$A \to \mathbb{N},\ p \mapsto \dim T_pA,$$

is upper semicontinous in the Zariski topology on A. That is, for any integer k, the subset

$$\{p \in A \mid \dim_{\mathbb{K}} T_pA \geq k\} \subset A$$

is Zariski closed. Indeed, this subset is the intersection of A with the locus of zeros of the $(n-k+1) \times (n-k+1)$ minors of the **Jacobian matrix** $\left(\frac{\partial f_i}{\partial x_j}\right)$. □

Example 4.1.8. Let $A = \mathrm{V}(xz, yz) = \mathrm{V}(x, y) \cup \mathrm{V}(z) =: L \cup P \subset \mathbb{A}^3$ be the union of the z-axis and the xy-plane:



If $o = (0,0,0) \in \mathbb{A}^3$ is the origin, and $p \in A$ is any point, then $\dim T_pA = 1$ if $p \in L \setminus \{o\}$, $\dim T_pA = 2$ if $p \in P \setminus \{o\}$, and $\dim T_pA = 3$ if $p = o$. □

According to our definition, a hypersurface $A \subset \mathbb{A}^n$ is smooth at a point $p \in A$ if the dimension of A equals the dimension of the tangent space T_pA. In extending this definition to an arbitrary algebraic set A, we have to take into account that, in contrast to the hypersurface case, A may have irreducible components of different dimension. On the other hand, the behavior of A near $p \in A$ is only effected by those components passing through p.

Definition 4.1.9. Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p \in A$ be a point. The **local dimension of *A* at *p***, written $\dim_p A$, is the maximum dimension of an irreducible component of A containing p. □

We always have

$$\dim_{\mathbb{K}} T_pA \geq \dim_p A. \tag{4.2}$$

In contrast to the hypersurface case, however, the result for arbitrary algebraic sets is not immediately clear from the definitions. We will prove it in a more general algebraic setting in Corollary 4.6.20 as a consequence of Krull's principal ideal theorem.

Definition 4.1.10. Let $A \subset \mathbb{A}^n$ be algebraic.

1. We say that $A \subset \mathbb{A}^n$ is **smooth** (or **nonsingular**) **at *p*** $\in A$ if

$$\dim_{\mathbb{K}} T_pA = \dim_p A.$$

We, then, refer to p as a **smooth** (or a **nonsingular**) **point** of A. Otherwise, we say that A is **singular at *p***, that p is a **singular point** of A, or that p is a **singularity** of A.

2. The set A_{sing} of singular points of A is called the **singular locus** of A. If A_{sing} is empty, that is, if A is smooth at each of its points, then A is called **smooth**. □

Remark 4.1.11. Let $A \subset \mathbb{A}^n$ be an algebraic set.

1. If A is smooth at p, then p is contained in a single component of A. In fact, if $A = V_1 \cup \dots \cup V_s$ is the decomposition of A into its irreducible components, then

$$A_{\text{sing}} = \bigcup_{i \neq j} (V_i \cap V_j) \cup \bigcup_i (V_i)_{\text{sing}}$$

(we will establish this in Corollary 4.6.26). In particular, A_{sing} is an algebraic subset of A since this is true in the case where A is irreducible. Indeed, in this case, $\dim_p A = \dim A$ for all $p \in A$, and we may apply part 2 of Remark 4.1.7, with $k = \dim A + 1$.

2. The singular locus A_{sing} and A have no irreducible component in common. That is, for any irreducible component V_i of A, we have $A_{\text{sing}} \cap V_i \subsetneq V_i$. Using Theorem 3.5.2 and the formula in part 1 above, we will deduce this fact in Corollary 4.2.16 from the hypersurface case. □

If generators $f_1, \dots, f_r$ for the vanishing ideal $\mathrm{I}(A)$ are given, and the local dimension $\dim_p A$ is known to us, we can decide whether A is smooth at p by computing $\dim_{\mathbb{K}} T_p A = n - \operatorname{rank}\left(\frac{\partial f_i}{\partial x_j}(p)\right)$, and comparing this number with $\dim_p A$. The Jacobian criterion, which we treat next, often allows one to test smoothness without having to check a priori that the given polynomials $f_1, \dots, f_r$ defining A actually generate $\mathrm{I}(A)$. In fact, under the assumptions of the corollary to the Jacobian criterion stated below, this will follow a posteriori. In this way, the corollary gives a powerful method for establishing that $f_1, \dots, f_r$ generate a radical ideal.

Theorem 4.1.12 (Jacobian Criterion). *Let $A \subset \mathbb{A}^n$ be an algebraic subset, let $p \in A$ a point, and let $f_1, \dots, f_r \in \mathrm{I}(A)$. Then*

$$n - \operatorname{rank}\left(\frac{\partial f_i}{\partial x_j}(p)\right) \geq \dim_p A.$$

If equality holds, then A is smooth at p.

Proof. This follows from the chain of inequalities

$$n - \operatorname{rank}\left(\frac{\partial f_i}{\partial x_j}(p)\right) \geq \dim_{\mathbb{K}} T_p A \geq \dim_p A.$$

□

Corollary 4.1.13. *Let $I = \langle f_1, \dots, f_r \rangle \subset \Bbbk[x_1, \dots, x_n]$ be an ideal such that $A = \mathrm{V}(I) \subset \mathbb{A}^n$ is equidimensional of dimension d, and let $I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right)$ denote the ideal generated by the $(n-d) \times (n-d)$ minors of the Jacobian matrix of the f_i. If $I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right) + I = \langle 1 \rangle$, then A is smooth and $I\,\mathbb{K}[x_1, \dots x_n] = \mathrm{I}(A)$. In particular, I is a radical ideal.*

Proof. The subset $\mathrm{V}(I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right)+I)=\{p\in A\mid n-\mathrm{rank}\left(\frac{\partial f_i}{\partial x_j}(p)\right)>d\}\subset A$ is empty by the assumption on $I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right)+I$ and Hilbert's Nullstellensatz. Since each irreducible component of A has dimension d, the Jacobian criterion implies that A is smooth. That $I\,\mathbb{K}[x_1,\dots x_n]=\mathrm{I}(A)$ will be established towards the end of Section 4.6. □

Under a stronger assumption, the Jacobian criterion can also be applied if $1\notin I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right)+I$:

Corollary 4.1.14. *Let $I=\langle f_1,\dots,f_r\rangle\subset\Bbbk[x_1,\dots,x_n]$ be an ideal of dimension d, and let $A=\mathrm{V}(I)\subset\mathbb{A}^n$. Suppose that $\Bbbk[x_1,\dots,x_n]/I$ is Cohen-Macaulay (by the Unmixedness Theorem 3.3.12, this implies that A is equidimensional of dimension d). With notation as in Corollary 4.1.13, if*

$$\dim\mathrm{V}(I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right)+I)<\dim\mathrm{V}(I)=d,$$

then $I\,\mathbb{K}[x_1,\dots x_n]=\mathrm{I}(A)$ and $\mathrm{V}(I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right)+I)=A_{\mathrm{sing}}$. In particular, I is a radical ideal.

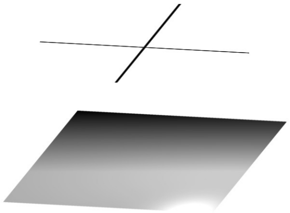
Proof. This will also be established towards the end of Section 4.6. □

The following example shows that the assumption of equidimensionality in Corollary 4.1.13 is really needed:

Example 4.1.15. Let $I=\langle f_1,f_2\rangle\subset\Bbbk[x_1,x_2,x_3]$ be the ideal generated by $f_1=x_1^2-x_1$ and $f_2=x_1x_2x_3$. Buchberger's criterion shows that f_1,f_2 form a lexicographic Gröbner basis for I. By Proposition 3.3.3, the composition $\Bbbk[x_2,x_3]\subset\Bbbk[x_1,x_2,x_3]\to\Bbbk[x_1,x_2,x_3]/I$ is a Noether normalization, so that

$$d=\dim\Bbbk[x_1,x_2,x_3]/I=2.$$

Though $1=(2x_1-1)\frac{\partial f_1}{\partial x_1}-4f_1\in I_1(\frac{\partial f_i}{\partial x_j})+I$, however, $A=\mathrm{V}(I)\subset\mathbb{A}^3$ is not smooth. In fact, $A=\mathrm{V}(x_1)\cup\mathrm{V}(x_1-1,x_2x_3)$ is the union of a plane and a pair of lines intersecting in a point which is necessarily a singular point of A.



□

Exercise 4.1.16. Consider the matrix

$$D = \begin{pmatrix} x_1 & x_2 & x_3^2 - 1 \\ x_2 & x_3 & x_1x_2 + x_3 + 1 \\ x_3^2 - 1 & x_1x_2 + x_3 + 1 & 0 \end{pmatrix}$$

and the ideal $I = \langle f_1, f_2 \rangle \subset \Bbbk[x_1, x_2, x_3]$ generated by $f_1 = \det D$ and the "first" 2×2 minor $f_2 = x_1x_3 - x_2^2$ of D. Verify by computation:

1. The algebraic set $A = \mathrm{V}(I) \subset \mathbb{A}^3$ is equidimensional of dimension $d = 1$.
2. The zero locus of the ideal $J = I_2(\frac{\partial f_i}{\partial x_j}) + I$ coincides with that of I. That is, $\mathrm{V}(J) = \mathrm{V}(I) = A$.
3. The vanishing ideal $\mathrm{I}(A) = (I : J)\,\mathbb{K}[x_1, x_2, x_3]$.
4. A is smooth.

The geometric interpretation of this is that the two hypersurfaces $\mathrm{V}(f_1)$ and $\mathrm{V}(f_2)$ touch each other along A.

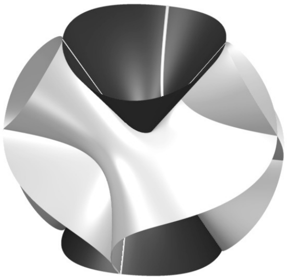


Fig. 4.1. *The cone $V(f_2)$ (dark surface) together with $V(f_1)$ (bright surface) and their intersection (white curve).*

□

Definition 4.1.6 treats the tangent space T_pA *externally*, that is, as a subspace of the ambient space $\mathbb{A}^n$. Hence, it is not obvious that under an isomorphism $\varphi : A \to B$ the tangent spaces at p and $\varphi(p)$ are isomorphic. To prove this, we give an *intrinsic* description of T_pA which only depends on the coordinate ring $\mathbb{K}[A]$.

We consider $T_p\mathbb{A}^n = \mathbb{A}^n$ as an abstract vector space with origin p and coordinates $X_i = x_i - a_i$, $i = 1, \dots, n$. Then $T_pA = \mathrm{V}(d_pf \mid f \in \mathrm{I}(A)) \subset T_p\mathbb{A}^n$ is a linear subspace. Indeed, for each $f \in \mathbb{K}[x_1, \dots, x_n]$, the differential d_pf is linear in the $x_i - a_i$. Moreover, the restriction of d_pf to T_pA depends only on the residue class $\overline{f} = f + \mathrm{I}(A)$ of f in $\mathbb{K}[A]$. We, thus, obtain a well-defined linear map

$$d_p : \mathbb{K}[A] \to T_p^*A,\ \overline{f} \mapsto d_pf|T_pA,$$

where $T_p^*A = \mathrm{Hom}(T_pA, \mathbb{K})$ is the dual vector space of T_pA. The map d_p is surjective since the d_pX_i form a basis for the dual vector space of $T_p\mathbb{A}^n$ and every linear form on T_pA is induced by a linear form on $T_p\mathbb{A}^n$. To describe T_p^*A and, thus, $T_pA = (T_p^*A)^*$ in terms of $\mathbb{K}[A]$, we need to identify the kernel of d_p. Since $d_p c = 0$ for each constant $c \in \mathbb{K}$, the map d_p is determined by its values on the maximal ideal

$$\mathrm{I}_A(p) := \mathrm{I}_A(\{p\}) = \{\overline{f} \in \mathbb{K}[A] \mid f(p) = 0\} \subset \mathbb{K}[A]$$

corresponding to p. We may, thus, as well study the restricted map

$$d_p : \mathrm{I}_A(p) \to T_p^* A, \ \overline{f} \mapsto d_p f | T_p A.$$

This map vanishes on the second power of $\mathrm{I}_A(p)$ (the terms of degree ≥ 2 in the Taylor expansion of f at p do not contribute to $d_p f$). In fact, we have the following result (the final version of this result, proved in Section 4.2, will lead us to the definition of the Zariski tangent space):

Theorem 4.1.17 (Zariski Tangent Space, Preliminary Version). *Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p \in A$ be a point. The $\mathbb{K}[A]$-module $\mathrm{I}_A(p)/\mathrm{I}_A^2(p)$ is naturally a $\mathbb{K}$-vector space. Moreover, the map d_p defines an isomorphism*

$$\mathrm{I}_A(p)/\mathrm{I}_A^2(p) \cong T_p^* A$$

of $\mathbb{K}$-vector spaces.

Proof. Since the $\mathbb{K}[A]$-module $\mathrm{I}_A(p)/\mathrm{I}_A^2(p)$ is annihilated by $\mathrm{I}_A(p)$, it is naturally a $\mathbb{K}[A]/\mathrm{I}_A(p)$-module. The first assertion follows since $\mathbb{K}[A]/\mathrm{I}_A(p) \cong \mathbb{K}$, where the isomorphism is defined by evaluating polynomial functions at p. To prove the theorem, it remains to show that $\ker d_p \subset \mathrm{I}_A^2(p)$. Let $\overline{f} \in \ker d_p$. That is, $\overline{f} \in \mathrm{I}_A(p)$ and $d_p f | T_p A = 0$. Then, if $f_1, \dots, f_r$ are generators for $\mathrm{I}(A)$, the differential $d_p f$ is a $\mathbb{K}$-linear combination of the $d_p f_i$:

$$d_p f = \sum_{i=1}^{r} \lambda_i d_p f_i.$$

Set $g = f - \sum_{i=1}^{r} \lambda_i f_i$. Then $g(p) = 0$ and $d_p g = 0$. We conclude that $g \in \mathrm{I}^2(p) \subset \mathbb{K}[x_1, \dots, x_n]$, so that $\overline{f} = \overline{g} \in \mathrm{I}_A^2(p) \subset \mathbb{K}[A]$. □

Let, now, $\varphi : A \to B$ be a morphism of affine algebraic sets, let $\varphi^* : \mathbb{K}[B] \to \mathbb{K}[A]$ be the induced map, let $p \in A$ be a point, and let $q = \varphi(p)$. Then

$$\varphi^*(\mathrm{I}_B(q)) \subset \mathrm{I}_A(p) \ \text{ and } \ \varphi^*(\mathrm{I}_B^2(q)) \subset \mathrm{I}_A^2(p).$$

Thus, φ defines a map $\varphi^* : \mathrm{I}_B(q)/\mathrm{I}_B^2(q) \to \mathrm{I}_A(p)/\mathrm{I}_A^2(p)$. The dual map

$$d_p \varphi : T_p A \cong (\mathrm{I}_A(p)/\mathrm{I}_A^2(p))^* \to (\mathrm{I}_B(q)/\mathrm{I}_B^2(q))^* \cong T_q B$$

is called the **differential** of φ at p. Note that if $\psi : B \to C$ is another morphism of affine algebraic sets, then

$$d_p(\psi \circ \varphi) = \mathrm{d}_{\varphi(p)}\, \psi \circ d_p \varphi.$$

Furthermore,

$$d_p(\mathrm{id}_A) = \mathrm{id}_{T_pA}.$$

These observations show that the tangent space is invariant under isomorphims:

Corollary 4.1.18. *If $\varphi : A \to B$ is an isomorphism of affine algebraic sets and $p \in A$ is a point, then*

$$d_p\varphi : T_pA \to T_{\varphi(p)}B$$

is an isomorphism of $\mathbb{K}$-vector spaces. □

4.2 Local Rings

In this section, given an algebraic set A and a point $p \in A$, we will describe the construction of the local ring $\mathcal{O}_{A,p}$. This ring is the basic invariant of A at p. We will use it to express smoothness in algebraic terms.

The elements of $\mathcal{O}_{A,p}$ are functions defined on A "near" p. More precisely, the functions are defined on Zariski open neighborhoods of p in A, and two such functions will be identified if they coincide on a sufficiently small neighborhood of p on which both functions are defined. In this sense, the elements of $\mathcal{O}_{A,p}$ are actually **germs of functions**.

What functions are allowed in the construction of $\mathcal{O}_{A,p}$? Since every Zariski neighborhood of p in A contains an open neighborhood of type $\mathrm{D}_A(f) = A \setminus \mathrm{V}_A(f)$, where $f \in \mathbb{K}[A]$ is not vanishing at p, we can restrict ourselves to describe the admissible functions on a neighborhood of this type. Now, note that on $\mathrm{D}_A(f)$, the function f and, thus, its powers f^m are invertible. It is therefore natural to associate to $\mathrm{D}_A(f)$ the $\mathbb{K}$-algebra $\mathbb{K}[A]_f$ of functions on $\mathrm{D}_A(f)$ obtained by adjoining $1/f$ to $\mathbb{K}[A]$. The elements of $\mathrm{D}_A(f)$ are, then, fractions of type g/f^m, where $g \in \mathbb{K}[A]$ and $m \geq 0$. Two such fractions g/f^m and $g'/f^{m'}$ define the same function on $\mathrm{D}_A(f)$ iff $gf^{m'} - g'f^m = 0$ as functions on $\mathrm{D}_A(f)$. Equivalently, $f(gf^{m'} - g'f^m) = 0$ on all of A. That is, $f(gf^{m'} - g'f^m) = 0 \in \mathbb{K}[A]$.

The desired local ring $\mathcal{O}_{A,p}$ is obtained by inverting all the functions in $\mathbb{K}[A]$ not vanishing at p. Its elements are fractions of type g/h, where $g, h \in \mathbb{K}[A]$, with $h(p) \neq 0$. Here, two such fractions g/h and g'/h' will be identified if $gh' - g'h = 0$ on some neighborhood of p contained in $\mathrm{D}_A(h) \cap \mathrm{D}_A(h')$. As pointed out above, we may choose this neighborhood to be of type $\mathrm{D}_A(f)$, where $f \in \mathbb{K}[A]$ is not vanishing at p. Thus, g/h and g'/h' will be identified if $f(gh' - g'h) = 0 \in \mathbb{K}[A]$ for some $f \in \mathbb{K}[A]$ with $f(p) \neq 0$.

The construction of both rings $\mathbb{K}[A]_f$ and $\mathcal{O}_{A,p}$ follows the same algebraic principle: we invert elements of a multiplicative closed subset U of a ring R (it is natural to invert elements from multiplicatively closed subsets since the product of two inverted elements is an inverse for the product).

The principle is familiar to us from Section 2.6 where we studied the quotient field of an integral domain R. In that case, $U = R \setminus \{0\}$. In the more general setting considered here, however, U may contain zerodivisors (such as x or y in $\mathbb{K}[x,z]/\langle xy \rangle$). Thus, we cannot conclude from an equation of type $f(gh' - g'h) = 0$ that $gh' - g'h = 0$.

Taking our cue from these considerations, we arrive at the following purely algebraic definition:

Remark-Definition 4.2.1. Let R be a ring, and let $U \subset R$ be a multiplicatively closed subset. The relation $\sim$ on $R \times U$ defined by

$$(r,u) \sim (r',u') \iff v(ru' - ur') = 0 \text{ for some } v \in U$$

is an equivalence relation (check this; observe that if we just had $ru' - ur' = 0$ in the definition of $\sim$, the transitivity law would fail if U contains zerodivisors). We write r/u for the equivalence class of (r,u) and

$$R[U^{-1}] = U^{-1}R = \{\frac{r}{u} \mid r \in R, u \in U\}$$

for the set of all equivalence classes. We make $R[U^{-1}]$ into a ring by defining

$$\frac{r}{u} + \frac{r'}{u'} = \frac{ur' + u'r}{uu'} \quad \text{and} \quad \frac{r}{u} \cdot \frac{r'}{u'} = \frac{rr'}{uu'}$$

(check that these definitions are independent of the choice of representatives). This ring is called the **localization of R at U**.

We have the natural ring homomorphism

$$\iota : R \to R[U^{-1}],\ r \mapsto \frac{r}{1},$$

which sends every element of U to a unit in $R[U^{-1}]$, and maps an element $r \in R$ to zero iff r is annihilated by an element of U. In particular, ι is injective iff U does not contain a zerodivisor, and $R[U^{-1}]$ is zero iff $0 \in U$. □

Exercise* 4.2.2 (Universal Property of Localization). Let R be a ring, and let $U \subset R$ be a multiplicatively closed subset. Show that if $\phi : R \to S$ is a homomorphism of rings which maps the elements of U to units, there exists a uniquely determined homomorphism $\Phi : R[U^{-1}] \to S$ such that the diagram

$$\begin{tikzcd} R \arrow[rr, "\phi"] \arrow[dr, "\iota"'] & & S \\ & R[U^{-1}] \arrow[ur, dashed, "\Phi"'] & \end{tikzcd}$$

commutes. □

Exercise* 4.2.3 (Localization Commutes with Passing to Quotients by Ideals). Let R and U be as above, let $I \subset R$ be an ideal, and let $\overline{U}$ be the image of U in R/I. Then show that the natural map

$$R \to R[U^{-1}] \to R[U^{-1}]/I\,R[U^{-1}]$$

induces an isomorphism

$$(R/I)[\overline{U}^{-1}] \cong R[U^{-1}]/I\,R[U^{-1}]. \qquad \square$$

Basic examples of localized rings are obtained by considering the multiplicative closed sets introduced earlier in this book:

Remark-Definition 4.2.4. Let R be a ring.

1. If R is an integral domain, and $U = R \setminus \{0\}$, then $R[U^{-1}]$ is the quotient field $\mathrm{Q}(R)$ of R, and any localization of R can be regarded as a subring of $\mathrm{Q}(R)$, with quotient field $\mathrm{Q}(R)$ (apply the universal property). If R is arbitrary, we may consider the multiplicatively closed set U of all nonzerodivisors of R. We, again, write $\mathrm{Q}(R) = R[U^{-1}]$, and call $\mathrm{Q}(R)$ the **total quotient ring** of R. Since U does not contain a zerodivisor, the natural ring homomorphism $\iota : R \to \mathrm{Q}(R)$ is injective, and we may consider R as a subring of $\mathrm{Q}(R)$ by means of ι.
2. If f is an element of R, then $U = \{f^m \mid m \geq 0\}$ is multiplicatively closed. We write $R_f = R[1/f] = R[U^{-1}]$ in this case.
3. If $\mathfrak{p}$ is a prime ideal of R, then $U = R \setminus \mathfrak{p}$ is multiplicatively closed. We write $R_{\mathfrak{p}} = R[U^{-1}]$ in this case, and call $R_{\mathfrak{p}}$ the **localization of R at $\mathfrak{p}$**. □

Example 4.2.5. By inverting all elements in $U = \mathbb{Z} \setminus \{0\}$, we obtain the field $\mathbb{Q}$ of rational numbers. Inverting fewer elements, we get subrings of $\mathbb{Q}$. For instance, if $n \in \mathbb{Z}$ is any number, we get the subring

$$\mathbb{Z}[1/n] = \{a/b \in \mathbb{Q} \mid b = n^k \text{ for some } k \in \mathbb{N}\}.$$

Or, if $p \in \mathbb{Z}$ is any prime number, we get the subring

$$\mathbb{Z}_{\langle p \rangle} = \{a/b \in \mathbb{Q} \mid p \text{ does not divide } b\}.$$

If p does not divide n, we have ring inclusions

$$\mathbb{Z} \subset \mathbb{Z}[1/n] \subset \mathbb{Z}_{\langle p \rangle} \subset \mathbb{Q}. \qquad \square$$

Remark 4.2.6. If $\mathfrak{p}$ is a prime ideal of a ring R, the nonunits of the ring $R_{\mathfrak{p}}$ form the ideal

$$\mathfrak{p}R_{\mathfrak{p}} = \{r/u \mid r \in \mathfrak{p}, u \in R \setminus \mathfrak{p}\}.$$

Taking Remark 1.3.8 into account, we find that $(R_{\mathfrak{p}}, \mathfrak{p}R_{\mathfrak{p}})$ is a local ring in the sense of Definition 1.3.7. By Exercise 4.2.3, the residue field is

$$R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}} \cong \mathrm{Q}(R/\mathfrak{p}). \qquad \square$$

Generalizing what we observed in the remark, our next result shows that the ideal theory of a localized ring is always a simplified version of the ideal theory of the original ring. This result is the main reason for the importance of rings of fractions in commutative algebra.

Theorem 4.2.7. *Let R be a ring, let $U \subset R$ be a multiplicative closed subset, and let $\iota : R \to R[U^{-1}]$, $r \mapsto r/1$, be the natural homomorphism.*

1. *If $I \subset R$ is an ideal, then*
$$\iota^{-1}(IR[U^{-1}]) = \{a \in R \mid ua \in I \text{ for some } u \in U\}.$$
2. *If $J \subset R[U^{-1}]$ is an ideal, then*
$$\iota^{-1}(J)R[U^{-1}] = J.$$
We, thus, get an injectice map of the set of ideals of $R[U^{-1}]$ into the set of ideals of R by sending J to $\iota^{-1}(J)$.
3. *If R is Noetherian, then so is $R[U^{-1}]$.*
4. *The injection $J \mapsto \iota^{-1}(J)$ restricts to a bijection between the set of prime ideals of $R[U^{-1}]$ and the set of prime ideals of R not meeting U.*

Proof. For part 1, observe that if $a \in R$, then $a \in \iota^{-1}(IR[U^{-1}]) \iff a/1 \in IR[U^{-1}] \iff ua \in I$ for some $u \in U$. For part 2, let $b/u \in R[U^{-1}]$, where $b \in R$ and $u \in U$. Then $b/u \in J \iff b/1 \in J \iff b \in \iota^{-1}(J) \iff b/u \in \iota^{-1}(J)R[U^{-1}]$. Part 3 follows from part 2 (for instance, use the ascending chain condition). For part 4, notice that if $\mathfrak{q}$ is a prime ideal of $R[U^{-1}]$, then $\mathfrak{p} = \iota^{-1}(\mathfrak{q})$ is a prime ideal of R. Moreover, $\mathfrak{p} \cap U = \emptyset$ since $\mathfrak{q}$ does not contain units. Conversely, let $\mathfrak{p}$ be a prime ideal of R such that $\mathfrak{p} \cap U = \emptyset$. If $a/u \cdot b/v \in \mathfrak{p}R[U^{-1}]$, with $u, v \in U$, then $wab \in \mathfrak{p}$ for some $w \in U$. Since $w \notin \mathfrak{p}$, we must have $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$ and, thus, $a/u \in \mathfrak{p}R[U^{-1}]$ or $b/v \in \mathfrak{p}R[U^{-1}]$. Moreover, $1 \notin \mathfrak{p}R[U^{-1}]$, so $\mathfrak{p}R[U^{-1}]$ is a prime ideal of $R[U^{-1}]$. The result follows from part 1 since $\iota^{-1}(\mathfrak{p}R[U^{-1}]) = \{a \in R \mid ua \in \mathfrak{p} \text{ for some } u \in U\} = \mathfrak{p}$. □

Exercise* 4.2.8 (Localization Commutes with Forming Radicals). If $I \subset R$ is an ideal, then show that $\operatorname{rad}(IR[U^{-1}]) = (\operatorname{rad} I)R[U^{-1}]$. Conclude that the injection $J \mapsto \iota^{-1}(J)$ restricts to a bijection between the set of primary ideals of $R[U^{-1}]$ and the set of primary ideals of R not meeting U. □

In the geometric setting, given an algebraic set A, we apply the constructions discussed in Example 4.2.4 to the coordinate ring $\mathbb{K}[A]$.

To begin with, the total quotient ring $\mathbb{K}(A) := \mathrm{Q}(\mathbb{K}[A])$ is the **ring of rational functions** on A. Here, the terminology introduced in Section 2.6 for rational functions on varieties carries over to rational functions on arbitrary algebraic sets. In particular, we define the **domain of definition** $\operatorname{dom}(f)$ of a rational function $f \in \mathbb{K}(A)$ as in Section 2.6, and view f as a function on $\operatorname{dom}(f)$. Note that $\operatorname{dom}(f)$ is open and, by Exercise 1.11.9, dense in the Zariski topology on A.

If $f \in \mathbb{K}[A]$, the localization $\mathbb{K}[A]_f$ is the $\mathbb{K}$-algebra of functions on $\mathrm{D}_A(f)$ considered in the introduction to this section.

Similarly, if $p \in A$ is a point, the local ring $\mathcal{O}_{A,p}$ is formally defined as the localization of $\mathbb{K}[A]$ at the maximal ideal of $\mathbb{K}[A]$ corresponding to p:

Remark-Definition 4.2.9. Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p \in A$ be a point. The **local ring of *A* at *p*,** written $\mathcal{O}_{A,p}$, is defined to be the localization

$$\mathcal{O}_{A,p} = \mathbb{K}[A]_{\mathfrak{m}},$$

where $\mathfrak{m} = \mathrm{I}_A(p) \subset \mathbb{K}[A]$ is the maximal ideal corresponding to p. Taking Remark 4.2.6 and part 3 of Proposition 4.2.7 into account, we find that $\mathcal{O}_{A,p}$ is a local Noetherian ring with maximal ideal

$$\mathfrak{m}_{A,p} := \{f/g \in \mathcal{O}_{A,p} \mid f(p) = 0\}.$$

Furthermore, by Exercise 4.2.3,

$$\mathcal{O}_{A,p} = \mathcal{O}_{\mathbb{A}^n,p}/\mathrm{I}(A)\mathcal{O}_{\mathbb{A}^n,p}.$$

□

Exercise 4.2.10. Let $B_1, B_2 \subset \mathbb{A}^n$ be algebraic sets, let $A = B_1 \cup B_2$, and let $p \in A$ be a point not lying on B_2. Then show that $\mathcal{O}_{A,p} \cong \mathcal{O}_{B_1,p}$. □

Remark 4.2.11. If V is an affine variety, the local rings $\mathcal{O}_{V,p}$, $p \in V$, are subrings of $\mathbb{K}(V)$ containing $\mathbb{K}[V]$. In fact, by Proposition 2.6.15,

$$\mathbb{K}[V] = \bigcap_{p \in V} \mathcal{O}_{V,p} \subset \mathbb{K}(V).$$

□

Remark 4.2.12. Instead of just considering local rings at points, it makes also sense to consider the **local ring of *A* along** a subvariety ***W*** of A. This ring, written $\mathcal{O}_{A,W}$, is the localization of $\mathbb{K}[A]$ at the prime ideal $\mathfrak{p} = \mathrm{I}_A(W)$. If $A = V$ is a variety, then $\mathcal{O}_{V,W}$ is a subring of $\mathbb{K}(V)$, namely the subring consisting of all rational functions on V that are defined at some point of W (and, hence, defined on a dense open subset of W). □

We postpone the further development of the general theory of localization to Section 4.5. Our next goal in this section is to characterize the smoothness of an algebraic set A at a point $p \in A$ in terms of the local ring $\mathcal{O}_{A,p}$. To begin with, we characterize the local dimension $\dim_p A$ in terms of $\mathcal{O}_{A,p}$:

Proposition 4.2.13. *If R is a ring, and $\mathfrak{p}$ is a prime ideal of R, then*

$$\dim R_{\mathfrak{p}} = \operatorname{codim} \mathfrak{p}.$$

In particular, if $A \subset \mathbb{A}^n$ is an algebraic set, and $p \in A$ is a point, then

$$\dim \mathcal{O}_{A,p} = \dim_p A.$$

Proof. By Proposition 4.2.7, there is a one-to-one correspondence between maximal chains of prime ideals of $R_{\mathfrak{p}}$ and maximal chains of prime ideals of R with largest ideal $\mathfrak{p}$:

$$\mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \ldots \subsetneq \mathfrak{p}_d = \mathfrak{p}.$$

This shows the first assertion. For the second assertion, note that if $R = \mathbb{K}[A]$, and $\mathfrak{p} = \mathrm{I}_A(p) \subset R$ is the maximal ideal corresponding to p, then a chain as above corresponds to a chain of subvarieties $W_i := \mathrm{V}_A(\mathfrak{p}_i) \subset A$ containing p. The variety W_0 is actually an irreducible component of A since otherwise we could insert a prime ideal strictly contained in $\mathfrak{p}_0$. Moreover,

$$\langle 0 \rangle \subsetneq \mathfrak{p}_1/\mathfrak{p}_0 \subsetneq \ldots \subsetneq \mathfrak{p}_d/\mathfrak{p}_0$$

is a maximal chain of prime ideals of $\mathbb{K}[W_0] \cong \mathbb{K}[A]/\mathfrak{p}_0$. Every such chain has length $\dim W_0$ by Corollary 3.4.9. Conversely, if $\mathfrak{p}_0 \subset \mathbb{K}[A]$ is a prime ideal such that $\mathrm{V}_A(\mathfrak{p}_0)$ is an irreducible component of A passing through p, then $\mathfrak{p}_0$ fits as smallest ideal into a maximal chain of prime ideals of $\mathbb{K}[A]$ with largest ideal $\mathfrak{p} = \mathrm{I}_A(p)$. □

Next, in the final version of Theorem 4.1.17, we describe the tangent space T_pA in terms of $\mathcal{O}_{A,p}$. For this, note that if $(R, \mathfrak{m})$ is a local ring with residue field $R/\mathfrak{m}$, then $\mathfrak{m}/\mathfrak{m}^2$ is naturally an $R/\mathfrak{m}$-module. That is, $\mathfrak{m}/\mathfrak{m}^2$ is an $R/\mathfrak{m}$-vector space.

Theorem-Definition 4.2.14 (Zariski Tangent Space, Final Version). *If $A \subset \mathbb{A}^n$ is an algebraic set, and $p \in A$ is a point, there is a natural isomorphism of $\mathbb{K}$-vector spaces*

$$(\mathfrak{m}_{A,p}/\mathfrak{m}_{A,p}^2)^* \cong T_pA.$$

*We call $(\mathfrak{m}_{A,p}/\mathfrak{m}_{A,p}^2)^*$ the **Zariski tangent space** to A at p.*

Proof. Let $f = g/h \in \mathbb{K}(x_1, \ldots, x_n)$ be a rational function such that $h(p) \neq 0$. In extending what we did for polynomials, we define the **differential $\boldsymbol{d_pf}$ of $\boldsymbol{f}$ at $\boldsymbol{p}$** by formally writing down the quotient rule:

$$d_pf := \frac{h(p)d_pg - g(p)d_ph}{h^2(p)}$$

(this is independent of the choice of representation for f as a fraction). Arguing, now, as in the proof of Theorem 4.1.17, we get a map

$$d_p : \mathfrak{m}_{A,p} \to T_p^*A,\ \overline{f} = \overline{g}/\overline{h} \mapsto d_pf|T_pA$$

whose kernel is $\mathfrak{m}_{A,p}^2$. □

Combining Proposition 4.2.13 and Theorem 4.2.14, we get:

Corollary 4.2.15. *Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p \in A$ be a point. Then A is smooth at p iff*

$$\dim_{\mathbb{K}} \mathfrak{m}_{A,p}/\mathfrak{m}_{A,p}^2 = \dim \mathcal{O}_{A,p}.$$

□

Corollary 4.2.16. *If $A \subset \mathbb{A}^n$ is an algebraic set, then A_{sing} and A have no irreducible component in common.*

Proof. As already pointed out in Remark 4.1.11, we will show in Corollary 4.6.26 that a point of A is singular iff it lies on the intersection of two irreducible components of A or is a singular point of one of the components. For our purposes here, it is, hence, enough to show that if V is such a component, then V contains V_{sing} properly. By Proposition 4.1.3, this is true in the hypersurface case. To reduce to this case, we apply Theorem 3.5.2: let $\phi : V \to W$ be a finite morphism onto a hypersurface $W \subset \mathbb{A}^{d+1}$ admitting a rational inverse $\psi : W \dashrightarrow V$. Then, since W_{sing} is a proper algebraic subset of W, the set $U := \text{dom}(\psi) \cap (W \setminus W_{\text{sing}})$ is Zariski dense in W. In particular, U is nonempty. But if $q = \phi(p)$ is a point of U, the isomorphism $\phi^* : \mathbb{K}(W) \to \mathbb{K}(V)$ restricts to an isomorphism $\mathcal{O}_{W,q} \cong \mathcal{O}_{V,p}$. Hence, we are done by Corollary 4.2.15. □

The inequality

$$\dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2 \geq \dim R \tag{4.3}$$

holds for any local Noetherian ring $(R, \mathfrak{m})$ (this is the. general algebraic form of inequality (4.1) on Page 143 which we will be prove in Corollary 4.6.20). The importance of Corollary 4.2.15 is emphasized by the following definition:

Definition 4.2.17 (Krull). A local Noetherian ring $(R, \mathfrak{m})$ is called **regular** if $\dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2 = \dim R$. □

Using this notion, we can restate Corollary 4.2.15 as follows:

Corollary 4.2.18. *Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p \in A$ be a point. Then A is smooth at p iff $\mathcal{O}_{A,p}$ is a regular local ring .* □

In most textbooks on commutative algebra, the definition of a regular local ring involves a characterization of $\dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2$ in terms of generators for $\mathfrak{m}$. This is obtained as an application of the following fundamental result:

Theorem 4.2.19 (Lemma of Nakayama). *Let $(R, \mathfrak{m})$ be a local ring, let M be a finitely generated R-module, and let $N \subset M$ be a submodule. Then*

$$N + \mathfrak{m}M = M \ \textit{ iff } \ N = M.$$

Proof. Replacing M by M/N, we reduce to the case $N = 0$. That is, it suffices to show that $\mathfrak{m}M = M$ implies $M = 0$ (the converse implication is clear). Let $m_1, \dots, m_r$ be a finite set of generators for M. If $\mathfrak{m}M = M$, we may write each m_i as an $\mathfrak{m}$-linear combination of the m_j:

$$m_i = \sum r_{ij} m_j, \text{ with all } r_{ij} \in \mathfrak{m}.$$

In matrix notation,

$$(E_r - B) \begin{pmatrix} m_1 \\ \vdots \\ m_r \end{pmatrix} = 0,$$

where $B = (r_{ij})$ and E_r is the $r \times r$ identity matrix. Arguing once more as in the proof of the Projection Theorem 3.1.2, we multiply with the matrix of cofactors of $(E_r - B)$, and obtain that $h = \det(E_r - B)$ annihilates each m_i. This implies that the m_i and, thus, M are zero. Indeed, h is a unit in R since $h \equiv 1 \mod \mathfrak{m}$. □

Starting from well-known facts on vector spaces, Nakayama's lemma allows us to deduce information on modules over local rings. In making this explicit, we use the following notation: If R is any ring, and M is any R-module, a **set of generators** for M is **minimal** if no proper subset generates M.

Corollary 4.2.20. *Let $(R, \mathfrak{m})$ and M be as in Nakayama's Lemma 4.2.19. Then $m_1, \dots, m_r \in M$ generate M as an R-module iff the residue classes $\overline{m}_i = m_i + \mathfrak{m}M$ generate $M/\mathfrak{m}M$ as an $R/\mathfrak{m}$-vector space. In particular, any minimal set of generators for M corresponds to an $R/\mathfrak{m}$-basis for $M/\mathfrak{m}M$, and any two such sets have the same number of elements.*

Proof. Let $N = \langle m_1, \dots, m_r \rangle \subset M$. Then $m_1, \dots, m_r$ generate M iff $N + \mathfrak{m}M = M$ iff $\operatorname{span}(\overline{m}_1, \dots, \overline{m}_r) = M/\mathfrak{m}M$. □

Corollary 4.2.21. *A local Noetherian ring $(R, \mathfrak{m})$ is* **regular** *iff $\mathfrak{m}$ can be generated by $\dim R$ elements.* □

The first part of the exercise below shows that the conclusion of Corollary 4.2.20 may be wrong over arbitrary rings:

Exercise 4.2.22. 1. Find an ideal of $\Bbbk[x_1, \dots, x_n]$ which admits minimal sets of generators differing in their number of elements.
2. Let $\mathcal{O}_{\mathbb{A}^2, o}$ be the local ring of $\mathbb{A}^2$ at the origin $o = (0, 0)$. For each $n \in \mathbb{N}$, find an ideal of $\mathcal{O}_{\mathbb{A}^2, o}$ which is minimally generated by n elements. □

Another application of Nakayama's lemma, which we present for later use, is a special case of Krull's intersection theorem (see Eisenbud (1995), Corollary 5.4 for the general case):

Theorem 4.2.23 (Krull's Intersection Theorem). *Let $(R, \mathfrak{m})$ be a local Noetherian ring. Then*

$$\bigcap_{k=0}^{\infty} \mathfrak{m}^k = \langle 0 \rangle.$$

Proof. In the polynomial ring $R[t]$, consider the subalgebra

$$S = R[\mathfrak{m}t] = R \oplus \mathfrak{m}t \oplus \mathfrak{m}^2t^2 \oplus \ldots \subset R[t].$$

Since R is Noetherian, $\mathfrak{m}$ is a finitely generated ideal of R. It follows that S is a finitely generated R-algebra and, thus, that S is Noetherian, too. In particular, if $J = \bigcap_{k=0}^{\infty} \mathfrak{m}^k$, the ideal

$$J \oplus Jt \oplus Jt^2 \oplus \ldots \subset S$$

is generated by finitely many *homogeneous* polynomials in $R[t]$ (take the homogeneous components of any finite set of generators). If r is the maximum degree in of the generators, then $\mathfrak{m}tJt^r = Jt^{r+1}$. That is,

$$\mathfrak{m} \bigcap_{k=0}^{\infty} \mathfrak{m}^k = \bigcap_{k=0}^{\infty} \mathfrak{m}^k \subset R.$$

The result follows from Nakayama's lemma. □

Example 4.2.24. The conclusion of the intersection theorem may not hold if R is not Noetherian. For instance, let R be the ring of germs of $\mathcal{C}^\infty$ functions defined on arbitrarily small ϵ-neighborhoods of the origin $0 \in \mathbb{R}$ (that is, the elements of R are obtained by identifying two functions if they coincide on a sufficiently small neighborhood of 0). Then R is local with maximal ideal $\mathfrak{m} = \langle x \rangle$, where x is (the germ of) the coordinate function. On the other hand, the function

$$g(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \end{cases}$$

defines a (nontrivial) element of $\bigcap_{k=0}^{\infty} \mathfrak{m}^k$: indeed, $g(x)/x^k$ is $\mathcal{C}^\infty$ for every k. In particular, R cannot be Noetherian by Krull's intersection theorem. □

We end this section as we have started it, namely by considering admissible functions. So far, given an algebraic set $A \subset \mathbb{A}^n$, we have described the functions allowed on distinguished open subsets of A. Now, taking our cue from Proposition 2.6.15, we deal with arbitrary open subsets:

Remark-Definition 4.2.25. Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $U \subset A$ be an open subset. A function $f : U \to \mathbb{K}$ is called **regular at a point** $p \in U$ if there are $g, h \in \mathbb{K}[A]$ such that $h(q) \neq 0$ and $f(q) = g(q)/h(q)$ for all $q \in U$. We say that f is **regular on** $\boldsymbol{U}$ if it is regular at every point of U. The set $\mathcal{O}(U)$ of all regular functions on U becomes a ring, with pointwise defined algebraic operations. That is, we add and multiply values in $\mathbb{K}$.

On distinguished open subsets, we get the functions already familiar to us:

Proposition 4.2.26. *Let $A \subset \mathbb{A}^n$ be an algebraic set. If $0 \neq h \in \mathbb{K}[A]$, then for each regular function f on $\mathrm{D}_A(h)$, there exist $g \in \mathbb{K}[A]$ and $m \geq 1$ such that $f(p) = g(p)/h(p)^m$ for all $p \in \mathrm{D}_A(h)$. That is, we may identify $\mathcal{O}(\mathrm{D}_A(h)) = \mathbb{K}[A]_h$. In particular, taking $h = 1$, we get $\mathcal{O}(A) = \mathbb{K}[A]$. That is, the regular functions on A are precisely the polynomial functions.*

Proof. Let f be a regular function on $\mathrm{D}_A(h)$. Since the Zariski topology is quasicompact, we can find a finite family of pairs of functions $g_i, h_i \in \mathbb{K}[A]$, say $i = 1, \ldots, N$, such that $\mathrm{D}_A(h) = \bigcup_{i=1}^N \mathrm{D}_A(h_i)$, and such that $f = g_i/h_i$ as functions on $\mathrm{D}_A(h_i)$, for all i. Then, for all i, j, we have $g_ih_j - g_jh_i = 0$ on $\mathrm{D}_A(h_i) \cap \mathrm{D}_A(h_j) = \mathrm{D}_A(h_ih_j)$ and, thus, $h_ih_j(g_ih_j - g_gh_i) = 0$ on all of A. Replacing g_i by g_ih_i and h_i by h_i^2 for all i, we may suppose that $g_ih_j = g_jh_i$ on A for all i, j.

Since $\mathrm{D}_A(h) = \bigcup_{i=1}^N \mathrm{D}_A(h_i)$, we have $\mathrm{V}_A(h) = \mathrm{V}_A(h_1, \ldots, h_N)$. The Nullstellensatz implies that $h^m \in \langle h_1, \ldots, h_N \rangle$ for some $m \geq 1$, say $h^m = \sum_{i=1}^N a_ih_i$, with $a_1, \ldots, a_N \in \mathbb{K}[A]$. Let $g = \sum_{i=1}^N a_ig_i$. Then for all j,

$$h^m g_j = \sum_{i=1}^N a_ih_ig_j = \sum_{i=1}^N a_ig_ih_j = gh_j$$

and, thus, $f = g_j/h_j = g/h^m$ as functions on $\mathrm{D}_A(h_j)$. The result follows since $\mathrm{D}_A(h) = \bigcup_{i=1}^N \mathrm{D}_A(h_i)$. □

Exercise 4.2.27. Show that regular functions are continous when $\mathbb{K}$ is identified with $\mathbb{A}^1$ in its Zariski topology.
Hint: The property that a subset Y of a topological space X is closed is a **local property** in the sense that Y is closed if it can be covered by open subsets U of X such that $Y \cap U$ is closed in Y for all U. □

Exercise 4.2.28 (Characterization of Rational Functions). Let A be an algebraic set. Let Σ be the set of pairs (U, f), where U is a Zariski dense open subset of A, and where $f \in \mathcal{O}(U)$. Show that the relation $\sim$ on Σ defined by

$$(U, f) \sim (U', f') \iff f|U \cap U' = f'|U \cap U'$$

is an equivalence relation. Show that the set of all equivalence classes is a ring which is naturally isomorphic to $\mathbb{K}(A)$ (the sum and product of two classes represented by pairs (U, f) and (U', f') are obtained by adding and multiplying f and f' on $U \cap U'$, respectively). Conclude that if $A = V_1 \cup \cdots \cup V_s$ is the decomposition of A into its irreducible components, then

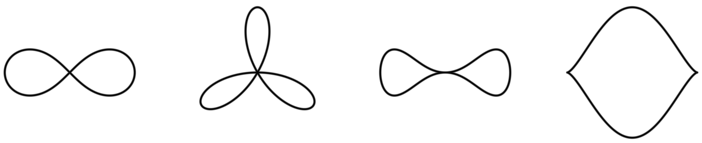
$$\mathbb{K}(A) \cong \mathbb{K}(V_1) \times \cdots \times \mathbb{K}(V_s).$$

□

4.3 Intersection Multiplicities of Plane Curves

In Section 5, we will prove Bezout's Theorem which says that if C, D are two plane curves of degrees d, e without a common component, then C and D intersect in precisely $d \cdot e$ points – provided we work in the right setting, and provided we count the intersection points with appropriate multiplicities. The right setting will be created in Section 5.1 by adding points at infinity. How to define the multiplicities will be explained now. We begin by fixing some terminology for dealing with singularities of plane curves.

Example 4.3.1. The following picture shows plane curves with different types of singularities:



node triple point tacnode cusps

□

Plane curves correspond to nonconstant square-free polynomials $f \in \Bbbk[x, y]$, where f is determined up to multiplication by a nonzero scalar. For reasons which will become clear later in this section, however, it is convenient to allow f to have multiple factors in the following definitions.

Definition 4.3.2. Let $f \in \Bbbk[x, y]$ be a nonconstant polynomial, and let $p = (a, b) \in \mathbb{A}^2$ be a point. Let

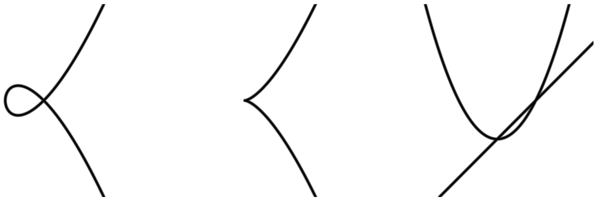
$$f = f_0 + f_1 + f_2 + \ldots + f_d \in \mathbb{K}[x, y]$$

be the Taylor expansion of f at p, where, for each i, the polynomial f_i collects the degree-i terms of f in $x-a$ and $x-b$. The **multiplicity of *f* at *p***, written $\operatorname{mult}(f, p)$, is defined to be the least m such that $f_m \neq 0$. By convention, $\operatorname{mult}(0, p) = \infty$.

If f is square-free, and $C = \mathrm{V}(f) \subset \mathbb{A}^2$ is the corresponding curve, we write $\operatorname{mult}(C, p) = \operatorname{mult}(f, p)$, and call this number the **multiplicity of *C* at *p***. □

Note that $p \in \mathrm{V}(f)$ iff $\operatorname{mult}(f, p) \geq 1$. If f is square-free, and $C = \mathrm{V}(f)$, then $\operatorname{mult}(C, p) = 1$ iff p is a smooth point of C. We speak of a **double point** if the multiplicity m is 2, of a **triple point**, if $m = 3$, and a **quadruple point**, if $m = 4$.

Example 4.3.3. The origin is a double point of each curve shown below:



$y^2 = x^3 + x^2$ $y^2 = x^3$ $y^2 = xy + x^2y - x^3$

□

Different types of singularities of plane curves can often be distinguished by considering the tangent lines at these points. To introduce tangent lines at singular points, we remark that over the algebraically closed field $\mathbb{K}$, every homogeneous polynomial in two variables can be written as a product of linear factors. Indeed, if $g = y^s h \in \mathbb{K}[x, y]$, where y does not divide h, the dehomogenized polynomial $g(x, 1) = h(x, 1)$ is univariate and decomposes, hence, into linear factors: $g(x, 1) = h(x, 1) = \prod_{i=1}^{r-1}(\lambda_i x - \mu_i)^{e_i} \in \mathbb{K}[x, y]$. Homogenizing the factors, we get $g = y^s \prod_{i=1}^{r-1}(\lambda_i x - \mu_i y)^{e_i}$.

Definition 4.3.4. Let $f \in \Bbbk[x, y]$ be a nonconstant polynomial, and let $p = (a, b) \in \mathbb{A}^2$ be a point. Let

$$f = f_m + \ldots + f_d \in \mathbb{K}[x, y]$$

be the Taylor expansion of f at p as in Definition 4.3.2, where $m = \operatorname{mult}(f, p)$. Decompose f_m over $\mathbb{K}$ into pairwise different linear factors in $x - a$ and $y - b$:

$$f_m = \prod_{i=1}^{r}(\lambda_i(x - a) - \mu_i(y - b))^{e_i} \in \mathbb{K}[x, y].$$

The **tangent lines to f at p** are defined to be the lines

$$L_i = \mathrm{V}(\lambda_i(x - a) - \mu_i(y - b)) \subset \mathbb{A}^2,$$

and e_i is the **multiplicity** of L_i.

If f is square-free, and $C = \mathrm{V}(f) \subset \mathbb{A}^2$ is the corresponding curve, the tangent lines to f at p are also called the **tangent lines to C at p**. □

At a smooth point of C, the multiplicity $m = 1$, and the definition above yields precisely the tangent line introduced in Section 4.1. If C has $m \geq 2$ *distinct* tangent lines (of multiplicity 1) at p, we say that p is an **ordinary multiple point** of C. An ordinary double point is called a **node**.

Example 4.3.5. In Example 4.3.3, the origin o is a node of $\mathrm{V}(y^2 - x^2 - x^3)$, with tangent lines $\mathrm{V}(x+y)$ and $\mathrm{V}(x-y)$. Similarly, o is a node of the reducible curve $C = \mathrm{V}(y^2 - xy - x^2y + x^3)$: the two different tangent lines are the line $\mathrm{V}(x - y)$, which is one of the components of C, and the x-axis, which is the tangent line at o to the other component $\mathrm{V}(y - x^2)$ of C. In contrast, the curve $\mathrm{V}(y^2 - x^3)$ has a tangent line of multiplicity 2 at o. □

Exercise 4.3.6. The curves in Example 4.3.1 are defined by the polynomials below:

$$y^2 = (1 - x^2)^3, \quad y^2 = x^2 - x^4, \quad y^3 - 3x^2y = (x^2 + y^2)^2, \quad y^2 = x^4 - x^6.$$

Which curve corresponds to which polynomial? □

Before turning to intersection multiplicities, we present a result which shows that the ideals of local rings of plane curves at smooth points are easy to handle. We need the following notation:

Definition 4.3.7. A **discrete valuation** on a field K is a surjective map $v\colon K \setminus \{0\} \to \mathbb{Z}$ such that, for all $a, b \in K \setminus \{0\}$,

1. $v(ab) = v(a) + v(b)$, and
2. $v(a + b) \geq \min(v(a), v(b))$. □

Note that the first condition of the definition means that $v : K \setminus \{0\} \to \mathbb{Z}$ is a group homomorphism. In particular, $v(1) = 0$. By convention, $v(0) = \infty$. The set

$$R := \{a \in K \mid v(a) \geq 0\}$$

is, then, a subring of K to which we refer as the **valuation ring** of v.

Definition 4.3.8. An integral domain R is called a **discrete valuation ring** (**DVR** for short) if R is the valuation ring of a discrete valuation on its quotient field. □

Example 4.3.9. The ring $\Bbbk[[x]]$ of formal power series $f = \sum_{i=0}^{\infty} a_i x^i$ with coefficients $a_i \in \Bbbk$ is a DVR. Indeed, it is an integral domain with quotient field $\Bbbk((x))$, where

$$\Bbbk((x)) = \{\sum_{i=n}^{\infty} a_i x^i \mid a_i \in \Bbbk \text{ for all } i\}$$

is the field of formal Laurent series with coefficients in $\Bbbk$. The desired valuation on $\Bbbk((x))$ is obtained by setting $v(f) = n$ if $f = \sum_{i=n}^{\infty} a_i x^i$ with $a_n \neq 0$. Using the same terminology as for convergent power and Laurent series in complex analysis, we say that $v(f)$ is the **vanishing order** of a formal power series $f \in \Bbbk[[x]]$ and that a formal Laurent series $f \in \Bbbk((x)) \setminus \Bbbk[[x]]$ has a **pole** of order $-v(f)$. □

If R is a DVR with quotient field K and corresponding discrete valuation v on K, its set of nonunits, which is the set

$$\mathfrak{m} := \{a \in K \mid v(a) \geq 1\},$$

is an ideal of R. Hence, $(R, \mathfrak{m})$ is a local ring. Furthermore, R is a PID: Since v is surjective, there is an element $t \in \mathfrak{m}$ such that $v(p) = 1$, and we claim that every nonzero ideal I of R is of type $I = \langle t^k \rangle = \mathfrak{m}^k = \{a \in R \mid v(a) \geq k\}$, where k is minimal among all $v(g)$, $g \in I$. Indeed, to see this, just note that if a, b are two elements of R, then $v(a) = v(b)$ iff $v(ab^{-1}) = 0$ iff ab^{-1} is a unit of R iff $\langle a \rangle = \langle b \rangle$.

Exercise* 4.3.10. Let R be a local Noetherian integral domain with maximal ideal $\mathfrak{m}$. Suppose that R contains a field L such that the composite map $L \to R \to R/\mathfrak{m}$ is an isomorphism. Then all quotients $\mathfrak{m}^k/\mathfrak{m}^{k+1}$ are L-vector spaces. In this situation, show that R is a DVR iff the following two conditions hold:

1. $\dim_L \mathfrak{m}^k/\mathfrak{m}^{k+1} = 1$ for all $k \geq 0$;
2. $\dim_L R/\mathfrak{m}^k = k$ for all $k \geq 1$. □

Proposition 4.3.11. *Let R be a local ring. Then the following are equivalent:*

1. *R is a DVR.*
2. *R is regular of dimension 1.*

Proof. $1 \implies 2$: If R is a DVR with maximal ideal $\mathfrak{m}$, the only chain of prime ideals of R is $\langle 0\rangle \subsetneq \mathfrak{m}$. So R has Krull dimension one. Moreover, as already pointed out in the discussion preceeding Exercise 4.3.10 , $\mathfrak{m}$ is generated by just one element. So R is regular.

$2 \implies 1$: Conversely, suppose that R is regular of dimension one, and let t be a generator for the maximal ideal $\mathfrak{m}$. To show that R is a DVR, we first observe that $t^r \neq 0$ for all r. Indeed, otherwise, $\mathfrak{m} = \langle t\rangle$ would be the only prime ideal of R, so that R would be zerodimensional. Let, now, $0 \neq g \in R$. By Krull's intersection theorem, g cannot be contained in all powers of $\mathfrak{m}$. Let $k = \max\{r \mid g \in \mathfrak{m}^r\}$. Then $g = ut^k$ for some element $u \in R \setminus \mathfrak{m}$, which necessarily is a unit of R. Similarly, if $0 \neq h$ is another element of R, write h as a product vt^ℓ, for some unit v and some ℓ. Then $gh = uvt^{k+\ell}$ is nonzero, and we conclude that R is an integral domain. Furthermore, any element f of the quotient field $\mathrm{Q}(R)$ has a unique representation of type $f = wt^m$, for some unit w and some $m \in \mathbb{Z}$. Setting $v(f) = m$, we get the desired discrete valuation on $\mathrm{Q}(R)$. □

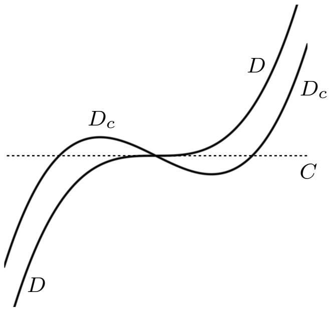
Taking Corollary 4.2.18 into account, we get:

Corollary 4.3.12. *An irreducible curve $C \subset \mathbb{A}^2$ is smooth at a point $p \in C$ iff $\mathcal{O}_{C,p}$ is a discrete valuation ring.* □

If C is smooth at p, we occasionally write $v_{C,p}$ for the corresponding discrete valuation on $\mathbb{K}(C)$. Motivated by Example 4.3.9, we say that $v_{C,p}(f)$ is the **vanishing order** of an element $f \in \mathcal{O}_{C,p}$, and that a rational function $f \in \mathbb{K}(C) \setminus \mathcal{O}_{C,p}$ has a **pole** of order $-v_{C,p}(f)$ at p.

We will, now, define intersection multiplicities. There are several ways of doing this, some of which go back to Newton and his contemporaries (see Fulton (1998), Chapter 7, Notes and References for some historical remarks).

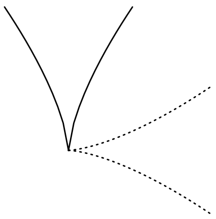
Example 4.3.13. Consider the curves $C = \mathrm{V}(y)$ and $D = \mathrm{V}(y - x^r)$ in $\mathbb{A}^2(\mathbb{C})$. Intuitively, we should count the origin $o = (0, 0)$ as an intersection point of multiplicity r. Indeed, if we perturb the equations defining C and D slightly, we get r distinct intersection points near o:



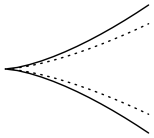
The case $r = 3$.

For a more precise statement, consider, for instance, a perturbation of the defining equation $f_0 = y - x^r$ for D, say $f_c = y - x^r + c_1 x^{r-1} + \ldots + c_r$, where $c = (c_1, \ldots, c_r)$ is a tuple of complex numbers, and let $D_c = \mathrm{V}(f_c) \subset \mathbb{A}^2(\mathbb{C})$. Given a sufficiently small $\epsilon > 0$, there is, then, a number $\delta > 0$ such that for any sufficiently general c with $|c_i| < \delta$, the curve D_c intersects C in r distinct points in the ϵ-neighborhood of the origin (we will prove this in the context of Bertini's theorem in Chapter 6). □

Example 4.3.14. Now, consider the pairs of curves $y^2 - x^3$ and $x^2 - y^3$, respectively $y^2 - x^3$ and $2y^2 - x^3$:



transversal cusps



tangential cusps

In both cases, can you find the intersection multiplicity at the origin? □

It is not immediately clear that the **dynamic** point of view taken in the examples above gives well-defined intersection multiplicities. Furthermore, computing intersection multiplicities in this way can be quite elaborate.

Following Macaulay (1916), we will work with a purely algebraic definition of intersection multiplicities which is **static** in that we do not vary the given equations. The definition is less intuitive, but turns out to be just right.

Definition 4.3.15. Let $f, g \in \Bbbk[x, y]$ be nonconstant polynomials, and let $p \in \mathbb{A}^2$ be a point. The **intersection multiplicity of *f* and *g* at *p***, written $i(f, g; p)$, is defined to be

$$i(f, g; p) = \dim_{\mathbb{K}} \mathcal{O}_{\mathbb{A}^2,p} / \langle f, g \rangle \mathcal{O}_{\mathbb{A}^2,p}.$$

If f, g are square-free, and $C = \mathrm{V}(f), D = \mathrm{V}(f) \subset \mathbb{A}^2$ are the corresponding curves, we write $i(C, D; p) = i(f, g; p)$, and call this number the **intersection multiplicity of $\boldsymbol{C}$ and $\boldsymbol{D}$ at $\boldsymbol{p}$**. □

The calculations in Example 4.3.17 below rely on the following observation:

Remark 4.3.16. Let $I \subset \Bbbk[x_1, \dots, x_n]$ be an ideal such that $\mathrm{V}(I) \subset \mathbb{A}^n$ consists of a single $\Bbbk$-rational point $p = (a_1, \dots, a_n)$. Then there is a natural isomorphism of $\mathbb{K}$-algebras

$$R := \mathbb{K}[x_1, \dots, x_n]/I\,\mathbb{K}[x_1, \dots, x_n] \cong \mathcal{O}_{\mathbb{A}^n,p}/I\mathcal{O}_{\mathbb{A}^n,p} =: R'.$$

Indeed, R is a local ring with maximal ideal $\overline{\mathfrak{m}} = \langle x_1 - a_1, \dots, x_n - a_n \rangle / I$. Hence, by the universal property of localization, $R = R_{\overline{\mathfrak{m}}}$. But $R_{\overline{\mathfrak{m}}} \cong R'$ by Exercise 4.2.3. □

Example 4.3.17. 1. In accordance with Example 4.3.13, we have

$$i(y, y - x^r; o) = r.$$

Indeed, by Remark 4.3.16,

$$\mathcal{O}_{\mathbb{A}^2,o}/\langle y, y - x^r \rangle \mathcal{O}_{\mathbb{A}^2,o} \cong \mathbb{C}[x, y]/\langle y, y - x^r \rangle \cong \mathbb{C}[x]/\langle x^r \rangle.$$

2. For the transversal cusps in Example 4.3.14, we get

$$i(y^2 - x^3, x^2 - y^3; o) = 4.$$

Indeed, since $1 - xy$ is a unit in $\mathcal{O}_{\mathbb{A}^2,o}$, we have

$$\langle y^2 - x^3, x^2 - y^3 \rangle = \langle y^2 - x^3, x^2 - x^3 y \rangle = \langle y^2 - x^3, x^2 \rangle = \langle y^2, x^2 \rangle \subset \mathcal{O}_{\mathbb{A}^2,o},$$

and the result follows as above from Remark 4.3.16. Similarly, for the tangential cusps,

$$i(y^2 - x^3, 2y^2 - x^3; o) = 6$$

since

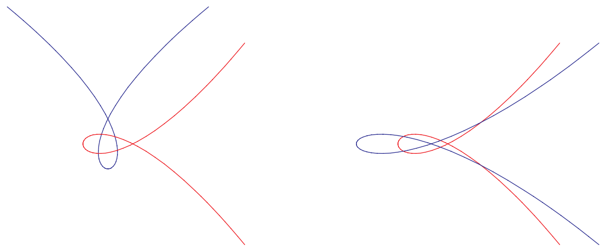
$$\langle y^2 - x^3, 2y^2 - x^3 \rangle = \langle y^2, x^3 \rangle \subset \mathcal{O}_{\mathbb{A}^2,o}.$$

To see this from the dynamical point of view, consider perturbed equations of type

$$y^2 - (x - c)^2 (x + c) = x^2 - (y - d)^2 (y + d) = 0$$

respectively

$$y^2 - (x - c)^2 (x + c) = 2y^2 - x^2 (x + d) = 0 :$$



4 intersection points

6 intersection points

□

Since we allow polynomials with multiple factors, it makes sense to extend some of the terminology used when working with curves to the more general case considered here. If $f \in \Bbbk[x,y]$ is a nonconstant polynomial, and $p \in \mathbb{A}^2$ is a point, we say that f **passes through $\boldsymbol{p}$** if $p \in \mathrm{V}(f)$. If $g \in \Bbbk[x,y]$ is another nonconstant polynomial, we say that f and g **intersect at $\boldsymbol{p}$** if $p \in \mathrm{V}(f) \cap \mathrm{V}(g)$ (equivalently, both multiplicities $\mathrm{mult}(f,p)$ and $\mathrm{mult}(g,p)$ are ≥ 1). We say that f and g **intersect transversally at $\boldsymbol{p}$** if $\mathrm{mult}(f,p) = \mathrm{mult}(g,p) = 1$ and the tangent line to f at p is different from the tangent line to g at p. Finally, if

$$f = \prod_{i=1}^{r} f_i^{e_i} \in \mathbb{K}[x,y]$$

is the decomposition of f into pairwise different irreducible factors f_i over $\mathbb{K}$, then each f_i is a **component of $\boldsymbol{f}$**, and e_i is the **multiplicity** of the component f_i.

Theorem 4.3.18 (Properties of Intersection Multiplicities). *Let $f, g \in \Bbbk[x,y]$ be nonconstant polynomials, and let $p = (a,b) \in \mathbb{A}^2$ be a point. Then:*

1. *$i(f,g;p) = 0$ iff f and g do not intersect at p.*
2. *$i(f,g;p) = \infty$ iff f and g have a common component passing through p.*
3. *$i(f,g;p) \geq \mathrm{mult}(f,p) \cdot \mathrm{mult}(g,p)$, with equality occuring iff f and g have no tangent line in common at p.*
4. *$i(f,g;p) = 1$ iff f and g intersect transversally at p.*
5. *$i(f,g;p) = i(g,f;p)$.*
6. *$i(f,g+hf;p) = i(f,g;p)$ for all $h \in \Bbbk[x,y]$.*
7. *If f is irreducible, and p is a smooth point of $C = \mathrm{V}(f) \subset \mathbb{A}^2$, then $i(f,g;p) = v_{C,p}(\overline{g})$, where $\overline{g} \in \mathbb{K}[C] \subset \mathcal{O}_{C,p}$ is the residue class of g.*
8. *$i(f,gh;p) = i(f,g;p) + i(f,h;p)$ for all $f,g,h \in \Bbbk[x,y]$.*

Proof. Parts 5 and 6 immediately follow from the definition. To show the remaining parts, we may suppose that all the components of f and g pass through p. Indeed, the other components are units in $\mathcal{O}_{\mathbb{A}^2,p}$ and do, hence, not contribute to $i(f,g;p)$. For simplicity, we write $\mathcal{O}_p = \mathcal{O}_{\mathbb{A}^2,p}$ and $\mathfrak{m}_p = \mathfrak{m}_{\mathbb{A}^2,p}$.

1. According to our definition, $i(f,g;p) = 0$ iff $\langle f,g \rangle \mathcal{O}_p = \mathcal{O}_p$. This, in turn, means that either f or g is a unit in $\mathcal{O}_p$ and, thus, that $p \notin \mathrm{V}(f) \cap \mathrm{V}(g)$.

2. If f and g have a common component h, then $\langle f,g \rangle \mathcal{O}_p \subset \langle h \rangle \mathcal{O}_p \subsetneq \mathcal{O}_p$. Hence, $i(f,g;p) \geq \dim_{\mathbb{K}} \mathcal{O}_p / \langle h \rangle \mathcal{O}_p$, and it suffices to show that the quotient of $\mathcal{O}_p$ modulo a proper principal ideal has infinite $\mathbb{K}$-dimension. We postpone the proof of this until we have formulated a version of Macaulay's Theorem 2.3.5 which holds in the ring $\mathcal{O}_p$. See Remark 4.4.24 in the next section.

For the converse, suppose that f and g have no common component. Then $\dim_{\mathbb{K}} \mathbb{K}[x,y]/\langle f,g \rangle$ is finite by Exercises 1.7.13 and 1.6.5. In particular, there is a unique $\langle x-a, y-b \rangle$-primary component of $\langle f,g \rangle \subset \mathbb{K}[x,y]$, which we denote by I. Then $\mathcal{O}_p / \langle f,g \rangle \mathcal{O}_p = \mathcal{O}_p / I\mathcal{O}_p$ (we will see this in Exercise 4.5.5, where we will study the behavior of primary decompositions under localization). Since, in turn, $\mathcal{O}_p/I\,\mathcal{O}_p \cong \mathbb{K}[x,y]/I$ by Remark 4.3.16, we conclude that $i(f,g;p) = \dim_{\mathbb{K}} \mathbb{K}[x,y]/I \leq \dim_{\mathbb{K}} \mathbb{K}[x,y]/\langle f,g\rangle < \infty$, as desired.

3. We will prove this part towards the end of the next section using Gröbner bases in the local case.

4. This special case of part 3 is easy to do directly. Indeed, applying Nakayama's lemma as in the proof of Corollary 4.2.20, we get: $i(f,g;p) = 1 \iff \langle f,g\rangle = \mathfrak{m}_p \iff \langle f,g\rangle + \mathfrak{m}_p^2 = \mathfrak{m}_p \iff \operatorname{span}(d_pf + \mathfrak{m}_p^2, d_pg + \mathfrak{m}_p^2\rangle = \mathfrak{m}_p/\mathfrak{m}_p^2$. Since $\mathfrak{m}_p/\mathfrak{m}_p^2$ is a two dimensional $\mathbb{K}$-vector space, $i(f,g;p) = 1$ iff d_pf and d_pg are $\mathbb{K}$-linearly independent, that is, iff C and D are smooth in p with different tangent lines.

7. According to our assumptions in this part, $\mathcal{O}_{C,p}$ is a DVR, with corresponding discrete valuation $v_{C,p}$ on $\mathbb{K}(C)$. Hence,

$$\mathcal{O}_p/\langle f,g\rangle\,\mathcal{O}_p \cong \mathcal{O}_{C,p}/\langle \overline{g}\rangle \cong \mathcal{O}_{C,p}/\langle t^k\rangle,$$

where $k = v_{C,p}(\overline{g})$. This shows the result since $\dim_{\mathbb{K}} \mathcal{O}_{C,p}/\langle t^k\rangle = k$ by Exercise 4.3.10.

8. Since the assertion follows from part 2 otherwise, we may suppose that f and gh have no common component. Consider, then, the sequence

$$0 \to \mathcal{O}_p/\langle f,h\rangle\,\mathcal{O}_p \xrightarrow{\phi} \mathcal{O}_p/\langle f,gh\rangle\,\mathcal{O}_p \xrightarrow{\psi} \mathcal{O}_p/\langle f,g\rangle\,\mathcal{O}_p \to 0, \qquad (4.4)$$

where ϕ is multiplication by g and ψ is induced by the identity on $\mathcal{O}_p$. By Exercise 2.8.4 on the additive behavior of $\mathbb{K}$-dimension, we are done if we show that (4.4) is exact.

For this, note that the syzygies on f,g over $\mathcal{O}_p$ are generated by the trivial syzygy $(g,-f)^t \in \mathcal{O}_p^2$. Indeed, given an $\mathcal{O}_p$-linear relation $Af+Bg=0$, choose a polynomial $u \in \mathbb{K}[x,y]$ with $u(p)=0$, and such that $a := uA \in \mathbb{K}[x,y]$ and $b := uB \in \mathbb{K}[x,y]$. Then $af + bg = 0 \in \mathbb{K}[x,y]$. Since $\mathbb{K}[x,y]$ is a UFD and f and g have no common component, b must be a multiple of f, so that $-b = cf$ for some $c \in \mathbb{K}[x,y]$. Then $(a,b)^t = c\cdot(g,-f)^t \in \mathbb{K}[x,y]^2$ and, thus, $(A,B)^t = C\cdot(g,-f)^t \in \mathcal{O}_p^2$, where $C = c/u$.

It follows that ϕ is injective: if $bg \in \langle f,gh\rangle\,\mathcal{O}_p$, say $bg = af + cgh$ with $a,c \in \mathcal{O}_p$, then $(a,-b+ch)^t$ is a syzygy on f,g, so that $b - ch \in f\mathcal{O}_p$ and, thus, $b \in \langle f,h\rangle\,\mathcal{O}_p$. Since, furthermore, ψ is surjective by its very definition, it remains to show that $\operatorname{im}\phi = \ker\psi$. This is completely straightforward and we leave it to the reader. □

Note that it are properties 6 and 8 which force us to allow polynomials with multiple factors in our definitions and statements. These properties are useful in that they often enable us to simplify the computation of intersection numbers. Let us, for instance, rewrite the last computation in Example 4.3.17. Property 6 (with the help of property 5) gives $i(y^2 - x^3, 2y^2 - x^3; o) = i(y^2, x^3; o)$. But $i(y^2, x^3; o) = 6$ by property 8. □

Exercise* **4.3.19.** Let $f \in \Bbbk[x,y]$ be a square-free polynomial, let $C = \mathrm{V}(f) \subset \mathbb{A}^2$ be the corresponding plane curve, and let $p \in C$ be a point.

1. Suppose that p is a double point at which C has precisely one tangent line L. Show that, then, $i(C, L; p) \geq 3$. We say that p is a **cusp** of C if $i(C, L; p) = 3$.
2. If p is the origin, and L is the x-axis, show that p is a cusp of C with tangent line L iff f is of type $f = ay^2 + bx^3 + \text{other terms of degree} \geq 3$, where $ab \neq 0$. □

4.4 Gröbner Bases in the Local Case

In this section, we will adjust the concept of Gröbner bases and Buchberger's algorithm to computations in the local ring of $\mathbb{A}^n$ at a given point of $\mathbb{A}^n$. This will, in particular, allow us to compute intersection multiplicities via Gröbner bases.

For our purposes, it is enough to consider the case where the given point is the origin $o \in \mathbb{A}^n$. Indeed, if $p = (a_1, \dots, a_n) \in \mathbb{A}^n$ is any point, we may translate p to o (on the level of rings, we have the isomorphism $\mathcal{O}_{\mathbb{A}^n,p} \cong \mathcal{O}_{\mathbb{A}^n,o}$ which extends the substitution homomorphism $\mathbb{K}[x_1, \dots, x_n] \to \mathbb{K}[x_1, \dots, x_n],\ x_i \mapsto x_i - a_i$). As usual, $\Bbbk \subset \mathbb{K}$ will be the ground field over which the generators of the ideals under consideration (and the originally given point p) are defined. Taking into account that Remark 2.7.1 on field extensions applies to the adjusted version of Buchberger's algorithm, too, we will be concerned with computations in the local ring

$$\mathcal{O}_o = \Bbbk[x_1, \dots, x_n]_{\langle x_1, \dots, x_n \rangle}.$$

Note that every ideal I of $\mathcal{O}_o$ can be generated by polynomials (choose any finite set of generators and clear denominators). Starting from a set of polynomial generators for I, the adjusted version of Buchberger's algorithm will compute a Gröbner basis for I consisting of polynomials, too. In fact, all computations in Buchberger's test will take place in the polynomial ring.

Reflecting the significance of the lowest degree terms of a polynomial f for local studies (as indicated by our treatment of singular points in the preceeding section), we will pick the leading term of f from among those terms. One way of making this precise would be to choose a degree-compatible monomial order such as the degree reverse lexicographic order, and pick the *least* term of f as the leading term. Pursuing an alternative approach, we will make use of monomial orders which are **degree-anticompatible**:

$$\deg x^\alpha < \deg x^\beta \Longrightarrow x^\alpha > x^\beta.$$

Example 4.4.1. The **local degree reverse lexicographic order** $>_{\mathrm{ldrlex}}$ on $\Bbbk[x_1, \dots, x_n]$ is defined by setting

$$x^\alpha >_{\text{ldrlex}} x^\beta \iff \deg x^\alpha < \deg x^\beta,\ \text{or } (\deg x^\alpha = \deg x^\beta \text{ and the last nonzero entry of } \alpha - \beta \in \mathbb{Z}^n \text{ is negative}).$$ □

A degree-anticompatible monomial order such as $>_{\text{ldrlex}}$ is never global. It is, in fact, local in the following sense:

Definition 4.4.2. A monomial order on $\Bbbk[x_1, \dots, x_n]$ is **local** if

$$x_i < 1 \quad \text{for} \quad i = 1, \dots, n.$$ □

Example 4.4.3. A weight order $>_w$ on $\Bbbk[x_1, \dots, x_n]$ is local iff the coefficients of w are strictly negative. □

Remark 4.4.4. Given a local monomial order $>$ on $\Bbbk[x_1, \dots, x_n]$, a polynomial $u \in \Bbbk[x_1, \dots, x_n]$ is a unit in $\mathcal{O}_o$ iff its leading monomial is 1. □

A drawback of local monomial orders is that they are not Artinian. As a consequence, the usual division process may not terminate. This is illustrated by Example 2.2.9 which we revisit now:

Example 4.4.5. In the case of one variable x, there is precisely one local monomial order:

$$1 > x > x^2 > \cdots$$

Dividing $g = x$ by $f_1 = x - x^2$ with respect to this order, we successively get the expressions $g = 1 \cdot f_1 + x^2$, $x^2 = x \cdot f_1 + x^3, \dots$. This may be interpreted by saying that the result of the division process, computed in *infinitely* many steps, is a standard expression whose quotient g_1 is the formal power series $\sum_{k=0}^{\infty} x^k$:

$$g = g_1 \cdot f_1 + 0 \in \Bbbk[[x]], \quad \text{where} \quad g_1 = \sum_{k=0}^{\infty} x^k. \tag{4.5}$$

On the other hand, expressing the fact that $1 - x$ is a multiplicative inverse to $\sum_{k=0}^{\infty} x^k$ in $\Bbbk[[x]]$, we have the **formal geometric series expansion**

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k.$$

We may, hence, rewrite (4.5) in a form which makes sense as an equation in the ring we are actually interested in:

$$g = \frac{1}{1-x} \cdot f_1 + 0 \in \Bbbk[x]_{\langle x \rangle}.$$

Multiplying both sides above by the unit $u = 1 - x \in \Bbbk[x]_{\langle x \rangle}$, we get the expression

$$u \cdot g = 1 \cdot f_1 + 0 \in \Bbbk[x] \tag{4.6}$$

which involves polynomials only. □

In what follows, we will discuss a division algorithm, designed by Mora (1982), which computes standard expressions such as (4.6). Based on this, we will formulate a version of Buchberger's criterion for $\mathcal{O}_o$. To prove the criterion, we will reduce to Buchberger's criterion for the formal power series ring $\mathbb{k}[[x_1, \dots, x_n]]$ (which, in turn, will be proved as in the polynomial case). Setting the stage for the reduction, we treat, now, power series expansion in general: given $f \in \mathcal{O}_o$, write f as a fraction of type $g/(1-h)$, with polynomials $g \in \mathbb{k}[x_1, \dots, x_n]$ and $h \in \langle x_1, \dots, x_n \rangle$, and set

$$f = \frac{g}{1-h} = g \sum_{k=0}^{\infty} h^k. \tag{4.7}$$

The crucial point is that the right hand side of (4.7) makes sense as an element of $\mathbb{k}[[x_1, \dots, x_n]]$. To verify this, we use a bit of topology.

Remark-Definition 4.4.6. Given any ring R and any ideal $\mathfrak{m}$ of R, it makes sense to define the **$\mathfrak{m}$-adic topology** on R by taking the cosets $f + \mathfrak{m}^k$ as a basis, where $f \in R$ and $k \geq 0$. The $\mathfrak{m}$-adic topology is Hausdorff iff $\bigcap_{k=0}^{\infty} \mathfrak{m}^k = \langle 0 \rangle$. Due to Krull's intersection theorem, this condition is, in particular, fulfilled if R is a local Noetherian ring with maximal ideal $\mathfrak{m}$. □

If we endow a ring R with the $\mathfrak{m}$-adic topology for some ideal $\mathfrak{m} \subset R$, we say that a sequence $(f_\nu) \subset R$ is a **Cauchy sequence** if for every $k \geq 0$, there exists a number ν_0 such that $f_\nu - f_\mu \in \mathfrak{m}^k$ for all $\nu, \mu \geq \nu_0$. In the same spirit, a **sequence** $(f_\nu) \subset R$ is called **convergent**, with **limes** f, if for every $k \geq 0$, there exists a number ν_0 such that $f_\nu - f \in \mathfrak{m}^k$ for all $\nu \geq \nu_0$. A **series** $\sum_{\nu=0}^{\infty} f_\nu$ in R is **convergent** if the sequence formed by its partial sums is convergent. If the $\mathfrak{m}$-adic topology is Hausdorff, every convergent sequence (f_ν) has a unique limes, denoted $\lim_{\nu \to \infty} f_\nu$. In particular, every convergent series constitutes, then, an element of R.

Definition 4.4.7. Given a ring R and an ideal $\mathfrak{m}$ of R, we say that R is **complete with respect to $\mathfrak{m}$** if the $\mathfrak{m}$-adic topology is Hausdorff, and if every Cauchy sequence converges. □

Proposition 4.4.8. *Let* $\mathfrak{m} = \langle x_1, \dots, x_n \rangle \subset \mathbb{k}[[x_1, \dots, x_n]]$. *Then:*

1. *The* $\mathfrak{m}$*-adic topology on* $\mathbb{k}[[x_1, \dots, x_n]]$ *is Hausdorff:*

$$\bigcap_{k=0}^{\infty} \mathfrak{m}^k = \langle 0 \rangle.$$

2. $\mathbb{k}[[x_1, \dots, x_n]]$ *is complete with respect to* $\mathfrak{m}$.
3. *A series* $\sum_{\nu=0}^{\infty} f_\nu$ *in* $\mathbb{k}[[x_1, \dots, x_n]]$ *converges with respect to the* $\mathfrak{m}$*-adic topology iff* $\lim_{\nu \to \infty} f_\nu = 0$.
4. $\mathbb{k}[[x_1, \dots, x_n]]$ *is a local ring with maximal ideal* $\mathfrak{m}$.

5. *There is a natural embedding of local rings $\mathcal{O}_o \subset \Bbbk[[x_1, \dots, x_n]]$ defined by power series expansion. The image of the maximal ideal of $\mathcal{O}_o$ under this embedding is contained in the maimal ideal $\mathfrak{m}$.*

Proof. 1. This is clear: if the power series $f = \sum a_\alpha x^\alpha$ is contained in $\mathfrak{m}^k$, then $a_\alpha = 0$ for all α with $|\alpha| < k$.

2. Given a Cauchy sequence $(f_\nu) = \left(\sum a_\alpha^{(\nu)} x^\alpha\right) \subset \Bbbk[[x_1, \dots, x_n]]$, define $f = \sum a_\alpha x^\alpha \in \Bbbk[[x_1, \dots, x_n]]$ as follows: for each $k \geq 1$, pick a number ν_0 such that $f_\nu - f_\mu \in \mathfrak{m}^k$ for all $\nu, \mu \geq \nu_0$, and set $a_\alpha = a_\alpha^{(\nu_0)}$ for all α with $|\alpha| = k - 1$. Then $f = \lim_{\nu \to \infty} f_\nu$.

3. This follows from part 2: the sequence formed by the partial sums of $\sum_{\nu=0}^{\infty} f_\nu$ is a Cauchy sequence iff $\lim_{\nu \to \infty} f_\nu = 0$.

4. We have to show that each element $f \in \Bbbk[[x_1, \dots, x_n]] \setminus \mathfrak{m}$ is a unit in $\Bbbk[[x_1, \dots, x_n]]$. For this, write $f = a_0 - h$, with $0 \neq a_0 \in \Bbbk$ and $h \in \mathfrak{m}$, and expand:

$$\frac{1}{a_0 - h} = \frac{1}{a_0} \sum_{k=0}^{\infty} (\frac{h}{a_0})^k.$$

Then, by part 3, the series on the right hand side converges and defines, thus, a multiplicative inverse to f.

5. This follows similarly: it is, now, clear that the series on the right hand side of (4.7) constitutes an element of $\Bbbk[[x_1, \dots, x_n]]$. □

Exercise* 4.4.9. Let S be a ring which is complete with respect to some ideal $\mathfrak{m}$. Given $s_1, \dots, s_n \in \mathfrak{m}$, show that there exists a unique homomorphism $\Phi : \Bbbk[[x_1, \dots, x_n]] \to S$ such that $\Phi(x_i) = s_i$ for all i. In fact, Φ is the map which sends a a power series f to the series $f(s_1, \dots, s_n) \in S$. As in the polynomial case, we refer to Φ as a **substitution homomorphism**. □

We, now, come to division with remainder and Gröbner bases in $\Bbbk[[x_1, \dots, x_n]]$. This topic is of theoretical interest and was first considered by Hironaka (1964) and, independently, Grauert (1972) who used the name **standard basis** instead of Gröbner basis. Our terminology will be the same as in Chapter 2. For instance, if $0 \neq f = \sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha \in \Bbbk[[x_1, \dots, x_n]]$, we call any $a_\alpha x^\alpha$ with $a_\alpha \neq 0$ a **term** of f. And, given a local monomial order $>$ on $\Bbbk[x_1, \dots, x_n] \subset \Bbbk[[x_1, \dots, x_n]]$, we define the **leading term** of f, written $\mathbf{L}(f) = \mathbf{L}_>(f)$, to be the largest term of f. This makes sense since every nonempty set X of monomials in $\Bbbk[x_1, \dots, x_n]$ has a *largest* element with respect to the *local* order $>$. Indeed, arguing as in the proof of Proposition 2.2.10, we may take the largest element of a finite set of monomial generators for the ideal $\langle X \rangle \subset \Bbbk[x_1, \dots, x_n]$. As usual, $\mathbf{L}_>(0) = \mathbf{L}(0) = 0$.

Since a global monomial order $>$ is Artinian, there is no sequence $(m_\nu)_{\nu \in \mathbb{N}}$ of monomials m_ν such that $m_1 > m_2 > \cdots$. In the local case, we have instead:

Lemma 4.4.10. *Let $\mathfrak{m} = \langle x_1, \dots, x_n \rangle$ be the maximal ideal of $\Bbbk[[x_1, \dots, x_n]]$, and let $>$ be a local monomial order on $\Bbbk[x_1, \dots, x_n] \subset \Bbbk[[x_1, \dots, x_n]]$.*

1. *If $(m_\nu)_{\nu\in\mathbb{N}}$ is a sequence of monomials in $\Bbbk[x_1,\dots,x_n]$ such that $m_1 > m_2 > \cdots$, then $\lim_{\nu\to\infty} m_\nu = 0$ with respect to the $\mathfrak{m}$-adic topology.*
2. *If $>$ is a local weight order $>_w$, and $(f_\nu)_{\nu\in\mathbb{N}}$ is a sequence of formal power series in $\Bbbk[[x_1,\dots,x_n]]$, then, with respect to the $\mathfrak{m}$-adic topology, we have:*

$$\lim_{\nu\to\infty} \mathbf{L}_{>_w}(f_\nu) = 0 \implies \lim_{\nu\to\infty} f_\nu = 0$$ □

Proof. Given k, only finitely many of the monomials in $\Bbbk[x_1,\dots,x_n]$ are not contained in $\mathfrak{m}^k$. In particular, there is an integer ν_0 such that $m_\nu \in \mathfrak{m}^k$ for all $\nu \geq \nu_0$. This shows part 1. For part 2, set

$$r = \min\{w(m) \mid m \text{ a monomial such that } m \notin \mathfrak{m}^k\}.$$

Then, if $\lim_{\nu\to\infty} \mathbf{L}_{>_w}(f_\nu) = 0$, there is a number ν_1 such that $w(\mathbf{L}_{>_w}(f_\nu)) < r$ for all $\nu \geq \nu_1$ (indeed, the coefficients of w are strictly negative by assumption). We conclude that $f_\nu \in \mathfrak{m}^k$ for all $\nu \geq \nu_1$, as desired. □

Theorem 4.4.11 (Grauert's Division Theorem). *Let $>$ be a local monomial order on $\Bbbk[x_1,\dots,x_n]$, write $R = \Bbbk[[x_1,\dots,x_n]]$, and let $f_1,\dots,f_r \in R\setminus\{0\}$. For every $g\in R$, there exists a uniquely determined expression*

$$g = g_1 f_1 + \ldots + g_r f_r + h, \text{ with } g_1,\dots,g_r,h \in R,$$

and such that:

(DD1) For $i > j$, no term of $g_i\,\mathbf{L}(f_i)$ is divisible by $\mathbf{L}(f_j)$.
(DD2) For all i, no term of h is divisible by $\mathbf{L}(f_i)$.

This expression is called a **Grauert standard expression** *for g with* **remainder** *h (in terms of the f_i, with respect to $>$).*

Proof. The *uniqueness* follows as in the polynomial case (see Theorem 2.2.12). For the *existence*, we first note that as in the polynomial case, the result clearly holds if $f_1,\dots,f_r$ are terms. In the general case, we get, thus, a unique expression

$$g^{(0)} := g = \sum_{j=1}^{r} g_j^{(0)}\,\mathbf{L}(f_j) + h^{(0)}$$

satisfying conditions (DD1) and (DD2). Then either $g^{(1)} := g - \sum_{j=1}^r g_j^{(0)} f_j - h^{(0)}$ is zero, and we are done, or $\mathbf{L}(g^{(0)}) > \mathbf{L}(g^{(1)})$. Recursively, we are either done in finitely many steps, or we get sequences $(g^{(\nu)})$, $(g_j^{(\nu)})$, $j = 0,\dots,r$, and $(h^{(\nu)})$ of formal power series such that, for all ν,

$$g^{(\nu+1)} = g - \sum_{j=1}^{r}\sum_{\mu=1}^{\nu} g_j^{(\mu)} f_j - \sum_{\mu=1}^{\nu} h^{(\mu)}.$$

In the latter case, the result will follow once we show that all our sequences converge to zero with respect to the $\langle x_1, \dots, x_n \rangle$-adic topology on $\Bbbk[[x_1, \dots, x_n]]$. For this, consider the monomial ideals $I_j \subset \Bbbk[x_1, \dots, x_n]$ generated by all terms of f_j except the $\mathbf{L}(f_j)$, $j = 1, \dots, r$. For each j, let X_j consist of the minimal (monomial) generators for I_j together with $\mathbf{L}(f_j)$. Then $X := \bigcup X_j$ is a finite set of monomials. By Exercise 2.2.11, there exists a local weight order $>_w$ on $\Bbbk[x_1, \dots, x_n]$ which coincides on X with the given local order $>$. Due to our construction of X, we have $\mathbf{L}_{>_w}(f_j) = \mathbf{L}_>(f_j)$ for all j. Hence, repeating the division process above with $>$ replaced by $>_w$, we get the same sequences $(g^{(\nu)})$, $(g_j^{(\nu)})$, and $(h^{(\nu)})$.

Since $\mathbf{L}(g^{(0)}) > \mathbf{L}(g^{(1)}) > \dots$, we have $\lim_{\nu \to \infty} \mathbf{L}(g^{(\nu)}) = 0$ by part 1 of Lemma 4.4.10. Then also $\lim_{\nu \to \infty} \mathbf{L}(g_j^{(\nu)}) = 0$ and $\lim_{\nu \to \infty} \mathbf{L}(h^{(\nu)}) = 0$ since $\mathbf{L}(g^{(\nu)}) \geq_w \mathbf{L}(g_j^{(\nu)} f_j) = \mathbf{L}(g_j^{(\nu)})\mathbf{L}(f_j)$ and $\mathbf{L}(g^{(\nu)}) \geq_w \mathbf{L}(h^{(\nu)})$ for all ν. We are, thus, done by part 2 of Lemma 4.4.10. □

Leading ideals, **standard monomials**, and **Gröbner bases** for ideals in $\Bbbk[[x_1, \dots, x_n]]$ are defined as for ideals in $\Bbbk[x_1, \dots, x_n]$. Making use of Gordan's lemma as in the polynomial case is one way of showing that $\Bbbk[[x_1, \dots, x_n]]$ **is Noetherian**. Furthermore, we have the following variant of Macaulay's Theorem 2.3.5:

Proposition 4.4.12. *Let $I \subset \Bbbk[[x_1, \dots, x_n]] =: R$ be an ideal, and let $>$ be a local monomial order on $\Bbbk[x_1, \dots, x_n]$. Then:*

1. *The standard monomials represent $\Bbbk$-linearly independent elements of R/I, and their residue classes generate a subspace of R/I which is dense with respect to the $\mathfrak{m}_{R/I}$-adic topology, where $\mathfrak{m}_{R/I}$ is the maximal ideal of R/I.*
2. *If $\dim_\Bbbk R/I < \infty$, the standard monomials represent a $\Bbbk$-vector space basis for R/I.*

Proof. 1. Let

$$\mathcal{B} := \{m + I \mid m \in R \text{ a standard monomial}\} \subset R/I,$$

and let W be the subspace of R/I generated by the elements of $\mathcal{B}$. Arguing as in the proof of Macaulay's Theorem 2.3.5, we find:

(a) The elements of $\mathcal{B}$ are $\Bbbk$-linearly independent.
(b) Given a power series $g \in R$, there is a power series $h = \sum_\alpha b_\alpha x^\alpha \in R$ whose terms involve only standard monomials, and such that $g + I = h + I$. In fact, h is uniquely determined by g, I, and $>$ as the remainder in a Grauert standard expression $g = \sum_{i=1}^r g_i f_i + h$, where $f_1, \dots, f_r$ is any Gröbner basis for I.

Statement (a) is precisely the first assertion of part 1 of the proposition. To show that W is dense in R/I, we note that in the situation of (b), given an integer $k \geq 0$, we have $h - \sum_{|\alpha|<k} b_\alpha x^\alpha \in \mathfrak{m}^k$, where $\mathfrak{m}$ is the maximal ideal of R. Hence, $g \equiv \sum_{|\alpha|<k} b_\alpha x^\alpha \mod I + \mathfrak{m}_{R/I}^k$, as desired.

If $\dim_{\Bbbk} R/I < \infty$, there are only finitely many standard monomials by (a). Hence, given $g \in R$, any power series h as in (b) is, in fact, a polynomial. Together with (a), this shows part 2. □

Definition 4.4.13. As in the polynomial case, we call the remainder h in the proof above the **normal form** of g mod I. □

Finally, we have a version of Buchberger's criterion for $\Bbbk[[x_1, \ldots, x_n]]$ whose statement and proof read word for word identically to what we did in the polynomial case (in particular, in the statement of the criterion, it is enough to consider standard expressions in the weak sense of Remark 2.2.16). We leave the details to the reader:

Exercise* 4.4.14. Let $R = \Bbbk[[x_1, \ldots, x_n]]$.

1. Formulate and prove versions of Grauert's division theorem and Buchberger's criterion for free R-modules.
2. Show that Hilbert's syzygy theorem holds for R: Every finitely generated R-module M has a finite free resolution of length at most n, by finitely generated free R-modules. □

As is already clear from Example 4.4.5, this does not give us an algorithm for computing Gröbner bases in power series rings: even if we start with polynomials, the remainder on Grauert division may be a power series, and it may take infinitely many steps to compute this series.

Next, we turn from $\Bbbk[[x_1, \ldots, x_n]]$ to $\mathcal{O}_o$. To begin with, we show by example that the strong condition (DD2) of Grauert's division theorem cannot always be achieved in $\mathcal{O}_o$:

Example 4.4.15. Consider the polynomials $f = x$ and $f_1 = x - x^2 - y$ in $\Bbbk[x, y] \subset \Bbbk[[x, y]]$, and fix a local monomial order $>$ on $\Bbbk[x, y]$ such that $\mathbf{L}(f_1) = x$ (for instance, take $>_{\text{ldrlex}}$). Suppose there is a standard expression $x = g_1 f_1 + h$ as in Grauert's division theorem, with $g_1, h \in \Bbbk[x, y]_{\langle x,y \rangle}$. Then no term of the remainder h is divisible by $\mathbf{L}(f_1) = x$. That is, $h \in \Bbbk[y]_{\langle y \rangle}$. This implies that $x = \mathbf{L}(x) = \mathbf{L}(g_1 f_1) = \mathbf{L}(g_1) \cdot x$ and, thus, that g_1 is a unit in $\Bbbk[x, y]_{\langle x,y \rangle}$ (that is, $g(0, 0) \neq 0$). Furthermore, substituting h for x in $x = g_1 f_1 + h$, we get the equality

$$g_1(h, y) \cdot (h - h^2 - y) = 0 \in \Bbbk[y]_{\langle y \rangle}.$$

On the other hand, since f and f_1 vanish at the origin, h cannot have a constant term. It follows that $g_1(h, y) \neq 0$ since $g(0, 0) \neq 0$. We conclude that

$$h - h^2 - y = 0. \tag{4.8}$$

This is impossible since regarding (4.8) as a quadratic equation in h and solving it, we do not get a rational function: $h = \frac{1 \pm \sqrt{1-4y}}{2}$. Arguing more formally (supposing that h does exist as a rational function), write h as a fraction $h = \frac{h_1}{1+h_2}$, with polynomials $h_1 \in \Bbbk[y]$ and $h_2 \in \langle y \rangle \subset \Bbbk[y]$. Then, from (4.8), we obtain

$$(1+h_2) \cdot h_1 - h_1^2 - y \cdot (1+h_2)^2 = 0 \in \Bbbk[y]. \tag{4.9}$$

A check on degrees gives a contradiction as follows: If $\deg h_1 \geq 1 + \deg h_2$, then $\deg h_1^2 > 1 + \deg(h_2^2) = \deg(y \cdot (1+h_2^2))$ and $\deg h_1^2 > \deg((1+h_2) \cdot h_1)$. If $\deg h_2 \geq \deg h_1$, then $\deg((1+h_2^2) \cdot y) > \deg((1+h_2) \cdot h_1) \geq \deg h_1^2$. Hence, in both cases, the degree of one of the three summands on the left hand side of (4.9) is strictly larger than the degree of any other summand, absurd. □

Our discussion of division with remainder and Gröbner bases in $\mathcal{O}_o$ is motivated by what we did in Example 4.4.5. Taking additionally into account that every ideal in $\mathcal{O}_o$ is generated by polynomials, our statements will be formulated such that they involve polynomial data only.

Theorem 4.4.16 (Mora's Division Theorem). *Let $>$ be a monomial order on $\Bbbk[x_1, \ldots, x_n]$, and let $f_1, \ldots, f_r \in \Bbbk[x_1, \ldots, x_n] \setminus \{0\}$. For every $g \in \Bbbk[x_1, \ldots, x_n]$, there exists an expression*

$$u \cdot g = g_1 f_1 + \ldots + g_r f_r + h,$$

where $u, g_1, \ldots, g_r, h \in \Bbbk[x_1, \ldots, x_n]$, with $\mathbf{L}(u) = 1$, and such that:

(ID1) $\mathbf{L}(g) \geq \mathbf{L}(g_i f_i)$ *whenever both sides are nonzero.*
(ID2) *If h is nonzero, then $\mathbf{L}(h)$ is not divisible by any $\mathbf{L}(f_i)$.*

Every such expression is called a **Mora standard expression** *for g with* **remainder** *h (in terms of the f_i, with respect to $>$).* □

The proof of the theorem consists of an algorithm for computing Mora standard expressions. In comparison with the division algorithms discussed in Chapter 2, the crucial new idea of Mora is to not only divide by $f_1, \ldots, f_r$, but also by some of the intermediate dividends. To decide whether an intermediate dividend should be stored as a possible divisor for division steps still to come, its ecart will be computed.

Definition 4.4.17. Let $>$ be a monomial order on $\Bbbk[x_1, \ldots, x_n]$. Given a nonzero polynomial $f \in \Bbbk[x_1, \ldots, x_n]$, the **ecart** of f (with respect to $>$), written $\operatorname{ecart}(f)$, is defined to be

$$\operatorname{ecart}(f) = \deg f - \deg \mathbf{L}(f).$$

□

In stating Mora's division algorithm, we focus on the computation of the remainder h. How to compute the unit u and the quotients g_i (this requires some extra bookkeeping) will be described in the correctness argument given in the proof below.

Algorithm 4.4.18 (Mora's Division Algorithm). *Let $>$ be a monomial order on $\Bbbk[x_1,\dots,x_n]$. Given nonzero polynomials $g, f_1,\dots,f_r \in \Bbbk[x_1,\dots,x_n]$, compute a remainder h of g on Mora division by $f_1,\dots,f_r$.*

```
1. Set $h := g$ and $D := \{f_1,\dots,f_r\}$.
2. while $\big(h \neq 0$ and $D(h) := \{f \in D \mid \mathbf{L}(h)$ is divisible by $\mathbf{L}(f)\} \neq \emptyset\big)$
       • choose $f \in D(h)$ with $\operatorname{ecart}(f)$ minimal;
       • if $(\operatorname{ecart}(f) > \operatorname{ecart}(h))$ then $D := D \cup \{h\}$;
       • set $h := h - \frac{\mathbf{L}(h)}{\mathbf{L}(f)} f$.
3. return($h$).
```

□

Remark 4.4.19. 1. If we apply Mora's algorithm to homogeneous polynomials $g, f_1,\dots,f_r$, all polynomials computed in the resulting division process are homogeneous, too. Hence, all ecart's are zero, and Mora's algorithm follows the steps of an indeterminate version of the usual division algorithm. In fact, as shown by the correctness argument in the proof below, the algorithm computes a standard expression of type $g = g_1 f_1 + \ldots + g_r f_r + h$.

2. If $>$ is a global monomial order, and $\mathbf{L}(h)$ is a multiple of $\mathbf{L}(f)$, then $\mathbf{L}(h) \geq \mathbf{L}(f)$. Hence, even if added to D in the division process, h will not be used in further division steps. Thus, we obtain again an indeterminate version of the usual division algorithm, but in the nonhomogeneous case, the freedom of choice is reduced. □

Proof (of termination and correctness). We write D_k and h_k respectively for the set of intermediate divisors and the intermediate dividend after the kth iteration of the `while` loop, starting with $D_0 = D$ and $h_0 = g$.

Termination. We proceed in two steps. In the first step, we show that the set D of divisors will be enlarged in at most finitely many iterations of the `while` loop. Then, taking our cue from the remark above, we homogenize with respect to an extra variable x_0 to reduce to the termination result for the usual division algorithm.

After k iterations, the algorithm continues with the `while` loop iff $0 \neq \mathbf{L}(h_k) \in \langle \mathbf{L}(f) \mid f \in D_k \rangle \subset \Bbbk[x_1,\dots,x_n]$. In this case, h_k is added to D_k iff $x_o^{\operatorname{ecart}(h_k)}\mathbf{L}(h_k)$ is not contained in the monomial ideal

$$I_k = \langle x_o^{\operatorname{ecart}(f)}\mathbf{L}(f) \mid f \in D_k \rangle \subset \Bbbk[x_0,\dots,x_n].$$

By Gordan's lemma, the ascending chain $I_1 \subset I_2 \ldots$ is eventually stationary, say $I_N = I_{N+1} = \cdots$ for some N. Then also $D_N = D_{N+1} = \cdots$. Say, $D_N = \{f_1,\dots,f_{r'}\}$.

Termination will follow once we show that after finitely many further iterations, either $h = 0$ or $D(h) = \emptyset$. For this, homogenize h_{N+1} and the f_i with respect to x_0: set

$$H_{N+1} = x_0^{\deg(h_{N+1})} h_{N+1}(x_1/x_0, \dots x_n/x_0) \text{ and}$$

$$F_i = x_0^{\deg(f_i)} f_i(x_1/x_0, \dots x_n/x_0),\ i = 1, \dots, r'.$$

On $\Bbbk[x_0, \dots, x_n]$, consider the monomial order $>_g$ defined by setting

$$\begin{aligned} x_o^c x^\alpha >_g x_o^d x^\beta \iff & \deg x_o^c x^\alpha > \deg x_o^d x^\beta, \text{ or} \\ & (\deg x_o^c x^\alpha = \deg x_o^d x^\beta \text{ and } x^\alpha > x^\beta). \end{aligned}$$

This order is global, and we have $\mathbf{L}_{>_g}(F_i) = x_o^{\operatorname{ecart}(f_i)} \mathbf{L}_>(f_i)$. Thus, if we divide h_{N+1} by the f_i, Mora's algorithm follows the steps of an indeterminate version of the division algorithm, as desired.

Correctness. Recursively, starting with $u_0 = 1$ and $g_i^{(0)} = 0$, $i = 1, \dots, r$, suppose that, due to the first $k-1$ iterations of the while loop, we already have expressions of type

$$u_\ell \cdot g = g_1^{(\ell)} f_1 + \ldots + g_r^{(\ell)} f_r + h_\ell, \quad \text{with } \mathbf{L}(u_\ell) = 1,$$

$\ell = 0, \dots, k-1$. Then, if the test condition for the k-th iteration of the while loop is fulfilled, choose a polynomial $f = f^{(k)}$ as in the statement of the algorithm, and set $h_k = h_{k-1} - m_k f^{(k)}$, where $m_k = \frac{\mathbf{L}(h_{k-1})}{\mathbf{L}(f^{(k)})}$. There are two possibilities: either,

(a) $f^{(k)}$ is one of $f_1, \dots, f_r$, or
(b) $f^{(k)}$ is one of $h_0, \dots, h_{k-1}$.

Accordingly, substituting $h_k + m_k f^{(k)}$ for h_{k-1} in the expression for $u_{k-1} \cdot g$, we get an expression of type

$$u_k \cdot g = g_1^{(k)} f_1 + \ldots + g_r^{(k)} f_r + h_k,$$

where either,

(a) $u_k = u_{k-1}$, or
(b) $u_k = u_{k-1} - m_k u_\ell$, for some ℓ.

In any case, $\mathbf{L}(u_k) = \mathbf{L}(u_{k-1}) = 1$ (in case (b), note that $\mathbf{L}(h_l) > \mathbf{L}(h_{k-1}) = \mathbf{L}(m_k \cdot h_\ell) = m_k \cdot \mathbf{L}(h_\ell)$, so that $\mathbf{L}(u_{k-1}) = 1 > m_k = \mathbf{L}(m_k \cdot u_\ell)$). We conclude that, upon termination, the algorithm outputs a Mora standard expression as desired (that the conditions (ID1) and (ID2) are fulfilled is clear). □

Example 4.4.20. Dividing $g = x$ by $f_1 = x - x^2$ with respect to the unique local monomial order on $\Bbbk[x]$, we successively get:

$$h_0 = x,\ D_0 = \{x - x^2\},\ 1 \cdot g = 0 \cdot f_1 + x,$$

$$f^{(1)} = x - x^2,\ D_1 = \{x - x^2, x\},\ h_1 = x^2,\ 1 \cdot g = 1 \cdot f_1 + x^2,$$

and

$$f^{(2)} = x,\ h_1 = 0,\ (1 - x) \cdot g = 1 \cdot f_1 + 0.$$ □

Exercise 4.4.21. Consider $>_{\text{ldrlex}}$ on $\Bbbk[x,y,z]$ and compute a Mora standard expression for $g = x^3y + x^5 + x^2y^2z^2 + z^6$ in terms of $f_1 = x^2 + x^2y$, $f_2 = y^3 + xyz$, $f_3 = x^3y^2 + z^4$. □

We, now, come to Gröbner bases in $\mathcal{O}_o$. Let $>$ be a local monomial order on $\Bbbk[x_1,\dots,x_n]$. Considering the embedding $\mathcal{O}_o \subset \Bbbk[[x_1,\dots,x_n]]$, we define the **leading term** of an element $f \in \mathcal{O}_o$, written $\mathbf{L}(f) = \mathbf{L}_>(f)$, to be the leading term of its power series expansion. Given an ideal $I \subset \mathcal{O}_o$, the **leading ideal** of I is the monomial ideal $\mathbf{L}(I) = \mathbf{L}_>(I) \subset \Bbbk[x_1,\dots,x_n]$ generated by the leading terms of the elements of I. **Standard monomials** and **Gröbner bases** for ideals in $\mathcal{O}_o$ are defined as in the polynomial case. In fact, we ask that the Gröbner basis elements are polynomials (otherwise, clear denominators). Based on Mora Division with remainder, we get the $\mathcal{O}_o$ analog of Buchberger's Criterion 2.3.9:

Theorem 4.4.22 (Buchberger's Criterion for $\mathcal{O}_o$). *Let $>$ be a local monomial order on $\Bbbk[x_1,\dots,x_n]$, and let $f_1,\dots,f_r \in \Bbbk[x_1,\dots,x_n] \setminus \{0\}$. For every $i = 2,\dots,r$ and every minimal monomial generator x^α for*

$$M_i = \langle \mathbf{L}(f_1),\dots,\mathbf{L}(f_{i-1})\rangle : \mathbf{L}(f_i) \subset \Bbbk[x_1,\dots,x_n],$$

choose an S-polynomial $\mathrm{S}(f_i,f_j)$ as in Buchberger's Criterion 2.3.9. Then $f_1,\dots,f_r$ form a Gröbner basis iff any such $\mathrm{S}(f_i,f_j)$ has a Mora standard expression with remainder zero.

Proof. The condition on the remainders is clearly necessary. It is also sufficient. Indeed, considering the syzygies arising from the Mora standard expressions with remainder zero and arguing as in the proof of Buchberger's criterion 2.3.9, we find for every nonzero $g \in I = \langle f_1,\dots,f_r\rangle \subset \mathcal{O}_o \subset \Bbbk[[x_1,\dots,x_n]]$ a Grauert standard expression in terms of the f_k with remainder zero. Hence, $\mathbf{L}(g)$ is divisible by one of the $\mathbf{L}(f_k)$. □

The $\mathcal{O}_o$ analog of Macaulay's Theorem 2.3.5 is part 2 below:

Proposition 4.4.23. *Let $>$ be a local monomial order on $\Bbbk[x_1,\dots,x_n]$. Then:*

1. *Let I be an ideal of $\mathcal{O}_o$, and let $f_1,\dots,f_r \in I$ be polynomials. Then the f_k form a Gröbner basis for I iff they form a Gröbner basis for the extended ideal $I\,\Bbbk[[x_1,\dots,x_n]]$.*
2. *Proposition 4.4.12 on standard monomials remains true if $\Bbbk[[x_1,\dots,x_n]]$ is replaced by $\mathcal{O}_o$.*

Proof. Let $I^e = I\,\Bbbk[[x_1,\dots,x_n]]$.

1. The implication from right to left is clear: Since $I \subset I^e$, we also have $\mathbf{L}(I) \subset \mathbf{L}(I^e)$.

Conversely, suppose that the f_k form a Gröbner basis for I. Then $f_1,\dots,f_r$ generate I and, hence, also I^e. It is, thus, enough to show that the f_k form a

Gröbner basis in $\Bbbk[[x_1,\dots,x_n]]$. By assumption, the f_k satisfy Buchberger's criterion for $\mathcal{O}_o$. That is, we have Mora standard expressions of type

$$u \cdot \mathrm{S}(f_i, f_j) = \sum g_k f_k,$$

where u is a unit in $\mathcal{O}_o$. Multiplying both sides by the power series expansion of the inverse of u, we get a standard expression (in the weak sense of Remark 2.2.16) for $\mathrm{S}(f_i, f_j)$ in $\Bbbk[[x_1,\dots,x_n]]$ with remainder zero. Hence, Buchberger's criterion is satisfied in the power series ring as well.

2. Let I be an ideal of $\mathcal{O}_o$. Given an element $g \in \mathcal{O}_o \subset \Bbbk[[x_1,\dots,x_n]]$, we consider the remainder $h = \sum_\alpha b_\alpha x^\alpha$ in a Grauert standard expression $g = \sum_{i=1}^r g_i f_i + h$, where $f_1,\dots,f_r$ is any Gröbner basis for I (and, thus, also for I^e by part 1). Then, if $\mathfrak{m}_o$ denotes the maximal ideal of $\mathcal{O}_o$, and k is an integer ≥ 0, we can replace h modulo $I + \mathfrak{m}_o^k$ by the polynomial $\sum_{|\alpha|<k} b_\alpha x^\alpha$. Arguing as in the proof of Proposition 4.4.12, we are done. □

The result on standard monomials gives us in particular:

Remark 4.4.24. If $n > 1$, and $\langle f\rangle \subsetneq \mathcal{O}_o$ is a proper principal ideal, then $\dim_\Bbbk \mathcal{O}_o/\langle f\rangle = \infty$ since there are infinitely many standard monomials for $\langle f\rangle$. This concludes the proof of part 2 of Theorem 4.3.18. □

As in Chapter 2, Buchberger's criterion gives us **Buchberger's test** and **Buchberger's algorithm** for computing Gröbner bases (being able to compute remainders, the termination of the algorithm only relies on the ascending chain condition for monomial ideals, but not on the fact that the given order is Artinian; see Corollary 2.3.11).

Exercise 4.4.25. Consider $\Bbbk[x,y]$ with $>_{\mathrm{ldrlex}}$. Compute Gröbner bases for the following ideals:

$$I = \langle x^3 - y^3, x^2y^2\rangle,\ J = \langle x^3 - y^3, x^2y^2 + xy^3\rangle,\ \text{ and }\ K = \langle x^3 - y^4, x^2y^2\rangle.$$

Hint: You should get

$$\{x^3-y^3, x^2y^2, y^5\},\ \{x^3-y^3, x^2y^2+xy^3, xy^4-y^5, y^6\},\ \text{ and }\ \{x^3-y^4, x^2y^2, y^6\}.$$

In the proof below, we will make use of the ideals I, J, and K to illustrate the main arguments by examples. □

Proof of Theorem 4.3.18, Part 3. Let $f, g \in R = \Bbbk[x,y]$ be nonconstant polynomials, let $m = \mathrm{mult}(f,o)$ and $n = \mathrm{mult}(f,o)$ be their multiplicities at the origin o, and let f_m and g_n be the homogeneous components of f and g of degrees m and n, respectively. We have to show that $i(f,g;o) \geq m\cdot n$, with equality occuring iff f and g have no tangent line in common at o. This is clear if $i(f,g;o) = \infty$. Writing $I_o = \langle f,g\rangle\mathcal{O}_o$, we may, therefore, assume that

$$i(f,g;o) = \dim_\Bbbk \mathcal{O}_o/I_o < \infty. \tag{4.10}$$

By part 2 of Theorem 4.3.18, the geometric meaning of this is that f and g do not have a common component passing through o.

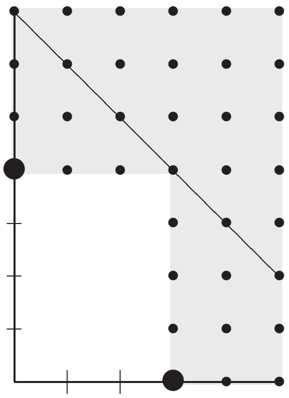
Given any local monomial order on $\Bbbk[x,y]$, it follows from (4.10) and part 2 of Proposition 4.4.23 that $i(f,g;p)$ is precisely the number of standard monomials for I_o. To compute this number, we fix the local degree reverse lexicographic order $>_{\text{ldrlex}}$. Then, since $>_{\text{ldrlex}}$ is degree-anticompatible, the leading terms $\mathbf{L}(f)$ and $\mathbf{L}(g)$ are among the terms of f_m and g_n, respectively. We may, hence, choose the coordinates such that $\mathbf{L}(f) = x^m$ and, then, suppose that $\mathbf{L}(g)$ is of type $\mathbf{L}(g) = x^{\beta_1}y^{\beta_2}$, where $m > \beta_1$ and $\beta_1 + \beta_2 = n$ (subtract a multiple of f from g and adjust constants, if necessary). To proceed, we distinguish two cases.

Case 1: Suppose f and g are homogeneous. That is, $f = f_m$ and $g = g_n$. Then f and g have no common tangent line at o (every such line would be a common component of f and g at o). Hence, in this case, we have to show that the number of standard monomials for I_o is $m \cdot n$.

If $\beta_1 = 0$, we are done right away: Since

$$\mathrm{S}(g,f) \in \langle x,y\rangle^d \subset \langle \mathbf{L}(f), \mathbf{L}(g)\rangle, \tag{4.11}$$

where d is the degree of the "corner" $\mathrm{LCM}(\mathbf{L}(g), \mathbf{L}(f)) = x^m y^n$, the remainder of $\mathrm{S}(g,f)$ on Mora division by f,g is zero. Hence, f,g form a Gröbner basis for I_o, and the monomials $x^{\alpha_1}y^{\alpha_2}$ with $0 \le \alpha_1 \le m-1$ and $0 \le \alpha_2 \le n-1$ are precisely the standard monomials:



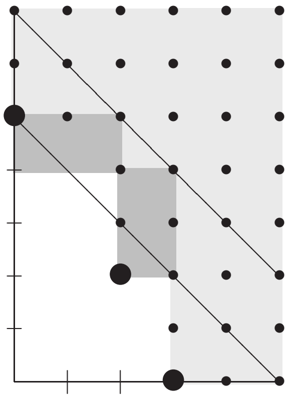
2 Gröbner basis elements

If $\beta_1 > 0$, however, then f,g do not form a Gröbner basis since this would imply that there are infinitely many standard monomials. Hence, the remainder of $S(g,f) = x^{(m-\beta_1)}g - y^{\beta_2}f$ on Mora division by f,g is nonzero and gives a new (homogeneous) Gröbner basis element h_3 for I whose leading term is a scalar times a monomial of type $x^{\gamma_1}y^{\gamma_2}$, with $\beta_1 > \gamma_1$ and $\gamma_1 + \gamma_2 = m + \beta_2$.

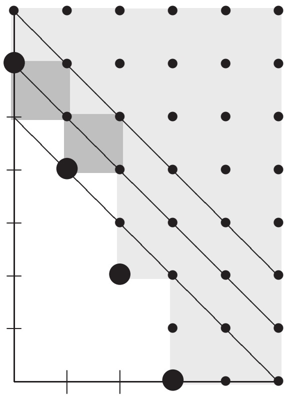
Applying Buchberger's criterion to f,g,h_3, the only new S-polynomial to be tested is $S(h_3,g)$ since $x^{m-\gamma_1}$ is divisible by $x^{\beta_1-\gamma_1}$. If nonzero, we add the remainder arising from this test to the set of generators and continue in this way. The resulting process yields (homogeneous) Gröbner basis elements $h_1 = f, h_2 = g, h_3, \ldots,$ where, at each stage of the process, only $S(h_k, h_{k-1})$ needs to be tested, and where the degree of the new generator h_{k+1} coincides with that of the "corner" $\mathrm{LCM}(\mathbf{L}(h_k), \mathbf{L}(h_{k-1}))$.

Eventually, we will get an element h_r such that $\mathbf{L}(h_r)$ is a scalar times a power of y. Then the remainder of $S(h_r, h_{r-1})$ on Mora division by the h_k is zero by reasons of degree (as in (4.11)). Hence, $h_1, \ldots, h_r$ form a Gröbner basis for I_0.

In visualizing the process just described, we may say that the leading monomials of the h_k determine a staircase which connects the x-axis with the y-axis. An elementary inductive argument shows that the area under the stairs has size $m \cdot n$, as in the case where $\beta_1 = 0$:



3 Gröbner basis elements

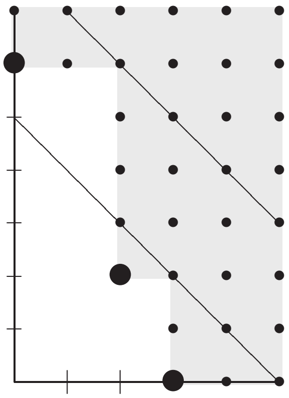


4 Gröbner basis elements

Case 2: Let, now, f and g be nonhomogeneous. As above, by computing a Gröbner basis $h_1 = f, h_2 = g, \ldots, h_r$ for I_o, we get a staircase of leading terms which connects the x-axis with the y-axis. Now, however, the Gröbner basis elements are not necessarily homogeneous. Let $\widetilde{h}_{k+1}$ be the part of h_{k+1} of degree $\deg \mathrm{LCM}(\mathbf{L}(h_k), \mathbf{L}(h_{k-1}))$, and let s be the least number k such that $\widetilde{h}_{k+1}$ is zero. Then $\widetilde{h}_1, \ldots, \widetilde{h}_s$ form a Gröbner basis for $\langle f_m, g_n \rangle \mathcal{O}_o$ such that $\mathbf{L}(h_k) = \mathbf{L}(\widetilde{h}_k)$ for all $k \leq s$ (recall that $>_{\mathrm{ldrlex}}$ is degree-anticompatible). We, hence, have two possibilities:

Case 2a: If f and g do not have a common tangent line at o, the $\mathbf{L}(\widetilde{h}_k)$ must reach the y-axis as well, which means that the staircase arising from f_m, g_n coincides with that arising from f, g. Then, again, there are precisely $m \cdot n$ standard monomials for I_o.

Case 2b: If, however, f and g do have a common tangent line at o, we must have $s < r$. Then $\deg \mathbf{L}(h_{s+1}) > \deg \mathrm{LCM}(\mathbf{L}(h_s), \mathbf{L}(h_{s-1}))$, so that for the staircase arising from f, g, the area under the stairs has size $> m \cdot n$:



This concludes the proof of Theorem 4.3.18. □

Exercise* 4.4.26 (Multiplicities in Terms of the Local Ring). Let $f \in \Bbbk[x_1, \ldots, x_n]$ be a nonconstant polynomial, let $p \in \mathbb{A}^n$ be a point, and let R be the local ring $R = \mathcal{O}_{\mathbb{A}^n,p}/\langle f \rangle\, \mathcal{O}_{\mathbb{A}^n,p}$ with its maximal ideal $\mathfrak{m}_R$. The **multiplicity of f at p**, written $\operatorname{mult}(f,p)$, is defined to be

$$\operatorname{mult}(f,p) = \min\{k \mid \dim_{\mathbb{K}} R/\mathfrak{m}_R^{k+1} < \binom{n+k}{k}\}.$$

Show that $\operatorname{mult}(f,p) \geq 1$ iff $p \in \mathrm{V}(f)$. If f is square-free, show that $\operatorname{mult}(f,p) = 1$ iff p is a smooth point of $\mathrm{V}(f)$. In case $n = 1$, show that $\operatorname{mult}(f,p)$ is the usual multiplicity of p as a root of f. In the case of plane curves, show that the definition of multiplicity given here coincides with the one given in Definition 4.3.2. □

We conclude this section with some remarks on convergent power series. Recall that in case $\Bbbk = \mathbb{C}$ (or $\Bbbk = \mathbb{R}$), a power series $f = \sum_\alpha f_\alpha x^\alpha \in \mathbb{C}[[x_1, \ldots, x_n]]$ is convergent if there exist a polyradius $\rho = (\rho_1, \ldots, \rho_n) \in \mathbb{R}^n_{>0}$ such that the series

$$||f||_\rho = \sum_\alpha |f_\alpha| \rho_1^{\alpha_1} \cdots \rho_n^{\alpha_n} < \infty$$

In this case, f converges absolutely on the polydisc $D_\rho = \{|x_1| \leq \rho_1, \ldots, |x_n| \leq \rho_n\}$ and $R_\rho = \{f \mid ||f||_\rho < \infty\}$ is a Banach space.

The set of convergent power series is a ring which we denote by $\mathbb{C}\{x_1, \ldots, x_n\}$. We, then, have a chain of ring inclusions

$$\mathbb{C}[x_1, \ldots, x_n] \subset \mathcal{O}_{\mathbb{A}^n(\mathbb{C}),o} \subset \mathbb{C}\{x_1, \ldots, x_n\} \subset \mathbb{C}[[x_1, \ldots, x_n]].$$

Proposition 4.4.27. *Let $>$ be a local monomial order on $\mathbb{C}[x_1, \ldots, x_n]$. If $g, f_1, \ldots, f_r$ are convergent power series, and $g = \sum g_i f_i + h$ is the unique exprssion satisfying the conditions (DD1) and (DD2) of Grauert's division* *theorem, then the g_i and h are convergent, too. In particular, the reduced Gröbner basis of an ideal in $\mathbb{C}\{x_1,\dots,x_n\}$ generated by convergent power series consists of convergent power series, too.* □

Proof. Let $>_w$ be local weight order on $\mathbb{C}[x_1,\dots,x_n]$ given by $\mathbb{Q}$-linear independent negative weights such that $\mathbf{L}_w(f_i) = \mathbf{L}(f_i)$. Without of generality we assume that the f_i are monique, say $\mathbf{L}(f_j) = x^{\alpha^j}$. Consider tupels

$$K = \{(g_1,\dots,g_r,h) \in \mathbb{C}[[x_1,\dots,x_n]]^{r+1} \mid \text{ satisfying condition DD2}\}$$

and the subspace K_ρ of tuples, which have finite norm

$$||(g_1,\dots,g_r,h)||_\rho := \sum ||g_i||_\rho \rho^{\alpha^j} + ||h||_\rho < \infty$$

Then the map

$$\psi : K_\rho \to R_\rho, (g_1,\dots,g_r,h) \mapsto \sum g_i x^{\alpha^i} + h$$

is an isometrie of Banach spaces. We claim that for suitable ρ the pertubation

$$\phi : K_\rho \to R_\rho, (g_1,\dots,g_r,h) \mapsto \sum g_i f_i + h$$

is still an isomorphism. For this we consider the weight order given by w and a polyradius $\rho(\tau) = (\tau^{-w_1},\dots,\tau^{-w_n})$ for $0 < \tau << 1$ such that $g, f_1,\dots,f_r$ converge in $D_{\rho(\tau)}$ and $q = \sum_i ||f_i - ini_w(f_i)||_{\rho(\tau)} \rho(\tau)^{-alpha^i} < 1$. Then $\phi \circ \psi^{-1} = id_{R_{\rho(\tau)}} + \epsilon$ with operator norm $||\epsilon||_{\rho(\tau)} \leq q < 1$. Hence $\sum_k (-1)^k \epsilon^k$ is a convergent series of operators, which gives $(id_{R_rho(\tau)} + \epsilon)^{-1}$.

Thus given $g \in \mathbb{C}\{x_1,\dots,x_n\}$ we can choose $0 < \tau << 1$ such that additionally $g \in R_{\rho(\tau)}$. Then $g_1, dots, g_r$ and h converge in this polydisc as well.

The rings $\Bbbk[[x_1,\dots,x_n]]$ and $\mathbb{C}\{x_1,\dots,x_n\}$. As for the polynomial ring, the proof uses induction and Gauss‘ Lemma., utilizing the Weierstrass Preparation Theorem which frequently is also used to prove the Noetherian property of these rings. We need the following notation: A power series $f \in \Bbbk[[x_1,\dots,x_n]]$ is called x_n**-general** if $f(0,x_n) \neq 0 \in \Bbbk[x_n]$.

Exercise 4.4.28 (Weierstrass Preparation Theorem). If $f \in \Bbbk[[x_1,\dots,x_n]]$ is a power series, show:

1. By a triangular change of coordinates, we can achieve that f is x_n-general.
2. If f is x_n-general, there exisits a local monomial order on $\Bbbk[x_1,\dots,x_n]$ such that $\mathbf{L}(f) = \mathbf{L}(f(0,x_n))$.
3. If f is x_n-general, then $\langle f \rangle$ is generated by a Weierstrass polynomial

   $$p = x_n^d + a_1(x_1,\dots,x_{n-1})x_n^{n-1} + \ldots + a_d(x_1,\dots,x_{n-1}) \in \Bbbk[[x_1,\dots,x_{n-1}]][x_n] \text{ with } p(0,x_n) = x_n^d,$$

   that is there exists a unit $u \in \Bbbk[[x_1,\dots,x_n]]$ with $f = up$. *Hint:* Grauert division gives an expression $x_n^d = uf + h$ satisfying conditions /DD1) and (DD2). Set $p_n = x_n^d - h$ and show that u is a unit. □

Exercise 4.4.29. Complete the proof of the fact that $\Bbbk[[x_1, \dots, x_n]]$ is factorial. □

Exercise 4.4.30. 1. Formal implicit mappimg theorem
2. Formal inverse function theorem

□

4.5 The Local-Global Principle

The technique of localization often allows one to reduce the proof of a result in commutative algebra to the local case, where the result is easier to establish (for instance, since we can apply Nakayama's lemma). We will see several examples of how this works in the next section. Now, in preparing the ground for some of the arguments, we extend localization from rings to modules, and study **properties** of a module M over a ring R which are **local** in the sense that M has the property iff $M_{\mathfrak{p}}$ has the property for all prime ideals $\mathfrak{p}$ of R. Here, $M_{\mathfrak{p}} = M[U^{-1}]$ is the localization of M at $U = R \setminus \mathfrak{p}$ in the following sense:

Remark-Definition 4.5.1. Let R be a ring, let $U \subset R$ be a multiplicatively closed subset, and let M be an R-module. As in case $M = R$, the relation

$$(m, u) \sim (m', u') \iff v(mu' - um') = 0 \text{ for some } v \in U$$

is an equivalence relation, and we write

$$M[U^{-1}] = U^{-1}M = \{\frac{m}{u} \mid m \in M, u \in U\}$$

for the set of all equivalence classes. We consider $M[U^{-1}]$ as an $R[U^{-1}]$-module, with addition defined as for $R[U^{-1}]$, and with the action

$$\frac{r}{u} \cdot \frac{m}{u'} = \frac{rm}{uu'}.$$

This module is called the **localization of *M* at *U***.

If $\varphi : M \to N$ is an R-module homomorphism, there is an induced homomorphism $\varphi[U^{-1}] : M[U^{-1}] \to N[U^{-1}]$ of $R[U^{-1}]$-modules taking m/u to $\varphi(m)/u$. We have:

1. $\mathrm{id}_M[U^{-1}]) = \mathrm{id}_{M[U^{-1}]}$.
2. If
$$M' \xrightarrow{\varphi} M \xrightarrow{\psi} M''$$
are maps of R-modules, then
$$(\psi \circ \varphi)[U^{-1}] = \psi[U^{-1}] \circ \varphi[U^{-1}].$$

These properties are usually referred to by saying that U^{-1} is a **functor** from the category of R-modules to the category of $R[U^{-1}]$-modules.

Finally, note that if $I \subset R$ is an ideal, then

$$IR[U^{-1}] = I[U^{-1}].$$

Indeed, this is clear since every element $\sum f_i/u_i$ with $f_i \in I$ and $u_i \in U$ for all i can be brought to a common denominator. □

In what follows, let R and U be as above.

Exercise 4.5.2. If M is an R-module, show that

$$M[U^{-1}] \cong M \otimes_R R[U^{-1}].$$ □

Proposition 4.5.3. *The* **functor** U^{-1} *is* **exact**. *That is, if a sequence of R-modules*

$$M' \xrightarrow{\varphi} M \xrightarrow{\psi} M''$$

is exact at M, then the induced sequence of $R[U^{-1}]$-modules

$$M'[U^{-1}] \xrightarrow{\varphi[U^{-1}]} M[U^{-1}] \xrightarrow{\psi[U^{-1}]} M''[U^{-1}]$$

is exact at $M[U^{-1}]$.

Proof. By assumption and since U^{-1} is a functor, $0 = (\psi\circ\varphi)[U^{-1}] = \psi[U^{-1}]\circ \varphi[U^{-1}]$. Hence, $\operatorname{im}\varphi[U^{-1}] \subset \ker\psi[U^{-1}]$. To show the opposite inclusion, let $m/u \in \ker\psi[U^{-1}]$. That is, $0 = \psi[U^{-1}](m/u) = \psi(m)/u$. Then there is an element $v \in U$ such that $0 = v\psi(m) = \psi(vm)$. Hence, $vm \in \ker\psi = \operatorname{im}\varphi$ and, thus, $vm = \varphi(m')$ for some $m' \in M'$. We conclude that

$$m/u = vm/vu = \varphi(m')/vu = \varphi[U^{-1}](m'/vu) \in \operatorname{im}\varphi[U^{-1}].$$ □

The proposition implies, in particular, that if N is a submodule of M, then the induced map $N[U^{-1}] \to M[U^{-1}]$ is injective. We may, thus, regard $N[U^{-1}]$ as a submodule of $M[U^{-1}]$.

Exercise* 4.5.4. Show that localization commutes with forming sums and intersections of submodules.That is, if N and N' are submodules of an R-module M, then:

1. $(N + N')[U^{-1}] = N[U^{-1}] + N'[U^{-1}]$.
2. $(N \cap N')[U^{-1}] = N[U^{-1}] \cap N'[U^{-1}]$. □

Proposition 4.5.5 (Primary Decomposition and Localization). *Let R be a Noetherian ring, let $I \subset R$ be an ideal, let $U \subset R$ be a multiplicatively closed subset, and let $\iota : R \to R[U^{-1}]$ be the natural homomorphism. If $I = \bigcap_{i=1}^t \mathfrak{q}_i$ is a minimal primary decomposition, then*

$$I[U^{-1}] = \bigcap_{\mathfrak{q}_i\cap U=\emptyset} \mathfrak{q}_i[U^{-1}] \;\; \textit{and} \;\; \iota^{-1}(I[U^{-1}]) = \bigcap_{\mathfrak{q}_i\cap U=\emptyset} \mathfrak{q}_i$$

are minimal primary decompositions as well.

Proof. We write $\mathfrak{p}_i = \operatorname{rad} \mathfrak{q}_i$.

If $\mathfrak{q}_i \cap U \neq \emptyset$, then $\mathfrak{q}_i[U^{-1}] = R[U^{-1}]$ since the elements of U are sent to units in $R[U^{-1}]$. In contrast, if $\mathfrak{q}_i \cap U = \emptyset$, then $\mathfrak{q}_i[U^{-1}]$ is $\mathfrak{p}_i[U^{-1}]$-primary and $\iota^{-1}(\mathfrak{q}_i[U^{-1}]) = \mathfrak{q}_i$ (see Exercise 4.2.8). Taking Exercise 4.5.4 into account, we find that

$$I[U^{-1}] = \bigcap_{\mathfrak{q}_i \cap U = \emptyset} \mathfrak{q}_i[U^{-1}]$$

and

$$\iota^{-1}(I[U^{-1}]) = \bigcap_{\mathfrak{q}_i \cap U = \emptyset} \iota^{-1}(\mathfrak{q}_i[U^{-1}]) = \bigcap_{\mathfrak{q}_i \cap U = \emptyset} \mathfrak{q}_i$$

are primary decompositions. These decompositions are minimal since the original decomposition of I is minimal (apply Theorem 4.2.7 to see that the involved prime ideals are distinct). □

Exercise* 4.5.6. Prove the 2nd Uniqueness Theorem 1.8.9 for primary decomposition. □

Now, we give some examples of local properties:

Proposition 4.5.7. *If M is an R-module, the following are equivalent:*

1. $M = 0$.
2. $M_{\mathfrak{p}} = 0$ *for all prime ideals* $\mathfrak{p}$ *of* R.
3. $M_{\mathfrak{m}} = 0$ *for all maximal ideals* $\mathfrak{m}$ *of* R.

Proof. The only nontrivial part of the proof is to show that condition 3 implies condition 1. For this, suppose that $M \neq 0$, and let $m \in M$ be a nonzero element. Then the annihilator $\operatorname{Ann}(m)$ is a proper ideal of R which is necessarily contained in a maximal ideal $\mathfrak{m} \subset R$. It follows that $m/1 \in M_{\mathfrak{m}}$ cannot be zero since otherwise $vm = 0$ for some $v \in R \setminus \mathfrak{m}$, a contradiction to $\operatorname{Ann}(m) \subset \mathfrak{m}$. In particular, $M_{\mathfrak{m}} \neq 0$, as desired. □

In the proposition below, if $\mathfrak{p}$ is a prime ideal of R and $U = R \setminus \mathfrak{p}$, we write $\phi_{\mathfrak{p}} = \phi[U^{-1}]$.

Proposition 4.5.8. *If $\phi : M \to N$ is a homomorphism of R-modules, the following are equivalent:*

1. ϕ *is injective.*
2. $\phi_{\mathfrak{p}} : M_{\mathfrak{p}} \to N_{\mathfrak{p}}$ *is injective for all prime ideals* $\mathfrak{p}$ *of* R.
3. $\phi_{\mathfrak{m}} : M_{\mathfrak{m}} \to N_{\mathfrak{m}}$ *is injective for all maximal ideals* $\mathfrak{m}$ *of* R.

The same holds if we replace "injective" by "surjective" in all statements.

Proof. 1 $\Longrightarrow$ 2: This follows by applying Proposition 4.5.3 to the exact sequence

$$0 \to M \to N.$$

2 $\Longrightarrow$ 3: This is clear.

$3 \implies 1$: Applying Proposition 4.5.3 to the exact sequence

$$0 \to \ker\phi \to M \to N,$$

we find that the localized sequences

$$0 \to (\ker\phi)_{\mathfrak{m}} \to M_{\mathfrak{m}} \to N_{\mathfrak{m}}$$

are exact for all maximal ideals $\mathfrak{m}$ of R. Since all the $(\ker\phi)_{\mathfrak{m}}$ are zero by assumption, also $\ker\phi$ is zero by Proposition 4.5.7.

The surjectivity part follows in the same way. □

Exercise 4.5.9. Show that being normal is a local property of integral domains. □

4.6 Artinian Rings and Krull's Principal Ideal Theorem

In practical applications, we might wish to compute intersection numbers in cases where the intersection points are not rational over the given field of definition of our curves.

Example 4.6.1. In $\mathbb{A}^2(\mathbb{C})$, consider the parabola $C = \mathrm{V}(y^2 - x)$ and the graph $D = \mathrm{V}(x^3 - 6x^2 + 2xy + 9x - 6y + 1)$ of the rational function which sends x to $\frac{x^3-6x^2+9x+1}{6-2x}$.

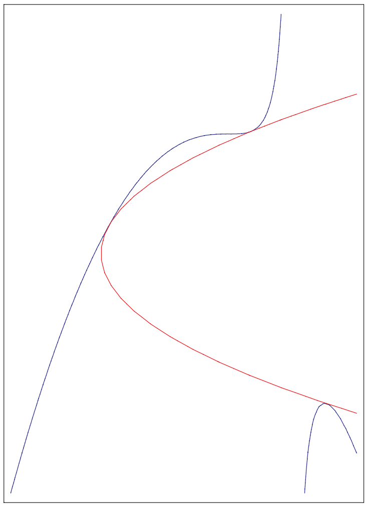


Fig. 4.2. *Three intersection points of multiplicity 2.*

Both curves are defined over $\mathbb{Q}$. Plugging in y^2 for x in the equation defining D, we find that the y-coordinates of the intersection points satisfy the equation $(y^3-3y+1)^2=0$. Hence, we have three intersection points, say $p_i=(a_i,b_i)$, $i=1,2,3$. Since the polynomial y^3-3y+1 is irreducible over $\mathbb{Q}$, the p_i are not defined over $\mathbb{Q}$. They are, in fact, defined over the number field

$$\mathbb{Q}(b_i)\cong\mathbb{Q}[y]/\langle y^3-3y+1\rangle$$

which is an extension field of $\mathbb{Q}$ of degree 3. Intuitively, considering Figure 4.2, each intersection point should be counted with multiplicity 2. Checking this for p_i using Definition 4.3.15, we would have to extend our ground field from $\mathbb{Q}$ to $\mathbb{Q}(b_i)$ and work in $\mathbb{Q}(b_i)[x,y]_{\langle x-a_i,y-b_i\rangle}$.

In what follows, we will describe an alternative way of defining intersection multiplicities which, in the example here, compares the ring

$$R=\mathbb{Q}[x,y]/\langle y^2-x,x^3-6x^2+2xy+9x-6y+1\rangle\cong\mathbb{Q}[y]/\langle y^3-3y+1\rangle^2$$

with its quotient

$$R/\langle\overline{y}^3-3\overline{y}+1\rangle\cong\mathbb{Q}[y]/\langle y^3-3y+1\rangle. \qquad \square$$

In making the alternative definition of intersection multiplicities, we will rely on the concept of length. This provides a measure for the size of a module and constitutes, thus, one way of extending the concept of dimension from vector spaces to modules. Here is the relevant terminology.

Let R be any ring, and let M be any R-module. A **normal series** of M is a sequence

$$M=M_0\supsetneq M_1\supsetneq M_2\supsetneq\ldots\supsetneq M_k=\langle 0\rangle$$

of submodules of M with strict inclusions. The number k of inclusions is called the **length** of the normal series. A **composition series** of M is a maximal normal series, that is, a normal series which cannot be extended to a normal series of greater length by inserting an extra submodule. Equivalently, each **factor** M_i/M_{i+1} is simple. Here, an R-module $0\neq M$ is called **simple** if it has no submodules other than $\langle 0\rangle$ and M itself. Note that simple modules (over *commutative* rings) are fields:

Lemma 4.6.2. *A module $0\neq M$ over a ring R is simple iff M can be written as a quotient $R/\mathfrak{m}$, where $\mathfrak{m}\subset R$ is a maximal ideal.*

Proof. If $M\cong R/\mathfrak{m}$ is a field, then it is clearly simple. For the converse, choose any element $0\neq m\in M$. Then $M=mR$ and, hence, $M\cong R/\mathfrak{m}$, where $\mathfrak{m}=\operatorname{Ann}(m)$. Necessarily, $\mathfrak{m}$ is a maximal ideal since otherwise M would contain a proper nonzero submodule. □

Definition 4.6.3. A module M over a ring R is said to be a **module of finite length** if it has a composition series. In this case, the length of the series is called the **length of M**, written $\ell(M)$. If no composition series exists, set $\ell(M)=\infty$. A **ring R** is **of finite length** if it is of finite length as an R-module. □

We show that $\ell(M)$ is well defined:

Theorem 4.6.4 (Jordan-Hölder). *Let M be a module over a ring R. Suppose that M has a composition series. Then any two such series have the same length. Furthermore, any normal series of M can be extended to a composition series.*

Proof. Let $M = M_0 \supsetneq M_1 \supsetneq M_2 \supsetneq \ldots \supsetneq M_\ell = \langle 0 \rangle$ be any composition series of M. Both statements of the theorem follow from the claim that every normal series of M has length $\leq \ell$. Indeed, the first statement is obtained by applying the claim to a composition series of minimum length. For the second statement, given a normal series of M which is not maximal, note that the process of inserting extra submodules must stop as soon as we reach length l.

To establish the claim, observe that the cases $\ell = 0$ (that is, $M = \langle 0 \rangle$) and $\ell = 1$ (that is, M is simple) are trivial. We consider, therefore, the case $\ell \geq 2$, and suppose inductively that the claim holds for all R-modules with a composition series of length $\leq \ell - 1$.

Let $M = N_0 \supsetneq N_1 \supsetneq N_2 \supsetneq \ldots \supsetneq N_k = \langle 0 \rangle$ be any normal series of M. If $N_1 \subset M_1$, the induction hypothesis applied to M_1 yields $k-1 \leq \ell-1$ since M_1 has a composition series of length $\ell-1$. If $N_1 \not\subset M_1$, we must have $N_1 + M_1 = M$ since M/M_1 is simple. Then $N_1/(M_1 \cap N_1) \cong (N_1 + M_1)/M_1 \cong M/M_1$ is simple as well. On the other hand, applying, once more, the induction hypothesis to M_1, we find that all normal series of the proper submodule $M_1 \cap N_1$ of M_1 must have length $\leq \ell - 2$. It follows that N_1 has a composition series of length $\leq \ell - 2 + 1 = \ell - 1$ since $N_1/(M_1 \cap N_1)$ is simple. As above, we conclude that $k - 1 \leq \ell - 1$. □

Exercise* 4.6.5. Let R be a ring, let M be an R-module of finite length, and let $M = M_0 \supsetneq M_1 \supsetneq M_2 \supsetneq \ldots \supsetneq M_\ell = \langle 0 \rangle$ be a composition series of M. If $\mathfrak{m}$ is a maximal ideal of R, show that the length of the $R_\mathfrak{m}$-module $M_\mathfrak{m}$ is the number of quotients M_i/M_{i+1} isomorphic to $R/\mathfrak{m}$. □

Our next goal is to characterize modules of finite length in terms of chain conditions. For this, we not only consider the ascending chain condition, but also the descending chain condition:

Definition 4.6.6. A **module** M over a ring R is called **Artinian** if it satifies the **descending chain condition**. That is, every chain

$$M = M_0 \supset M_1 \supset M_2 \supset \ldots M_k \supset \ldots$$

of submodules of M is eventually stationary. A **ring** R is called **Artinian** if it is Artinian as an R-module. That is, R satisfies the descending chain condition on ideals. □

As in Exercise 1.4.5 one shows that M is Artinian iff the **minimal condition** on submodules holds: Every nonempty set of ideals of R has a minimal element with respect to inclusion.

Proposition 4.6.7. *Let M be a module over a ring R. Then the following are equivalent:*

1. *M is of finite length.*
2. *M is Artinian and Noetherian.*

Proof. $1 \implies 2$: If $\ell(M) < \infty$, the length of any normal series of M is bounded by $\ell(M)$. Hence, both chain conditions hold.

$2 \implies 1$: Since M is Noetherian, it satisfies the maximal condition. In particular, there is a maximal submodule $M_1 \subsetneq M = M_0$ which, necessarily, is Noetherian as well. Applying the same argument to M_1 and so forth, we get a descending chain $M = M_0 \supsetneq M_1 \supsetneq \dots$ which, since M is Artinian, is eventually stationary. It is, hence, a composition series of M. □

Exercise* 4.6.8. Let R be a ring, and let

$$0 \to M' \to M \to M'' \to 0$$

be a short exact sequence of R-modules. Show:

1. M is Artinian (respectively Noetherian) iff both M' and M'' are Artinian (respectively Noetherian).
2. M is of finite length iff both M' and M'' are of finite length. In this case,

$$\ell(M) = \ell(M') + \ell(M'').$$

□

The examples in the following exercise illustrate our definitions:

Exercise* 4.6.9. Show:

1. If M is a module over a field K, that is, M is a K-vector space, then M is Noetherian iff M is Artinian iff M is of finite length iff $\dim_K M < \infty$.
2. If I is an ideal of a ring R, then R/I is of finite length as a ring iff it is of finite length as an R-module.
3. An affine $\Bbbk$-algebra $\Bbbk[x_1, \dots, x_n]/I$ is of finite length iff it has finite dimension as a $\Bbbk$-vector space. Geometrically, this is the case where the vanishing locus $\mathrm{V}(I) \subset \mathbb{A}^n$ consists of finitely many points.
4. The $\Bbbk[x]$-module $M = \Bbbk[x, x^{-1}]/\Bbbk[x]$ is Artinian, but not Noetherian. □

Definition 4.6.10. Let $f, g \in \Bbbk[x, y]$ be nonconstant polynomials, and let $\mathfrak{m}$ be a maximal ideal of $\Bbbk[x, y]$. The **intersection multiplicity of f and g at $\mathfrak{m}$**, written $i(f, g; \mathfrak{m})$, is defined to be

$$i(f, g; \mathfrak{m}) = \ell(\Bbbk[x, y]_{\mathfrak{m}} / \langle f, g\rangle \Bbbk[x, y]_{\mathfrak{m}}).$$

□

As a consequence of the definition, the following facts are easy to prove:

Exercise 4.6.11 (Properties of Intersection Multiplicities). Let $f, g \in \Bbbk[x, y]$ be nonconstant polynomials, and let $\mathfrak{m} \subset \Bbbk[x, y]$ be a maximal ideal. Then show:

1. $i(f,g;\mathfrak{m}) = 0$ iff $\mathrm{V}(\mathfrak{m}) \not\subset \mathrm{V}(f) \cap \mathrm{V}(g) \subset \mathbb{A}^2$.
2. $i(f,g;\mathfrak{m}) = \infty$ iff f and g have a common factor contained in $\mathfrak{m}$.
3. If $\mathrm{V}(f) \cap \mathrm{V}(g) \subset \mathbb{A}^2$ is finite, then $i(f,g;\mathfrak{m})$ is the number of quotients in a composition series of $\Bbbk[x,y]/\langle f,g\rangle$ which are isomorphic to $\Bbbk[x,y]/\mathfrak{m}$.
4. If the field extension $\Bbbk[x,y]/\mathfrak{m} \supset \Bbbk$ is separable, then $\mathrm{V}(\mathfrak{m}) \subset \mathbb{A}^2$ consists of $[\Bbbk[x,y]/\mathfrak{m} : \Bbbk]$ points (which form an orbit under the natural action of the Galois group of $\overline{\Bbbk}$ over $\Bbbk$). For each such point p,
$$i(f,g;p) = i(f,g;\mathfrak{m}).$$
5. If $\Bbbk[x,y]/\mathfrak{m} \supset \Bbbk$ is inseparable, then $\mathrm{V}(\mathfrak{m})$ consists of $[\Bbbk[x,y]/\mathfrak{m} : \Bbbk]_{sep}$ points. For each such point p,
$$i(f,g;p) = i(f,g;\mathfrak{m}) \cdot [\Bbbk[x,y]/\mathfrak{m} : \Bbbk]_{insep}.$$
Here, the subscripts *sep* and *insep* refer to the separable and inseparable degrees, respectively. □

Example 4.6.12. The affine $\mathbb{Q}$-algebra

$$R = \mathbb{Q}[x,y]/\langle y^2 - x, x^3 - 6x^2 + 2xy + 9x - 6y + 1\rangle \cong \mathbb{Q}[y]/\langle y^3 - 3y + 1\rangle^2$$

from Example 4.6.1 has finite length since it has finite dimension as a $\mathbb{Q}$-vector space. In fact, $R \supsetneq \langle \overline{y}^3 - 3\overline{y} + 1\rangle \supsetneq \langle 0\rangle$ is a composition series. Note that both factors are isomorphic to $\mathbb{Q}[y]/\langle y^3 - 3y + 1\rangle$. Taking parts 4, 3 of Exercise 4.6.11 into account, we find, as expected, that the curves C, D from Example 4.6.1 have three intersection points, each of which has multiplicity 2. □

Exercise 4.6.13. Let $I \subset \Bbbk[x_1,\dots,x_n]$ be an ideal such that $\mathrm{V}(I) \subset \mathbb{A}^n$ is finite, and let $\mathfrak{m} \subset \Bbbk[x_1,\dots,x_n]$ be a maximal ideal. Express

$$\ell(\Bbbk[x_1,\dots,x_n]_{\mathfrak{m}} / I\,\Bbbk[x_1,\dots,x_n]_{\mathfrak{m}})$$

in terms of the sequence $\dim_{\Bbbk} \Bbbk[x_1,\dots,x_n]/I_k$, $k \geq 0$, where I_k is defined inductively by $I_0 = I$ and $I_k = I_{k-1} : \mathfrak{m}$. □

Exercise 4.6.14. Some examples for intersection number computations. □

Despite the formal symmetry between the ascending and the decending chain condition, the notions of Noetherian and Artinian rings are quite different. In fact, our next result shows that every Artinian ring is Noetherian, but of a very special kind (so that most Notherian rings are not Artinian):

Theorem 4.6.15. *For a ring R, the following are equivalent:*

1. *R is Noetherian and* $\dim R = 0$.
2. *R has finite length.*
3. *R is Artinian.*

If these conditions are satisfied, then R has only finitely many maximal ideals.

Proof. 1 $\implies$ 2: Suppose that R is Noetherian. If R is not of finite length, the set

$$\mathit{\Gamma} := \{I \subset R \text{ ideal } \mid R/I \text{ is not of finite length}\}$$

is nonempty since $\langle 0\rangle \in \mathit{\Gamma}$. Hence, since R is Noetherian, $\mathit{\Gamma}$ contains a maximal element $\mathfrak{p}$. We show that $\mathfrak{p}$ is a prime ideal. For this, let $f, g \in R$ be elements such that $fg \in \mathfrak{p}$, but $f \notin \mathfrak{p}$. Consider the exact sequence

$$0 \to R/(\mathfrak{p} : f) \xrightarrow{\cdot f} R/\mathfrak{p} \to R/(\mathfrak{p} + \langle f\rangle) \to 0.$$

Since $\mathfrak{p} + \langle f\rangle \supsetneqq \mathfrak{p}$, the module $R/(\mathfrak{p} + \langle f\rangle)$ must have finite length by the maximality of $\mathfrak{p}$ as an element of $\mathit{\Gamma}$. If g would not be an element of $\mathfrak{p}$, then $\mathfrak{p} : f$ would contain $\mathfrak{p}$ properly, and $R/(\mathfrak{p} : f)$ would have finite length as well. But, then, $R/\mathfrak{p}$ would have finite length by Exercise 4.6.8, a contradiction to our choice of $\mathfrak{p}$.

Now, suppose not only that R is Notherian, but also that $\dim R = 0$. Then all prime ideals of R are maximal. In particular, if R were not of finite length, the prime ideal $\mathfrak{p}$ just constructed would be a maximal ideal, so that $R/\mathfrak{p}$ would be a field. This contradicts, again, the fact that $R/\mathfrak{p}$ is not of finite length.

2 $\implies$ 3: This is clear.

3 $\implies$ 1: Now, suppose that R is Artinian. To show that R satifies condition 1, we proceed in four steps.

Step 1. We show that $\dim R = 0$. For this, consider a nested pair of prime ideals $\mathfrak{p}_1 \subset \mathfrak{p}_2 \subset R$, and let f be any element of $\mathfrak{p}_2/\mathfrak{p}_1 \subset R/\mathfrak{p}_1$. Since $R/\mathfrak{p}_1$ is Artinian as well, the descending chain condition yields a number m such that $\langle f^m\rangle = \langle f^{m+1}\rangle$. Then $f^m = gf^{m+1}$ for some $g \in R/\mathfrak{p}_1$. That is, $(1 - gf)f^m = 0$. Since $R/\mathfrak{p}_1$ is an integral domain and $f \in \mathfrak{p}_2/\mathfrak{p}_1 \subsetneqq R/\mathfrak{p}_1$ is not a unit, we conclude that $f = 0$. It follows that $\mathfrak{p}_1 = \mathfrak{p}_2$ and, thus, that $\dim R = 0$, as claimed.

Step 2. The ring R has only finitely many maximal ideals since any infinite sequence $\mathfrak{m}_1, \mathfrak{m}_2, \mathfrak{m}_3, \ldots$ of maximal ideals of R would yield an infinite descending chain of ideals

$$\mathfrak{m}_1 \supset \mathfrak{m}_1 \cap \mathfrak{m}_2 \supset \ldots \supset \mathfrak{m}_1 \cap \mathfrak{m}_2 \cap \ldots \cap \mathfrak{m}_k \supset \ldots$$

with strict inclusions (by part 2 of Exercise 1.3.4). Writing $\mathfrak{m}_1, \ldots, \mathfrak{m}_s$ for the distinct maximal ideals of R and taking into account that every prime ideal of R is maximal by step 1, we conclude from Exercise 3.2.11 that

$$\operatorname{rad}\langle 0\rangle = \mathfrak{m}_1 \cap \ldots \cap \mathfrak{m}_s. \tag{4.12}$$

Step 3. For any i, the descending chain of ideals $\mathfrak{m}_i \supset \mathfrak{m}_i^2 \supset \mathfrak{m}_i^3 \supset \ldots$ is eventually stationary. We may, hence, choose a number N such thst $\mathfrak{m}_i^N = \mathfrak{m}_i^{N+1}$ for all i. Consider the ideal

$$I = \prod_{i=1}^{s} \mathfrak{m}_i^N.$$

Then $I^2 = I$. We use this to show that $I = \langle 0 \rangle$. Suppose the contrary. Then the set

$$\Gamma := \{J \subsetneq R \mid JI \neq \langle 0 \rangle\}$$

contains I since $I^2 = I \neq \langle 0 \rangle$. Hence, since R is Artinian, Γ contains a minimal element J_0. Let f be an element of J_0 such that $fI \neq \langle 0 \rangle$. Then $\langle f \rangle = J_0$ by the minimality of J_0. The same argument gives $fI = J_0 = \langle f \rangle$ since $(fI)I = fI^2 = fI \neq 0$. Choose an element $g \in I$ such that $fg = f$. Then $f = fg = fg^2 = \ldots = fg^m = 0$ for some $m \geq 1$ since every element of I is nilpotent by (4.12). This contradiction proves that $I = \langle 0 \rangle$, as claimed.

Step 4. Each of the successive quotients in the descending chain of ideals

$$R \supset \mathfrak{m}_1 \supset \ldots \supset \mathfrak{m}_1^N \supset \mathfrak{m}_1^N \mathfrak{m}_2 \supset \ldots \supset \prod_{i=1}^{s} \mathfrak{m}_i^N = \langle 0 \rangle \tag{4.13}$$

is a vector space over some field $R/\mathfrak{m}_i$. Hence, taking part 1 of Exercise 4.6.8 and part 1 of Exercise 4.6.9 into account, we get the following chain of eqivalences: R is Artinian $\Longleftrightarrow$ each quotient in (4.13) is Artinian $\Longleftrightarrow$ each quotient in (4.13) is Noetherian $\Longleftrightarrow$ R is Noetherian. This concludes the proof. □

Next, we establish a structure result for Artinian rings. Then, following Krull, we will apply Theorem 4.6.15 above to prove the principal ideal theorem which is fundamental to the dimension theory of Noetherian rings.

Theorem 4.6.16 (Structure Theorem for Artinian Rings). *Let R be an Artinian ring, and let $\mathfrak{m}_1, \ldots, \mathfrak{m}_s$ be the distinct maximal ideals of R. Then*

$$R \cong \prod_{i=1}^{s} R_{\mathfrak{m}_i}$$

is a finite direct product of local Artinian rings.

Proof. To begin with, we conclude from Theorem 4.2.7 that any localization of an Artinian ring is again Artinian. Now, as in the preceeding proof, choose a number N such that $\prod_{i=1}^{s} \mathfrak{m}_i^N = \langle 0 \rangle$. Since the $\mathfrak{m}_i$ are pairwise coprime, the $\mathfrak{m}_i^N$ are pairwise coprime as well (see part 4 of Exercise 1.5.12). Hence, the natural map

$$R \to \prod_{i=1}^{s} R/\mathfrak{m}_i^N \tag{4.14}$$

is an isomorphism by the Chinese remainder theorem (see Exercise 1.3.9). To conclude the proof, we localize both sides of (4.14) and find that $R_{\mathfrak{m}_i} \cong (R/\mathfrak{m}_i^N)_{\mathfrak{m}_i} \cong R/\mathfrak{m}_i^N$ (indeed, $(R/\mathfrak{m}_j^N)_{\mathfrak{m}_i} = 0$ for $j \neq i$ and $R/\mathfrak{m}_i^N$ is a local ring). □

In the geometric context, the structure theorem extends Remark 4.3.16:

Corollary 4.6.17. *Let $I \subset \mathbb{k}[x_1, \dots, x_n]$ be an ideal such that $\mathrm{V}(I) \subset \mathbb{A}^n$ is finite, say $\mathrm{V}(I) = \{p_1, \dots, p_s\}$. Then there is a natural isomorphism of $\mathbb{K}$-algebras*

$$\mathbb{K}[x_1, \dots, x_n]/I\,\mathbb{K}[x_1, \dots, x_n] \cong \prod_{i=1}^{s} \mathcal{O}_{\mathbb{A}^n,p_i}/I\mathcal{O}_{\mathbb{A}^n,p_i}.$$

□

Theorem 4.6.18 (Krull's Principal Ideal Theorem, First Version). *Let R be a Noetherian ring, and let $f \in R$. Then every minimal prime $\mathfrak{p}$ of $\langle f \rangle$ satisfies*

$$\operatorname{codim} \mathfrak{p} \leq 1.$$

If f is not a zerodivisor of R, then equality holds.

Proof. To show the first statement of the theorem, we will localize and apply Nakayama's lemma. To begin with, recall from Proposition 4.2.13 that if $\mathfrak{p}$ is any prime ideal of any ring R, then $\operatorname{codim} \mathfrak{p} = \dim R_{\mathfrak{p}}$. With our assumptions here, we have, in addition, that $\mathfrak{p}R_{\mathfrak{p}}$ is a minimal prime of $\langle f \rangle R_{\mathfrak{p}}$. Replacing R by $R_{\mathfrak{p}}$, we may, hence, assume that R is local ring with maximal ideal $\mathfrak{p}$. The first statement of the theorem will follow once we show that $\operatorname{codim} \mathfrak{q} = \dim R_{\mathfrak{q}} = 0$ for every prime ideal $\mathfrak{q} \subsetneq \mathfrak{p}$.

For this, given $\mathfrak{q}$, consider the ideals

$$\mathfrak{q}^{(n)} = \{a \in R \mid ua \in \mathfrak{q}^n \text{ for some } u \notin \mathfrak{q}\},\ n \geq 1.$$

Then, by part 1 of Proposition 4.2.7, $\mathfrak{q}^{(n)}$ is the preimage of $\mathfrak{q}^n R_{\mathfrak{q}}$ under the localization map $R \to R_{\mathfrak{q}}$. Since the maximal ideal $\mathfrak{p} + \langle f \rangle$ of the quotient ring $R/\langle f \rangle$ is also minimal, this ring is zerodimensionial. Being also Noetherian, it is Artinian by Theorem 4.6.15. Hence, the descending chain

$$\mathfrak{q}^{(1)} + \langle f \rangle \supset \mathfrak{q}^{(2)} + \langle f \rangle \supset \dots$$

is eventually stationary, say $\mathfrak{q}^{(n)} + \langle f \rangle = \mathfrak{q}^{(n+1)} + \langle f \rangle$. As a consequence, any element $g \in \mathfrak{q}^{(n)}$ can be written as a sum $g = h + af$ with $h \in \mathfrak{q}^{(n+1)}$ and $a \in R$. Then $af \in \mathfrak{q}^{(n)}$. Since $\mathfrak{p}$ is a minimal prime of $\langle f \rangle$, we have $f \notin \mathfrak{q}$ and, thus, $a \in \mathfrak{q}^{(n)}$ by the very definition of $q^{(n)}$. This shows that

$$\mathfrak{q}^{(n)} = f\mathfrak{q}^{(n)} + \mathfrak{q}^{(n+1)}.$$

Since f is contained in the maximal ideal $\mathfrak{p}$ of R, Nakayama's lemma yields $\mathfrak{q}^{(n)} = \mathfrak{q}^{(n+1)}$. Then $\mathfrak{q}^n R_{\mathfrak{q}} = \mathfrak{q}^{n+1} R_{\mathfrak{q}}$ by part 2 of Proposition 4.2.7. Applying Nakayama's lemma in $R_{\mathfrak{q}}$, we, hence, get $\mathfrak{q}^n R_{\mathfrak{q}} = \langle 0 \rangle$. We conclude that $\dim R_{\mathfrak{q}} = 0$, as desired.

The second statement of the theorem follows from the first one. Indeed, the Noetherian ring R contains only finitely many minimal prime ideals, say $\mathfrak{p}_1, \dots, \mathfrak{p}_r$. Thus, if f is a not a zerodivisor of R, it is not contained in any of the $\mathfrak{p}_i$ by Exercise 3.2.12. This implies that $\operatorname{codim} \mathfrak{p} \geq 1$. □

Theorem 4.6.19 (Krull's Principal Ideal Theorem, General Version). *Let R be a Noetherian ring. If $I = \langle f_1, \dots, f_c\rangle \subset R$ is an ideal which is generated by c elements, then every minimal prime $\mathfrak{p}$ of I satisfies*

$$\operatorname{codim} \mathfrak{p} \leq c.$$

Conversely, if $\mathfrak{p} \subset R$ is a prime ideal such that $\operatorname{codim} \mathfrak{p} = c$*, there exist elements $y_1, \dots, y_c \in R$ such that $\mathfrak{p}$ is a minimal prime of $\langle y_1, \dots, y_c\rangle$.*

Proof. To show the first statement of the theorem, let $\mathfrak{p}$ be a minimal prime of I. As in the preceeding proof, we may assume that R is a local ring with maximal ideal $\mathfrak{p}$. We do induction on c.

If $c = 0$, there is nothing to show. If $c > 0$, since R is Noetherian, we may find a prime ideal $\mathfrak{q} \subsetneq \mathfrak{p}$ such that no other prime ideal is between $\mathfrak{q}$ and $\mathfrak{p}$. Since $\mathfrak{p}$ is a minimal prime of $I = \langle f_1, \dots, f_c\rangle$, at least one of the f_i is not contained in $\mathfrak{q}$, say $f_c \notin \mathfrak{q}$. Then the maximal ideal $\mathfrak{p} + (\mathfrak{q} + \langle f_c\rangle)$ of the quotient ring $R/(\mathfrak{q} + \langle f_c\rangle)$ is also minimal, so that this ring is an Artinian local ring. In particular, all the f_i are nilpotent mod $\mathfrak{q} + \langle f_c\rangle$. Say,

$$f_i^N = g_i + a_i f_c \text{ with } g_i \in \mathfrak{q} \text{ snd } a_i \in R,\ i = 1, \dots, c-1.$$

Then $\mathfrak{p} \supset \langle g_1, \dots, g_{c-1}, f_c\rangle$, and the image $\overline{\mathfrak{p}}$ of $\mathfrak{p}$ in $R/\langle g_1, \dots, g_{c-1}\rangle$ is a minimal prime of the principal ideal $\langle \overline{f}_c\rangle$. Hence, $\overline{\mathfrak{p}}$ has codimension at most 1 by the first version of the principal ideal theorem. In R, this shows that $\mathfrak{q}$ is a minimal prime of $\langle g_1, \dots, g_{c-1}\rangle$. The induction hypothesis gives codim $\mathfrak{q} \leq c-1$ and, thus, codim $\mathfrak{p} \leq c$.

For the converse statement, given $\mathfrak{p}$ as in the statement, we choose the y_i one at a time. Inductively, with $0 \leq k < c$, suppose that $y_1, \dots, y_k \in \mathfrak{p}$ have already been chosen to generate an ideal of codimension k. Then, by prime avoidance, it is possible to pick an element $y_{k+1} \in \mathfrak{p}$ not contained in any of the finitely many minimal primes of $\langle y_1, \dots, y_k\rangle$ (indeed, any such prime does not contain $\mathfrak{p}$ since its codimension is $\leq k < c$ by the first statement of the theorem). Clearly, $\operatorname{codim}\langle y_1, \dots, y_k, y_{k+1}\rangle = k + 1$, and the result follows. □

We are, now, ready to prove inequality (4.1) in its general form (4.3):

Corollary 4.6.20. *Let $(R, \mathfrak{m})$ be a local Noetherian ring. Then*

$$\dim R = \min\{d \mid \text{ there exists an } \mathfrak{m}\text{-primary ideal } \langle y_1, \dots, y_d\rangle\}. \qquad (4.15)$$

In particular,

$$\dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2 \geq \dim R.$$

Proof. The last statement follows from the first one since $\mathfrak{m}$ is generated by $\dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2$ elements (see Corollary 4.2.20 to Nakayama's lemma).

For the first statement, let $d = \dim R = \operatorname{codim} \mathfrak{m}$, and let d' be the minimum on the right hand side of (4.15). Then $d \leq d'$ respectively $d' \leq d$ follow from the first respectively second statement of the generalized principal ideal theorem. □

Its applications to geometry make Corollary 4.6.20 an important result of commutative algebra, where, in the situation of the corollary, a sequence of $d = \dim R$ elements $y_1, \dots, y_d \in \mathfrak{m}$ is called a **system of parameters** for R if it generates an $\mathfrak{m}$-primary ideal. If $(R, \mathfrak{m})$ is regular, that is, if $\dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2 = d$, then, by Corollary 4.2.20, every minimal set of generators for $\mathfrak{m}$ is a system of parameters consisting of d elements. Such a system is called a **regular system of parameters** for R. A typical example is given below:

Corollary 4.6.21. *The formal power series ring $\Bbbk[[x_1, \dots, x_n]]$ is regular of dimension n. In fact, $x_1, \dots, x_n$ form a regular system of parameters.*

Proof. Since $\Bbbk[[x_1, \dots, x_n]]$ is an integral domain, $\dim \Bbbk[[x_1, \dots, x_n]]/\langle x_n \rangle = \dim \Bbbk[[x_1, \dots, x_n]] - 1$ by Krull's principal ideal theorem. On the other hand, $\Bbbk[[x_1, \dots, x_n]]/\langle x_n \rangle \cong \Bbbk[[x_1, \dots, x_{n-1}]]$. Hence, we conclude by induction on n that $\dim \Bbbk[[x_1, \dots, x_n]] = n$. The result follows. □

Remark 4.6.22. We mention in passing that every regular local ring $(R, \mathfrak{m})$ is an integral domain (to prove this, induct on $\dim R$ and use Nakayama's lemma). This, in turn, implies that if $y_1, \dots, y_d$ is a regular system of parameters for R, then $y_1, \dots, y_d$ is a **regular sequence** on R. That is, each y_i represents a nonzerodivisor of $R/\langle y_1, \dots, y_{i-1} \rangle$, $i = 1, \dots, d$. See Eisenbud (1995), Corollaries 10.14, 10.15 for details and further reading. □

At this point, the general definition of a Cohen-Macaulay ring deserves mentioning (though we will not need it in this book). According to this definition and the remark above, every regular local ring is Cohen-Macaulay.

Definition 4.6.23. A local Noetherian ring $(R, \mathfrak{m})$ is called **Cohen-Macaulay** if it has a system of parameters which is at the same time a regular sequence for R. An arbitrary Noetherian ring is called **Cohen-Macaulay** iff its lcoalization $R_{\mathfrak{p}}$ is Cohen-Macaulay for every prime ideal $\mathfrak{p}$ of R. □

The first statement made in Remark 4.6.22 says, in particular, that the local ring of an algebraic set at a smooth point is an integral domain. In the next two propositions, we give a direct proof for this fact:

Proposition 4.6.24. *Let $p = (a_1, \dots, a_n) \in \mathbb{A}^n$ be a point, let $f_1, \dots, f_r \in \Bbbk[x_1, \dots, x_n]$ be polynomials vanishing at p, where $1 \le r \le n$, and let $R := \mathcal{O}_{\mathbb{A}^n,p}/\langle f_1, \dots, f_r \rangle \mathcal{O}_{\mathbb{A}^n,p}$. Suppose the matrix $M = \left(\frac{\partial f_i}{\partial x_j}(p)\right)_{1 \le i,j \le r}$ has maximal rank r. Then R is isomorphic to a subring of $\mathbb{K}[[x_{r+1} - a_{r+1}, \dots, x_n - a_n]]$. In particular, R is an integral domain.*

Proof. By translating p to the origin o, we may assume that $p = o$. We write $M^{-1} = (a_{ki})$ and set $g_k = \sum_{i=1}^{r} a_{ki} f_i$, $k = 1, \dots r$. Then each g_k is of type x_k + terms of degree ≥ 2. In particular, by Buchberger's criterion, the g_k form a Gröbner basis for the ideal generated by the f_i in $\mathbb{K}[[x_1, \dots, x_n]]$ (fix a degree-anticompatible monomial order on $\mathbb{K}[x_1, \dots, x_n]$). Given any

$g \in \mathcal{O}_{\mathbb{A}^n,o} \subset \mathbb{K}[[x_1,\dots,x_n]]$, the uniquely determined remainder h on Grauert division of g by the g_k is contained in $\mathbb{K}[[x_{r+1},\dots,x_n]]$. Sending g to h defines, thus, a map $\mathcal{O}_{\mathbb{A}^n,o} \to \mathbb{K}[[x_{r+1},\dots,x_n]]$ whose kernel is $\langle f_1,\dots,f_r\rangle\mathcal{O}_{\mathbb{A}^n,o}$. The result follows. □

Proposition 4.6.25. *Let $A \subset \mathbb{A}^n$ be an algebraic set, let $p \in A$ be a point, and let $d = \dim_p A$. Suppose we can find polynomials $f_1,\dots,f_{n-d} \in \mathrm{I}(A)$ such that the matrix $M = \left(\frac{\partial f_i}{\partial x_j}(p)\right)_{1\le i,j\le n-d}$ has maximal rank $n-d$. Then $\mathcal{O}_{A,p} \cong \mathcal{O}_{\mathbb{A}^n,p}/\langle f_1,\dots,f_{n-d}\rangle\mathcal{O}_{\mathbb{A}^n,p}$, and this ring is a regular local ring.*

Proof. Of course, up to renumbering the variables, the assumption just means that p is a smooth point of A. To establish the result, we consider the natural epimorphism of local rings

$$\phi : R := \mathcal{O}_{\mathbb{A}^n,p}/\langle f_1,\dots,f_{n-d}\rangle\mathcal{O}_{\mathbb{A}^n,p} \to \mathcal{O}_{A,p}.$$

Corollary 4.6.20 gives us $d = \dim \mathcal{O}_{A,p} \le \dim R \le \dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2 = d$, where $\mathfrak{m}$ is the maximal of R (for the latter equality, note that $d_p f_1,\dots,d_p f_{n-d}$ are $\mathbb{K}$-linearly independent by virtue of the assumption on the matrix M). Since R is an integral domain by the preceeding proposition, we conclude that $\ker\phi$ is zero (which completes the proof): if $f \in \ker\phi$ were a nonzero element, Krull's principal ideal theorem would give us $d = \dim \mathcal{O}_{A,p} \le \dim R/\langle f\rangle \le \dim R - \operatorname{codim}\langle f\rangle = d-1$. □

We can, now, prove part 2 of Remark 4.1.11:

Corollary 4.6.26. *Let A be an algebraic set. If $A = V_1 \cup \dots \cup V_s$ is the decomposition of A into its irreducible components, then*

$$A_{\mathrm{sing}} = \bigcup_{i\neq j}(V_i \cap V_j) \cup \bigcup_i (V_i)_{\mathrm{sing}}.$$

Proof. Let $p \in A$ be a smooth point of A. Then, since $\mathcal{O}_{A,p}$ is an integral domain by the preceeding propositions, p lies on a unique component V_i of A. It is, then, a smooth point of V_i. We conclude that $A \setminus A_{\mathrm{sing}} \subset (\bigcup_i V_i \setminus (V_i)_{\mathrm{sing}}) \setminus \bigcup_{i\neq j}(V_i \cap V_j)$. The converse inclusion is clear. □

Furthermore, we can show the corollaries to the Jacobian criterion. For this, let $I = \langle f_1,\dots,f_r\rangle \subset \Bbbk[x_1,\dots,x_n]$ be an ideal, let $A = \mathrm{V}(I) \subset \mathbb{A}^n$, and let $I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right)$ be the ideal generated by the $(n-d)\times(n-d)$ minors of the Jacobian matrix of the f_i. Moreover, let $I^e = I\,\mathbb{K}[x_1,\dots,x_n]$.

Proof of Corollary 4.1.13, conclusion. Supposing that A is equidimensional of dimension d, we have to show: If

$$I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right) + I = \langle 1\rangle,$$

then $I^e = \mathrm{I}(A)$.

Let $\mathfrak{m} \subset \mathbb{K}[x_1, \ldots, x_n]$ be any maximal ideal, and let $p \in \mathbb{A}^n$ be the corresponding point. Since $I^e \subset \mathrm{I}(A) \subset \mathbb{K}[x_1, \ldots, x_n]$, also $I^e_{\mathfrak{m}} \subset \mathrm{I}(A)_{\mathfrak{m}} \subset \mathcal{O}_{\mathbb{A}^n,p}$ by the injectivity part of Proposition 4.5.8, and our claim will follow from the surjectivity part of that proposition once we show that $I^e_{\mathfrak{m}} = \mathrm{I}(A)_{\mathfrak{m}}$. For this, we distinguish two cases.

If $p \in \mathbb{A}^n \setminus A$, there is a polynomial $f \in I^e \subset \mathrm{I}(A)$ which is not contained in $\mathfrak{m}$. Then f is a unit in $\mathcal{O}_{\mathbb{A}^n,p}$, which implies that $I^e_{\mathfrak{m}} = \mathrm{I}(A)_{\mathfrak{m}} = \mathcal{O}_{\mathbb{A}^n,p}$.

If $p \in A$, then $I^e \subset \mathrm{I}(A) \subset \mathfrak{m}$. By assumption, at least one $(n-d) \times (n-d)$ minor of $\left(\frac{\partial f_i}{\partial x_j}(p)\right)$ is nonzero, say $\det \left(\frac{\partial f_i}{\partial x_j}(p)\right)_{1 \leq i,j \leq n-d} \neq 0$. Then $(\langle f_1, \ldots, f_{n-d}\rangle \mathbb{K}[x_1, \ldots, x_n])_{\mathfrak{m}} = \mathrm{I}(A)_{\mathfrak{m}}$ by Proposition 4.6.25 and, thus, also $I^e_{\mathfrak{m}} = \mathrm{I}(A)_{\mathfrak{m}}$. □

Proof of Corollary 4.1.14. Supposing that $\Bbbk[x_1, \ldots, x_n]/I$ is Cohen-Macaulay of dimension d, we have to show: If

$$\dim \mathrm{V}(I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right) + I) < \dim \mathrm{V}(I) = d,$$

then $I^e = \mathrm{I}(A)$ and $\mathrm{V}(I_{n-d}\left(\frac{\partial f_i}{\partial x_j}\right) + I) = A_{\text{sing}}$.

Arguing as in the previous proof, we see that the equality $I^e_{\mathfrak{m}} = \mathrm{I}(A)_{\mathfrak{m}}$ holds for the maximal ideal $\mathfrak{m}$ of any point $p \in A$ which is not contained in $B := \mathrm{V}(I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I)$. On the other hand, by virtue of the Cohen-Macaulay assumption, we conclude from the Unmixedness Theorem 3.3.12 that I^e has only isolated primary components, all of dimension d. In particular, by the 2nd uniqueness theorem for primary decomposition, I^e admits a uniquely determined minimal primary decomposition, say, $I^e = \bigcap_{i=1}^t \mathfrak{q}_i$. The radicals $\mathfrak{p}_i = \operatorname{rad} \mathfrak{q}_i$ are the associated primes of $\mathrm{I}(A)$, and the vanishing loci $V_i = \mathrm{V}(\mathfrak{q}_i)$ are the irreducible components of A.

For each i, since $\dim V_i = d > \dim B$, there is a point $p_i \in V_i \setminus (B \cup \bigcup_{j \neq i} V_j)$. Localize $R = \mathbb{K}[x_1, \ldots, x_n]$ at the maximal ideal $\mathfrak{m}_i$ of p_i, and let $\iota : R \to R_{\mathfrak{m}_i}$ be the natural homomorphism. Then, by Proposition 4.5.5, we have $\mathfrak{q}_i = \iota^{-1}(I^e_{\mathfrak{m}_i}) = \iota^{-1}(\mathrm{I}(A)_{\mathfrak{m}_i}) = \mathfrak{p}_i$. This shows that $I^e = \mathrm{I}(A)$.

Replacing I by $\mathrm{I}(A)$ in the definition of B, we see that $\dim T_pA > d$ iff $p \in B$. Hence, $B = A_{\text{sing}}$ since A is equidimensional of dimension d. □

4.7 Analytic Type and Tangent Cone

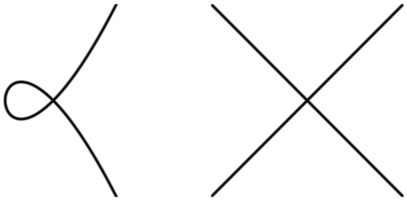
So far, we have defined two invariants of an algebraic set A at a point $p \in A$, namely the local ring $\mathcal{O}_{A,p}$ with its maximal ideal $\mathfrak{m}_{A,p}$, and the Zariski tangent space $T_pA \cong (\mathfrak{m}_{A,p}/\mathfrak{m}^2_{A,p})^*$. In this section, motivated by the fact that both invariants have their drawbacks at singular points, we will introduce two further invariants of A at p.

To begin, we note that the concept of the local ring is too fine at singular points in that two rings $\mathcal{O}_{A,p}$ and $\mathcal{O}_{B,q}$ may differ although our intuition is that locally, near p respectively q, the algebraic sets A and B look alike.

Example 4.7.1. For the plane curves

$$C = \mathrm{V}(y^2 - x^2 - x^3) \subset \mathbb{A}^2(\mathbb{C}) \quad \text{and} \quad D = \mathrm{V}(v^2 - u^2) \subset \mathbb{A}^2(\mathbb{C}),$$

our intuitive understanding is that C and D look alike near the origin o:



$y^2 - x^2 - x^3 = 0$ $\qquad$ $v^2 - u^2 = 0$

Nevertheless, the local rings $\mathcal{O}_{C,o}$ and $\mathcal{O}_{D,o}$ are not isomorphic. In fact, since C is irreducible, $\mathcal{O}_{C,o}$ is a subring of the rational function field $\Bbbk(C)$ and, thus, an integral domain. In contrast, reflecting the fact that o is contained in two irreducible components of D, the ring $\mathcal{O}_{D,o}$ contains zerodivisors: $(v-u)(v+u) = 0 \mod \langle v^2 - u^2 \rangle$. □

From a geometric point of view, the problem in the example is that near the origin, both curves consist of two different "branches", but for the curve C, the decomposition into branches does not happen in a *Zariski* neighborhood of the origin. In terms of functions, the polynomial $y^2 - x^2 - x^3$ cannot be factored in $\mathcal{O}_{C.o}$. Naively, to overcome the problem, we should work with smaller neighborhoods and, correspondingly, a larger class of functions. This is easy to establish in case $\mathbb{K} = \mathbb{C}$ where we may consider arbitraryly small Euclidean neighborhoods and allow convergent power series as functions on these:

$$y^2 - x^2 - x^3 = (y + x\sqrt{1+x}) \cdot (y - x\sqrt{1+x}),$$

where the Taylor series

$$\sqrt{1+x} = \sum_{k=0}^{\infty} \binom{1/2}{k} x^k$$

is convergent for $|x| < 1$. Ring theoretically, this suggests to consider the local ring

$$\mathbb{C}\{x_1 - a_1, \dots, x_n - a_n\} / \mathrm{I}(A)\, \mathbb{C}\{x_1 - a_1, \dots, x_n - a_n\}$$

instead of the local ring

$$\mathcal{O}_{A,p} \cong \mathcal{O}_{\mathbb{A}^n(\mathbb{C}),p} / \mathrm{I}(A)\, \mathcal{O}_{\mathbb{A}^n(\mathbb{C}),p}.$$

Over an arbitrary field $\mathbb{K}$, there is no analogue to the Euclidean topology, and it is not meaningful to speak of convergent power series. We, may, however, consider the local ring

$$\mathbb{K}[[x_1 - a_1, \ldots, x_n - a_n]]/\mathrm{I}(A)\,\mathbb{K}[[x_1 - a_1, \ldots, x_n - a_n]].$$

It turns out that this ring is naturally obtained from the local ring $\mathcal{O}_{A,p}$ by completing $\mathcal{O}_{A,p}$ with respect to the $\mathfrak{m}_{A,p}$-adic topology.

In what follows, we describe the construction of the completion in a general algebraic context: Let R be any ring, and let $\mathfrak{m}$ be any ideal of R. Considering the $\mathfrak{m}$-adic topology on R, we call two Cauchy sequences $(f_\nu), (g_\nu) \subset R$ equivalent if the sequence of differences $(f_\nu - g_\nu)$ converges to zero. The set of all equivalence classes of Cauchy sequences carries a natural ring structure: If $(f_\nu), (g_\nu) \subset R$ are Cauchy sequences, then so are $(f_\nu + g_\nu)$ and $(f_\nu \cdot g_\nu)$, and the classes of these depend only on the classes of (f_ν) and (g_ν). Suppressing the ideal $\mathfrak{m}$ in our notation, we write $\widehat{R}$ for the resulting ring, and call it the **completion of R with respect to $\mathfrak{m}$**. For each $f \in R$, the class of the constant sequence (f) is an element $\iota(f) \in \widehat{R}$. This defines a ring homomorphism $\iota : R \to \widehat{R}$. The kernel of ι is the ideal $\bigcap_{k=0}^{\infty} \mathfrak{m}^k$. Hence, we may consider R as a subring of $\widehat{R}$ if this ideal is zero, that is, if R is Hausdorff with respect to the $\mathfrak{m}$-adic topology. By Krull's intersection theorem, this holds, in particular, if $(R, \mathfrak{m})$ is a local Noetherian ring.

In treating the completion of affine rings and, similarly, that of $\mathcal{O}_{A,p}$, we make use of the following lemma.

Lemma 4.7.2. *Let $I \subset \Bbbk[x_1, \ldots, x_n]$ be an ideal, let $>$ be a degree-anticompatible monomial order on $\Bbbk[x_1, \ldots, x_n]$, and let $f_1, \ldots, f_r$ form a Gröbner basis for I. Then, for any $k \geq 1$, the f_i together with the monomials of degree k form a Gröbner basis for the ideal $I + \langle x_1, \ldots, x_n \rangle^k$.*

Proof. We write $\mathcal{G}$ for the set of proposed Gröbner basis elements. By assumption, the remainder in any standard expression for an S-polynomial of type $\mathrm{S}(f_i, f_j)$ in terms of $\mathcal{G}$ is zero. On the other hand, each term of an S-polynomial of type $\mathrm{S}(f_i, x^\alpha)$, where $|\alpha| = k$, has degree $\geq k$ since with respect to $>$, $\mathbf{L}(f_i)$ is chosen among the lowest degree terms of f_i. Hence, also in this case, Buchberger's test yields a remainder which is zero. □

Proposition 4.7.3. *If $R = \Bbbk[x_1, \ldots, x_n]/I$ is an affine ring, the completion of R with respect to the maximal ideal $\mathfrak{m} = \langle \overline{x}_1, \ldots, \overline{x}_n \rangle \subset R$ is*

$$\widehat{R} \cong \Bbbk[[x_1, \ldots, x_n]]/I\,\Bbbk[[x_1, \ldots, x_n]].$$

Proof. Let $I^e = I\,\Bbbk[[x_1, \ldots, x_n]]$. Given a power series $g = \sum_\alpha a_\alpha x^\alpha \in \Bbbk[[x_1, \ldots, x_n]]$, we write $g^{(\nu)}$ for the truncation $\sum_{|\alpha| \leq \nu} a_\alpha x^\alpha$. Associating to each g the sequence of truncations $(g^{(\nu)})$ and taking residue classes, we get a homomorphism

$$\phi : \Bbbk[[x_1, \dots, x_n]] \to \widehat{R}$$

with $I^e \subset \ker \phi$. The proposition will follow once we show that $I^e = \ker \phi$, and that ϕ is surjective. For this, fix a degree-anticompatible monomial order on $\Bbbk[x_1, \dots, x_n]$.

We first show that $I^e = \ker \phi$. Given $g \in \ker \phi$, let $h \in \Bbbk[[x_1, \dots, x_n]]$ be the normal form of g mod I^e. Then, in particular, no term of h is contained in $\mathbf{L}(I)$. Moreover, since $\phi(g) = 0$, also $\phi(h) = 0$. In terms of the truncations $h^{(\nu)}$ this means that for all $k \geq 0$, there is an index ν_0 such that $h^{(\nu)} + I \in \mathfrak{m}^k$ for all $\nu \geq \nu_0$. By Lemma 4.7.2, the latter condition is equivalent to $h^{(\nu)} \in I + \langle x_1, \dots, x_n \rangle^k$ for all $\nu \geq \nu_0$. Since k can be chosen arbitarily high, we have $\mathbf{L}(h) \in \mathbf{L}(I)$. By the choice of h, this is only possible if $h = 0$ and, thus, $g \in I^e$.

Next, we show that ϕ is surjective. For this, consider a sequence of polynomials (g_ν) in $\Bbbk[x_1, \dots, x_n] \subset \Bbbk[[x_1, \dots, x_n]]$ which represents a Cauchy sequence in R. For each ν, let $h_\nu \in \Bbbk[[x_1, \dots, x_n]]$ be the normal form of g_ν mod I^e. By Lemma 4.7.2, given $\nu, k \geq 0$, the truncation $h_\nu^{(k)}$ coincides with the normal form of g_ν mod $I + \langle x_1, \dots, x_n \rangle^{k+1}$. In particular, for each k, the sequence of polynomials $h_\nu^{(k)}$, $\nu \geq 0$, is ultimately constant, say $h_\nu^{(k)} = f_k$ for $\nu \gg 0$. Then $f_\ell - f_k \in \langle x_1, \dots, x_n \rangle^k$ for $\ell \geq k$. We conclude that the f_k constitute a power series whose image under ϕ in $\widehat{R}$ coincides with the class represented by (g_ν). □

Exercise 4.7.4. Let R be a ring, let $\mathfrak{m}$ be an ideal of R, and let $\widehat{R}$ be the completion of R with respect to $\mathfrak{m}$. Show:

1. If R is Noetherian, then so is $\widehat{R}$.
2. If R is Hausdorff with respect to the $\mathfrak{m}$-adic topology, then $\widehat{R}$ is complete with respect to $\mathfrak{m}\widehat{R}$.
3. If $\mathfrak{m}$ is a maximal ideal, then $\widehat{R}$ is a local ring with maximal ideal $\mathfrak{m}\widehat{R}$. Furthermore, $\widehat{R} = \widehat{R}_\mathfrak{m}$, where $\widehat{R}_\mathfrak{m}$ denotes the completion of the local ring $R_\mathfrak{m}$ with respect to its maximal ideal. □

Now, we focus on the completion of $\mathcal{O}_{A,p}$ with respect to $\mathfrak{m}_{A,p}$, denoted $\widehat{\mathcal{O}}_{A,p}$. By translating p to the origin and by either imitating the proof of Proposition 4.7.3 or by combining the proposition with part 3 of the exercise, we get:

Corollary 4.7.5. *Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p = (a_1, \dots, a_n) \in A$ be a point. Then*

$$\widehat{\mathcal{O}}_{A,p} \cong \mathbb{K}[[x_1 - a_1, \dots, x_n - a_n]]/\mathrm{I}(A)\,\mathbb{K}[[x_1 - a_1, \dots, x_n - a_n]]. \qquad \square$$

With respect to dimension, we have:

Corollary 4.7.6. *Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p \in A$ be a point. Then*

$$\dim \mathcal{O}_{A,p} = \dim \widehat{\mathcal{O}}_{A,p}. \qquad \square$$

Exercise 4.7.7. Prove Corollary 4.7.6.
Hint. Consider systems of parameters in both rings $\mathcal{O}_{A,p}$ and $\widehat{\mathcal{O}}_{A,p}$. Furthermore, consider the natural projection $\mathbb{K}[[x_1-a_1,\dots,x_n-a_n]] \to \widehat{\mathcal{O}}_{A,p}$ from Corollary 4.7.5 and make use of Exercise 1.9.3. □

Our next result refines Proposition 4.6.24. In particular, we show once more that the local ring of an algebraic set at a smooth point is an integral domain.

Corollary 4.7.8. *Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p \in A$ be a point.*

1. *If p is a smooth point of A, then*
$$\widehat{\mathcal{O}}_{A,p} \cong \mathbb{K}[[t_1,\dots,t_d]], \quad \textit{where} \ \ d = \dim_p A = \dim \mathcal{O}_{A,p}.$$
2. *More generally, if p is arbitrary, we have a representation of $\widehat{\mathcal{O}}_{A,p}$ as a quotient*
$$\widehat{\mathcal{O}}_{A,p} \cong \mathbb{K}[[t_1,\dots,t_e]]/J, \quad \textit{where} \ \ e = \dim_{\mathbb{K}} T_pA,$$
and where J is an ideal of $\mathbb{K}[[t_1,\dots,t_e]]$ such that $J \subset \langle t_1,\dots,t_e\rangle^2$.

Proof. We assume that $p = o$ is the origin.

1. By part 1 and the principal ideal theorem, any quotient of $\mathbb{K}[[t_1,\dots,t_d]]$ by a nonzero ideal J has dimension $< d$ since $\mathbb{K}[[t_1,\dots,t_d]]$ is an integral domain. Hence, part 1 is a special case of part 2.

2. If $\mathrm{I}(A) = \langle f_1,\dots,f_r\rangle$, then $T_pA = \mathrm{V}(d_pf_i \mid i = 1,\dots,r) \subset \mathbb{A}^n$. We may, hence, choose coordinates $x_1,\dots,x_n$ such that $d_pf_i = x_i$, for $i = 1\dots,n-e$, and such that $f_i \in \langle x_1,\dots,x_n\rangle^2$, for $i > n-e$. Sending the t_i to the x_{n-e+i} and composing with the natural projection $\mathbb{K}[[x_1-a_1,\dots,x_n-a_n]] \to \widehat{\mathcal{O}}_{A,p}$ from Corollary 4.7.5, we get a ring homomorphism
$$\phi : \mathbb{K}[[t_1,\dots,t_e]] \to \mathbb{K}[[x_1,\dots,x_n]] \to \widehat{\mathcal{O}}_{A,o}.$$

To show that ϕ is surjective, fix a degree-anticompatible monomial order on $\mathbb{K}[x_1,\dots,x_n]$. Given an element $\widehat{g} \in \widehat{\mathcal{O}}_{A,o}$, choose a power series $g \in \mathbb{K}[[x_1,\dots,x_n]]$ representing $\widehat{g}$, and let h be the normal form of g mod $\mathrm{I}(A)$. Then h also represents $\widehat{g}$. Moreover, no term of h is contained in $\mathbf{L}(\mathrm{I}(A))$. Since $\mathbf{L}(f_i) = x_i$ for $i = 1,\dots,n-e$, it follows that h is in the image of $\mathbb{K}[[t_1,\dots,t_e]] \to \mathbb{K}[[x_1,\dots,x_n]]$.

To finish the proof, we note that $J := \ker\phi$ is contained in $\langle t_1,\dots,t_e\rangle^2$ since $f_{n-e+1},\dots,f_r \in \langle x_1,\dots,x_n\rangle^2$. □

In the situation of the corollary, the number $e = \dim_{\mathbb{K}} T_pA$ is called the **embedding dimension** of the pair (A,p). Note that always $n \geq e$. We say that (A,p) is **minimally embedded** in $(\mathbb{A}^n,p)$ if $n = e$.

Exercise 4.7.9. For $\mathcal{O}_o$ and $\mathbb{K}[[x_1,\dots,x_n]]$, formulate and prove statements analogous to those in Propositions 3.3.3 and 3.3.11 on Noether normalization respectively to those in the Unmixedness Theorem 3.3.12. □

Definition 4.7.10. *Given affine algebraic sets A, B and points $p \in A$, $q \in B$, we call the pairs (A, p) and (B, q)* **analytically isomorphic** *if $\widehat{\mathcal{O}}_{A,p} \cong \widehat{\mathcal{O}}_{B,q}$ as $\mathbb{K}$-algebras.* □

Example 4.7.11. In Example 4.7.1, the pairs (C, o) and (D, o) are analytically isomorphic. Indeed, by the formal inverse function theorem (see Exercise 4.4.30), the homomorphism

$$\phi : \mathbb{C}[[u, v]] \to \mathbb{C}[[x, y]]$$

obtained by substituting

$$u \mapsto x\sqrt{1+x} = x \sum_{k=0}^{\infty} \binom{1/2}{k} x^k, \ v \mapsto y,$$

is an isomorphism. Since ϕ maps $v^2 - u^2$ to $y^2 - x^2 - x^3$, it induces the desired isomorphism

$$\widehat{\mathcal{O}}_{D,o} \cong \mathbb{C}[[u, v]]/\langle v^2 - u^2 \rangle \to \mathbb{C}[[x, y]]/\langle y^2 - x^2 - x^3 \rangle \cong \widehat{\mathcal{O}}_{C,o}.$$

□

In particular, the analytic type is a coarser invariant than the local ring. It is finer than the tangent space: If $R = \widehat{\mathcal{O}_{A,p}}$, and $\mathfrak{m}$ is the maximal ideal of R, then $\mathfrak{m}/\mathfrak{m}^2 \cong \mathfrak{m}_{A,p}/\mathfrak{m}_{A,p}^2$. Indeed there is a well-defined map $\mathfrak{m}_{A,p} \to \mathfrak{m} \to \mathfrak{m}/\mathfrak{m}^2$ of $\mathcal{O}_{A,p}$-modules which is surjective with kernel $\mathfrak{m}_{A,p}^2$.

Remark 4.7.12. Let $\mathbb{K} = \mathbb{C}$, let A, B be analytic sets, and let $p \in A$, $q \in B$ be points. Suppose that (A, p), (B, q) are minimally embedded in $(\mathbb{A}^e, o)$. Moreover, suppose that (A, p) and (B, q) are analytically isomorphic, where the isomorphism $\widehat{\mathcal{O}_{B,q}} \to \widehat{\mathcal{O}_{A,p}}$ is given by an e-tuple of *convergent* power series $(z_1, \dots, z_e)$. In this case, there are neighborhoods U of $p \in \mathbb{A}^e(\mathbb{C})$ and V of $q \in \mathbb{A}^e(\mathbb{C})$ in the Euclidean topology such that

$$z : U \to V,\ a \mapsto (z_1(a), \dots, z_e(a)),$$

is biholomorphic, and with $z(A \cap U) = B \cap V$. □

Exercise 4.7.13. Let p be a point of a plane curve $C \subset \mathbb{A}^2$.

1. Assume $\operatorname{char} \mathbb{K} \neq 2$. Show that p is a node respectively a cusp of C iff (C, p) is analytically isomorphic to $\mathrm{V}(y^2 - x^2)$ respectively $\mathrm{V}(y^2 - x^3)$.
2. Show that p is an ordinary triple point iff (C, p) is analytically isomorphic to $\mathrm{V}(xy(x - y))$. □

The precise definition of a tacnode is as follows (see Examples 4.3.1 and 4.3.6):

Definition 4.7.14. Assume $\operatorname{char} \mathbb{K} \neq 2$. A point p of a plane curve $C \subset \mathbb{A}^2$ is called a **tacnode** if (C, p) is analytically isomorphic to $(\mathrm{V}(y^2 - x^4), o)$. □

Exercise 4.7.15. Let $f \in \mathbb{K}[x, y]$ be a square-free polynomial, and let $C = \mathrm{V}(f) \subset \mathbb{A}^2$.

1. Show that C has at most nodes as singularities iff $\langle f, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\rangle \subset \mathbb{K}[x, y]$ is a radical ideal.
2. Show that C has at most double points as singularities iff
$$\langle f, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y^2}\rangle = \langle 1\rangle \subset \mathbb{K}[x, y].$$
3. Formulate and prove a criterion for C to have at most nodes and cusps as singularities.
4. The curve defined by
$$f = x^4 + y^4 - 8x^3 + 18xy^2 + 18x^2 + \frac{27}{2}y^2 - 27$$
has only nodes and cusps as singularities. How many of each type are there? □

We, now, turn from the local ring to the tangent space. The drawback of T_pA is that it fails to approximate A near a singular point $p \in A$. In fact, in this case, the dimension of T_pA, which determines T_pA as a $\mathbb{K}$-vector space up to isomorphism, is simply too big. In this sense, T_pA is too coarse at a singular point. To overcome this failure, we introduce our second new invariant of A at p which is the tangent cone TC_pA. This coincides with T_pA at a smooth point, but is better behaved than T_pA at a singular point.

Recall that according to our definitions, the tangent space at a smooth point is the union of lines which can be seen as the analogue of limiting positions of secant lines in calculus. Mimicing this if A is not necessarily smooth at p gives the tangent cone.

We suppose for simplicity that $p = o = (0, \dots, 0) \in A$ is the origin. Then the lines through p admit parametrizations of type $t \to tv$, where $v \in \mathbb{A}^n$, and every secant line to A through p gives a point $tv \in A$ with $t \neq 0$. We are interested in what is happening if t tends to zero. Consider the set
$$B = \{(v, t) \in \mathbb{A}^n \times \mathbb{A}^1 \mid tv \in A\} \subset \mathbb{A}^n \times \mathbb{A}^1 \cong \mathbb{A}^{n+1}.$$

As we will see more clearly in the proof of proposition 4.7.16 below, B is an algebraic set. Obviously, $B_1 = \mathbb{A}^n \times \{o\}$ is an irreducible component of B (if $B \neq \mathbb{A}^n$). We write $B_2 = \overline{B \setminus B_1}$ for the residual algebraic set. The **tangent cone** of A at o is defined to be the algebraic set
$$TC_oA = B_1 \cap B_2 \subset \mathbb{A}^n \times \{o\} \cong \mathbb{A}^n.$$

In determining equations for the tangent cone, given a polynomial $f \in \Bbbk[x_1, \dots, x_n]$, we write $m = \mathrm{mult}(f, o)$, and denote by f_i the homogeneous component of f of degree i.

Proposition 4.7.16. *Let $I \subset \Bbbk[x_1, \dots, x_n]$ be an ideal, and let $A = \mathrm{V}(I) \subset \mathbb{A}^n$. Suppose that A contains the origin o. Then, with notation as above, the tangent cone $TC_oA \subset \mathbb{A}^n$ is the locus of zeros of the ideal*

$$J = \langle \{ f_m \mid f \in I \} \rangle.$$

Proof. The set $B \subset \mathbb{A}^n \times \mathbb{A}^1$ is the common vanishing locus of the polynomials

$$f(tx) = t^m f_m(x) + t^{m+1} f_{m+1}(x) + \ldots + t^d f_d(x), \quad f \in I$$

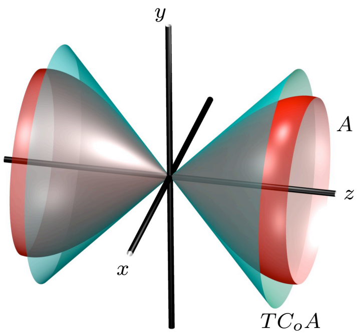
(note that $m \geq 1$ since $o \in A$). Saturating with respect to t, we obtain equations for the algebraic set residual to $B_1 = \mathrm{V}(t)$. That is, $B_2 \subset \mathbb{A}^n \times \mathbb{A}^1$ is the common vanishing locus of the polynomials

$$f_m(x) + t f_{m+1}(x) + \ldots + t^{d-m} f_d(x), \quad f \in I.$$

As a subset of $\mathbb{A}^n$, the intersection $B_1 \cap B_2 \subset \mathbb{A}^n \times \{o\} \cong \mathbb{A}^n$ is, then, defined by the ideal J. □

In particular, if $A = \mathrm{V}(f) \subset \mathbb{A}^n$ is a hypersurface with $o \in A$, then TC_oA is defined by the vanishing of the lowest degree part of f.

Example 4.7.17. If $A = \mathrm{V}(x^2 + y^2 - z^2 + z^4)$, then $TC_oA = \mathrm{V}(x^2 + y^2 - z^2)$.



□

Being defined by homogenous polynomials, TC_oA is the union of lines through the origin and, thus, indeed a cone: With notation as in the proposition, if $o \neq p \in TC_oA$ is a point, and $q = \lambda p$ is any point on the line $\overline{op}$, then $f_m(q) = \lambda^m f_m(p) = 0$ for all $f_m \in J$, so that $q \in TC_oA$ as well. See also Exercise 4.7.19, where we will give an alternative description of the tangent cone. Furthermore, note that TC_oA is contained in the tangent space T_oA. In fact, according to our definitions, if $I = \mathrm{I}(A) \subset \mathbb{K}[x_1, \dots, x_n]$ is the vanishing ideal of A, then the linear polynomials in J define the tangent space T_oA.

Exercise 4.7.18. In the situation of the proposition, let $f_1, \dots, f_r$ be a Gröbner basis for the ideal $I\mathcal{O}_o$ with respect to a degree-anticompatible monomial order on $\Bbbk[x_1, \dots, x_n]$. Then show that $TC_oA = \mathrm{V}((f_1)_m, \dots, (f_r)_m)$. □

Remark 4.7.19. In more abstract terms the ring of the tangent cone can be defined as the graded ring

$$grR = R/\mathfrak{m} \oplus \mathfrak{m}/\mathfrak{m}^2 \oplus \mathfrak{m}^2/\mathfrak{m}^3 \oplus \dots,$$

where R can be either the local ring or its completion. This shows that TC_pA depends only on $\widehat{\mathcal{O}_{A,p}}$. □

Example 4.7.20. Consider the algebraic set $A = \mathrm{V}(f_1, f_2, f_3, f_4) \subset \mathbb{A}^4$, where

$$\begin{aligned}
f_1 &= x_2^3 - x_1^2x_3 + x_1x_2x_4 - x_1x_3x_4 - x_2x_4^2 - x_1x_2,\\
f_2 &= x_1x_2^2 - x_1x_3^2 + 2x_2x_3x_4 - x_3^2x_4 - x_2x_3,\\
f_3 &= x_1^3 - x_1x_2x_3 + x_2^2x_4 + x_1x_4^2 - x_4^3 - x_1x_4,\\
f_4 &= x_1^2x_3 - x_2x_3^2 + x_1x_2x_4 + 2x_3x_4^2 - x_3x_4.
\end{aligned}$$

In the exercise below, we will show that these polynomials form a Gröbner basis with respect to $>_{\text{ldrlex}}$. Thus, the tangent cone of A at the origin $o \in \mathbb{A}^4$ is defined by the ideal

$$\langle x_1x_2, x_2x_3, x_1x_4, x_3x_4\rangle = \langle x_1, x_3\rangle \cap \langle x_2, x_4\rangle$$

which gives two planes in $\mathbb{A}^4$ intersecting at o. □

Exercise 4.7.21. Check the assertion about the Gröbner basis in Example 4.7.20. Then show that (A, o) and (TC_pA, o) are analytically isomorphic. □

In general, a singularity p of an algebraic set A is called an **improper node** if (A, o) and (TC_pA, o) are analytically isomorphic.

Exercise 4.7.22. Show that an ordinary quadrupel point is analytically isomorphic to a curve of type

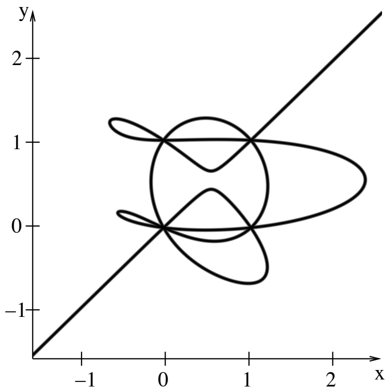
$$C_\lambda := \mathrm{V}(xy(y-x)(y-\lambda x)), \text{ where } \lambda \in \Bbbk \setminus \{0, 1\}.$$

Furthermore, show that two such curves C_λ and $C_{\lambda'}$ are analytically isomorphic iff

$$\lambda' \in \{\lambda,\, 1-\lambda,\, 1/\lambda, 1/(1-\lambda),\, (\lambda-1)/\lambda,\, \lambda/(\lambda-1)\}.$$ □

4.8 Additional Exercises

Exercise 4.8.1.
For the curve $\mathrm{V}(f) \subset \mathbb{A}^2(\mathbb{C})$ considered in part 2 of Exercise 4.1.5, determine the multiplicity at each singular point. Are all singular points ordinary multiple points?


y
2
1
0
−1
−1
0
1
2
x

□

Part II

Projective Algebraic Geometry

Chapter 5

Linear Systems of Plane Curves

This chapter provides a first impression of projective algebraic geometry. We will consider a new ambient space, projective n-space $\mathbb{P}^n(\Bbbk)$, which is obtained from affine n-space $\mathbb{A}^n(\Bbbk)$ by adding a "point at infinity in every direction". In this larger space, many geometric statements become simpler in that special cases are avoided.

The additional points form a hyperplane $H \subset \mathbb{P}^n(\Bbbk)$ which is often referred to as the "hyperplane at infinity". In fact, starting from a more formal definition of $\mathbb{P}^n(\Bbbk)$, we will see that there are many ways of writing $\mathbb{P}^n(\Bbbk)$ as the union of an "affine chart" $\mathbb{A}^n(\Bbbk)$ and a hyperplane at infinity. Local concepts can be extended from $\mathbb{A}^n(\Bbbk)$ to $\mathbb{P}^n(\Bbbk)$ by considering a covering of $\mathbb{P}^n(\Bbbk)$ by affine charts.

The introduction of homogeneous coordinates will allow us to define a projective algebraic set as the common locus of zeros of a collection of *homogeneous* polynomials. With respect to an affine chart, a projective algebraic set can be regarded as an affine algebraic set "completed" by adding relevant points at infinity (over the real or complex numbers, considering the Euclidean topology instead of the Zariski topology, the projective algebraic set is a natural compactification of the affine algebraic set). Postponing the general study of this and other facts about projective algebraic sets to the next chapter, we will, in this chapter, mainly focus on projective hypersurfaces, specifically on projective plane curves.

The natural parameter space for projective plane curves of a given degree d is a projective space itself. Its linear subspaces are classically known as linear systems of plane curves of degree d. They arise naturally in the context of a number of geometric questions. In fact, many geometric conditions on plane curves are linear in that the curves satisfying these conditions form a linear system. For instance, given a finite set of points in $\mathbb{P}^n(\Bbbk)$, we impose linear conditions by asking that the curves under consideration pass through these points (have multiplicities exceeding particular values at these points). After a basic treatment of linear systems in Section 5.3, we will use resultants to prove Bézout's theorem. Given two *projective* plane curves of degrees d, e

without a common component over an algebraically closed field, the theorem states that C and D intersect in $d \cdot e$ points, counted with multiplicity. As applications of Bézout's theorem, we will show how to bound the number of singular points of a plane curve and how to compute parametrizations of a rational plane curve with at most ordinary singularities.

In Section 5.5, we will treat Max Noether's fundamental theorem which, as we will see in Chapter 8, is central to the proof of the Riemann-Roch theorem given by Brill and Noether. Applications of Noether's result presented in this chapter are Pascal's theorem on the mystic hexagon and its generalizations.

In the final section of this chapter, we will define an addition law for points on cubic curves. We will use a general version of Pascal's theorem to show that this addition law is associative (and, thus, indeed a group law). We will, then, give a sketch of further results on cubic curves. In particular, we will adress the topology and the arithmetic of cubic curves.

5.1 Projective Space and Projective Algebraic Sets

In the *affine* plane, Bézout's theorem already fails in simple cases. For instance, two distinct circles have at most two points of intersection, even if we allow complex solutions and take multiplicities into account (see Exercise 5.3.10). Still simpler, two distinct lines do not intersect if they are parallel. The construction of the projective plane is custom-made to remedy the situation in the case of lines. As we will see in Section 5.4.8, it is universal enough to make Bézout's theorem hold in general.

Intuitively, we think of parallel lines as meeting at an "infinitely distant point" on the horizon (Renaissance painters referred to these points as *vanishing points* and used them as in Figure 5.2 to allow for perspective drawing):

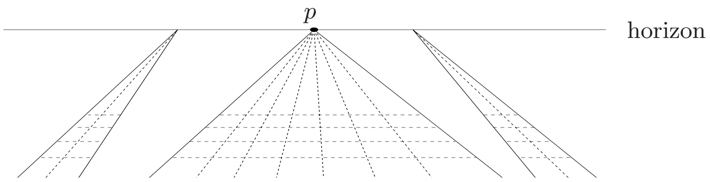


Fig. 5.1. *Vanishing points on the horizon*

Taking into account that the relation on lines in $\mathbb{A}^2(\mathbb{R})$ defined by 'is parallel to' is an equivalence relation, the idea is to require that all lines in a given equivalence class meet in the same point at infinity, with different classes corresponding to different points. Writing H for the set of all these points, we provisionally define the projective plane $\mathbb{P}^2(\mathbb{R})$ by setting

$$\mathbb{P}^2(\mathbb{R}) = \mathbb{A}^2(\mathbb{R}) \cup H.$$

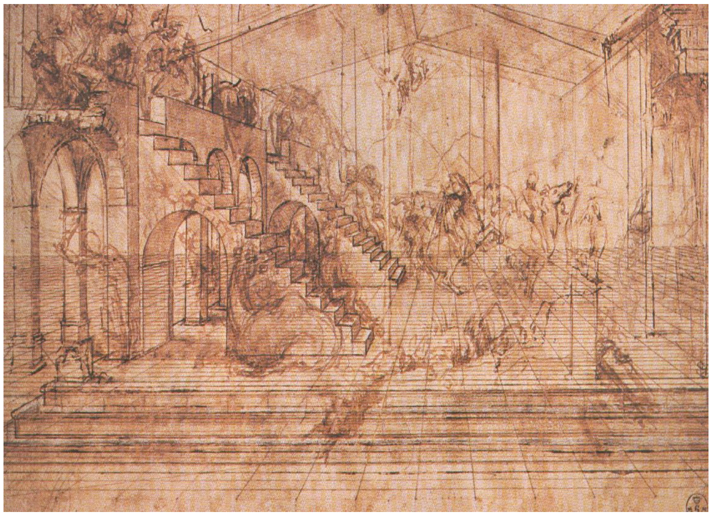


Fig. 5.2. *A sketch by Leonardo da Vinci*

A line in $\mathbb{P}^2(\mathbb{R})$ is, then, a line $L \subset \mathbb{A}^2(\mathbb{R})$ together with the common point at infinity of all lines parallel to L. Moreover, we regard H as a line in $\mathbb{P}^2(\mathbb{R})$, and call it the **line at infinity**. This makes sense since, now, any pair of distinct lines determines a unique point, and any pair of distinct points determines a unique line. Note that Figure 5.1 is somewhat inaccurate in that the horizon is not representing all points of H: It is missing the point at infinity of the lines “parallel to the horizon”.

Our provisional definition makes it cumbersome to work with $\mathbb{P}^2(\mathbb{R})$ since the points of $\mathbb{P}^2(\mathbb{R})$ are not treated on equal footing. To motivate the formal definition below, we write x_0, x_1, x_2 for the coordinates on the affine 3-space $\mathbb{A}^3(\mathbb{R})$, and choose $\mathrm{V}(x_0 - 1) \subset \mathbb{A}^3(\mathbb{R})$ as a reference plane for $\mathbb{A}^2(\mathbb{R})$:
Each point of $\mathbb{A}^2(\mathbb{R})$ determines, then, a line in $\mathbb{A}^3(\mathbb{R})$ through the origin o. In this way, we get all lines through o, except those lying in the plane $\mathrm{V}(x_0)$. The latter lines, in turn, form a copy of H. Indeed, the span of a given line $L \subset \mathbb{A}^2(\mathbb{R})$ and o intersects $\mathrm{V}(x_0)$ in a line through o which only depends on the class of lines parallel to L. We make the following general definition:

Definition 5.1.1. The **projective n-space** over the field $\Bbbk$ is the set

$$\begin{aligned}\mathbb{P}^n(\Bbbk) &= \{\text{lines through the origin in } \mathbb{A}^{n+1}(\Bbbk)\}\\ &= \{\text{one-dimensional linear subspaces of } \Bbbk^{n+1}\}.\end{aligned}$$

□

Considering a line L through the origin $o \in \mathbb{A}^{n+1}(\Bbbk)$ as an element of the new space $\mathbb{P}^n(\Bbbk)$, we call it a **point** of $\mathbb{P}^n(\Bbbk)$. If p denotes this point, then p is determined (or represented) by any point $(a_0, \dots, a_n) \in L \setminus \{o\}$. Accordingly, we write $p = [a_0 : \cdots : a_n]$, and call $a_0, \dots, a_n$ a set of **homogeneous**

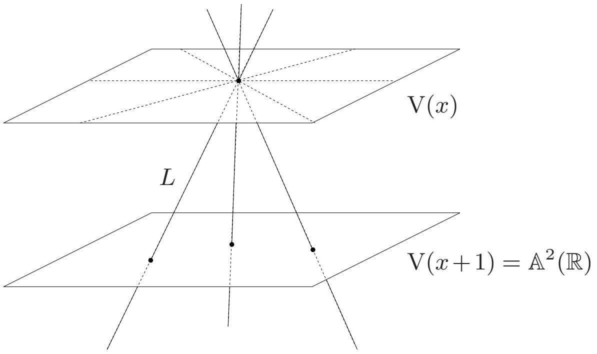


Fig. 5.3.

coordinates for p. Here, the colons and square brackets indicate that the homogeneous coordinates are determined up to a nonzero scalar multiple (if $a_i \neq 0$, the ratio $a_j : a_i$ depends on p only). Representing the points of $\mathbb{P}^n(\Bbbk)$ in this way means that we regard $\mathbb{P}^n(\Bbbk)$ as the quotient of $\mathbb{A}^{n+1}(\Bbbk) \setminus \{o\}$ modulo the equivalence relation defined by $(a_0, \dots, a_n) \sim (b_0, \dots, b_n)$ iff $(a_0, \dots, a_n) = \lambda(b_0, \dots, b_n)$ for some nonzero scalar λ:

$$\mathbb{P}^n(\Bbbk) \cong \big(\mathbb{A}^{n+1}(\Bbbk) \setminus \{o\}\big) \,/ \sim,$$

and we have the canonical projection

$$\pi : \mathbb{A}^{n+1}(\Bbbk) \setminus \{o\} \to \mathbb{P}^n(\Bbbk),\ (a_0, \dots, a_n) \mapsto [a_0 : \dots : a_n].$$

Remark-Definition 5.1.2. 1. It is often useful to have a basis-free definition of $\mathbb{P}^n$. If W is any $\Bbbk$-vector space of dimension $n+1$, then

$$\mathbb{P}(W) = \big\{\text{one-dimensional linear subspaces of } W\big\}$$

is called the **projective space of lines in *W***. Of course, after choosing a $\Bbbk$-basis for W, we can identify $\mathbb{P}(W)$ with $\mathbb{P}^n(\Bbbk)$, and regard the homogeneous coordinates on $\mathbb{P}^n(\Bbbk)$ as homogeneous coordinates on $\mathbb{P}(W)$.

2. If $(t_{ij}) \in \mathrm{GL}(n+1, \Bbbk)$ is an invertible matrix, the linear change of coordinates $x_i \mapsto \sum t_{ij} x_j$ induces a bijective map

$$T : \mathbb{P}^n(\Bbbk) \to \mathbb{P}^n(\Bbbk),\ [a_0 : \dots : a_n] \mapsto (\sum t_{0j} a_j : \dots : \sum t_{nj} a_j).$$

Any such map is called a **change of coordinates** of $\mathbb{P}^n(\Bbbk)$. Since multiples of the identity matrix act trivial, we are led to consider the group

$$\mathrm{PGL}(n+1, \Bbbk) := \mathrm{GL}(n+1, \Bbbk)/\Bbbk^*$$

which is called the **projective general linear group**. Later in the book, once we will have introduced morphisms between projective algebraic sets, we will see that any automorphism of $\mathbb{P}^n(\Bbbk)$ is an element of $\mathrm{PGL}(n+1, \Bbbk)$:

$$\mathrm{Aut}(\mathbb{P}^n(\Bbbk)) = \mathrm{PGL}(n+1, \Bbbk)$$

3. Two subsets $A, B \subset \mathbb{P}^n(\Bbbk)$ are called **projectively equivalent** if there is a change of coordinates T of $\mathbb{P}^n(\Bbbk)$ such that $T(A) = B$.
4. We say that $\mathbb{P}^1(\Bbbk)$ and $\mathbb{P}^2(\Bbbk)$ are the **projective line** and the **projective plane** over $\Bbbk$, respectively. □

In contrast to the affine case, the homogeneous coordinates do not constitute functions on $\mathbb{P}^n(\Bbbk)$. More generally, given any nonconstant polynomial $f \in \Bbbk[x_0, \dots, x_n]$, the value $f(a_0, \dots, a_n)$ depends on the choice of homogeneous coordinates for the point $p = [a_0 : \cdots : a_n] \in \mathbb{P}^n(\Bbbk)$ and can, therefore, not be called the value of f at p. Note, however, that if f is *homogeneous*, then $f(\lambda x_0, \dots, \lambda x_n) = \lambda^{\deg(f)} f(x_0, \dots, x_n)$ for all nonzero scalars λ, so that

$$f(a_0, \dots, a_n) = 0 \iff \forall\, \lambda \in \Bbbk \setminus \{0\} : \; f(\lambda a_0, \dots, \lambda a_n) = 0.$$

As a consequence, any homogeneous polynomial $f \in \Bbbk[x_0, \dots, x_n]$ has a well-defined **locus of zeros** (or **vanishing locus**) $\mathrm{V}(f)$ in $\mathbb{P}^n(\Bbbk)$. If f is nonconstant, we say that $\mathrm{V}(f)$ is a **hypersurface** in $\mathbb{P}^n(\Bbbk)$. A hypersurface in $\mathbb{P}^2(\Bbbk)$ is called a **projective plane curve**.

More generally, if $T \subset \Bbbk[x_1, \dots, x_n]$ is any subset of homogeneous polynomials, its **locus of zeros** (or **vanishing locus**) is the set

$$\mathrm{V}(T) = \{p \in \mathbb{A}^n(\Bbbk) \mid f(p) = 0 \text{ for all } f \in T\}.$$

If $T = \{f_1, \dots, f_r\}$ is finite, we write $\mathrm{V}(f_1, \dots, f_r) = \mathrm{V}(T)$.

Definition 5.1.3. A subset $A \subset \mathbb{P}^n(\Bbbk)$ is called an **algebraic subset** if $A = \mathrm{V}(T)$ for some subset $T \subset \Bbbk[x_0, \dots, x_n]$ of homogeneous polynomials. A **projective algebraic set** is an algebraic subset of some $\mathbb{P}^n(\Bbbk)$. □

Remark-Definition 5.1.4. As for $\mathbb{A}^n(\Bbbk)$, the **distinguished open sets**

$$\mathrm{D}(f) := \mathbb{P}^n(\Bbbk) \setminus \mathrm{V}(f), \quad f \in \Bbbk[x_0, \dots, x_n] \text{ homogeneous},$$

form the basis for a topology on $\mathbb{P}^n(\Bbbk)$ whose closed sets are the algebraic subsets of $\mathbb{P}^n(\Bbbk)$. This topology (the topology induced on any subset) is called the **Zariski topology** on $\mathbb{P}^n(\Bbbk)$ (on the subset). An algebraic subset of $\mathbb{P}^n(\Bbbk)$ is called **irreducible** (a **subvariety** of $\mathbb{P}^n(\Bbbk)$) if it cannot be written as a union of two strictly smaller closed subsets. A **projective variety** is a subvariety of some $\mathbb{P}^n(\Bbbk)$. Every nonempty Zariski open subset of a projective variety A is Zariski dense in A (see Proposition 1.11.8 and its proof). □

If not otherwise mentioned, subsets of $\mathbb{P}^n(\Bbbk)$ will carry the Zariski topology.

Exercise* 5.1.5. Recall that a map between topological spaces is said to be **open** if it sends open sets to open sets. Show: The canonical projection $\pi : \mathbb{A}^{n+1}(\Bbbk) \setminus \{o\} \to \mathbb{P}^n(\Bbbk)$ is continous and open with regard to the respective Zariski topologies. □

Remark-Definition 5.1.6. Given a subset of homogeneous polynomials $T \subset \Bbbk[x_0,\dots,x_n]$, rather than looking at the vanishing locus $A = \mathrm{V}(T) \subset \mathbb{P}^n$, we might also look at the vanishing locus of T in $\mathbb{A}^{n+1}$. This locus is a cone with vertex o: It is the union of all lines in $\mathbb{A}^{n+1}$ through o which correspond to points in A. We call this cone the **affine cone** over A, written $C(A)$. □

Classically, homogenous polynomials are known as **forms**. The adjectives **linear**, **quadratic**, **cubic**, **quartic**, **quintic** refer to forms of degree 1,2,3,4,5, respectively.

Example 5.1.7. The subsets of $\mathbb{P}^n(\Bbbk)$ defined by linear forms are precisely the subsets $\mathbb{P}(W) \subset \mathbb{P}^n(\Bbbk)$, where $W \subset \Bbbk^{n+1}$ is a linear subspace. Every such subset is called a **linear subspace** of $\mathbb{P}^n(\Bbbk)$ of dimension $\dim_{\Bbbk} W - 1$. Any two linear subspaces of the same dimension are projectively equivalent. Given a subset $\emptyset \neq X \subset \mathbb{P}^n(\Bbbk)$, there is a smallest linear subspace of $\mathbb{P}^n(\Bbbk)$ containing X. This subspace is called the **span** of X. A **line** in $\mathbb{P}^n(\Bbbk)$ is a linear subspace of dimension 1. A **plane** in $\mathbb{P}^n(\Bbbk)$ is a linear subspace of dimension 2. A **hyperplane** in $\mathbb{P}^n(\Bbbk)$ is a linear subspace of dimension $n-1$. □

Exercise 5.1.8. Let $p_0,\dots,p_n,p_{n+1} \in \mathbb{P}^n(\Bbbk)$ be a collection of $n+2$ points such that no subset of $n+1$ points is contained in a hyperplane. Show that there is a unique change of coordinates T of $\mathbb{P}^n(\Bbbk)$ such that

$$T(p_0) = [1:0:\dots:0], \dots, T(p_n) = [0:\dots:0:1], \\ \text{and } \; T(p_{n+1}) = [1:\dots:1].$$

The points $[1:0:\dots:0],\dots,[0:\dots:0:1]$ are known as the **coordinate points** of $\mathbb{P}^n(\Bbbk)$, and $[1:\dots:1]$ is the **scaling point**. □

Just as in our provisional definition of the real projective plane, we can write $\mathbb{P}^n(\Bbbk)$ as the union of $\mathbb{A}^n(\Bbbk)$ and a **hyperplane at infinity**:

$$\mathbb{P}^n(\Bbbk) = U_0 \cup H_0 \cong \mathbb{A}^n(\Bbbk) \cup \mathbb{P}^{n-1}(\Bbbk),$$

where

$$U_0 := \mathrm{D}(x_0) = \{[a_0:\dots:a_n] \in \mathbb{P}^n(\Bbbk) \mid a_0 \neq 0\},$$

and H_0 is the complement $H_0 = \mathbb{P}^n(\Bbbk) \setminus U_0 = \mathrm{V}(x_0)$. We identify H_0 with $\mathbb{P}^{n-1}(\Bbbk)$ by disregarding the first coordinate, and U_0 with $\mathbb{A}^n(\Bbbk)$ via

$$\varphi_0 : U_0 \to \mathbb{A}^n(\Bbbk),\; [a_0:\dots:a_n] = [1:\tfrac{a_1}{a_0}:\dots,\dots:\tfrac{a_n}{a_0}] \\ \mapsto (\tfrac{a_1}{a_0},\dots,\dots,\tfrac{a_n}{a_0}).$$

This map is bijective, with inverse

$$\mathbb{A}^n(\Bbbk) \to U_0,\; (b_1,\dots,b_n) \mapsto [1:b_1:\dots:b_n].$$

Given a point $p = [a_0:\dots:a_n] \in U_0$, the ratios a_i/a_0 are sometimes called the **affine coordinates** for p in U_0.

Example 5.1.9. In the special case of $\mathbb{P}^1(\mathbb{R})$, the map φ_0 sends a point $p = [a_0 : a_1] \in U_0$ to the slope a_1/a_0 of the line in $\mathbb{A}^2(\mathbb{R})$ corresponding to p.

The point [0:1] is the single point at infinity. It corresponds to the x_1-axis which is vertical and has, thus, slope ∞. □

The proof of our next proposition exhibits the geometric meaning of dehomogenization and homogenization.

Proposition 5.1.10. *The map $\varphi_0 : U_0 \to \mathbb{A}^n(\Bbbk)$ is a homeomorphism with regard to the respective Zariski topologies.*

Proof. Let $A \subset \mathbb{P}^n(\Bbbk)$ be a projective algebraic set. Then $A = \mathrm{V}(T)$ for some subset of homogeneous polynomials $T \subset \Bbbk[x_0, \dots, x_n]$. Let $T_a \subset \Bbbk[x_1, \dots, x_n]$ be obtained from T by dehomogenizing each element of T with respect to x_0. Then it is immediate from the definitions that $\varphi_0(A \cap U_0)$ is the algebraic set $\mathrm{V}_a(T_a) \subset \mathbb{A}^n(\Bbbk)$, where $\mathbf{V}_{\boldsymbol{a}}$ indicates that we look at the **affine vanishing locus**. Since the closed subsets of U_0 arise as intersections of type $A \cap U_0$, the map φ_0 is closed.

Conversely, let $A \subset \mathbb{A}^n(\Bbbk)$ be an affine algebraic set. Then $A = \mathrm{V}_a(T_a)$ for some subset of polynomials $T_a \subset \Bbbk[x_1, \dots, x_n]$, and it is easy to check that $\varphi_0^{-1}(A)$ is the closed subset $\mathrm{V}(T_a^h) \cap U_0$ of U_0, where $T_a^h \subset \Bbbk[x_0, \dots, x_n]$ is obtained from T_a by homogenizing each element of T_a with respect to x_0. Hence, the inverse map φ_0^{-1} is also closed. We conclude that φ_0 is a homeomorphism. □

Given an algebraic subset A of $\mathbb{P}^n(\Bbbk)$, we will identify $A \cap U_0$ with the algebraic set $\varphi_0(A \cap U_0) \subset \mathbb{A}^n(\Bbbk)$. Conversely, we will identify an algebraic subset A of $\mathbb{A}^n(\Bbbk)$ with $\varphi_0^{-1}(A) \subset \mathbb{P}^n(\Bbbk)$. Hence, the following definition makes sense:

Definition 5.1.11. If $A \subset \mathbb{A}^n(\Bbbk) \cong U_0$ is an algebraic subset, its Zariski closure $\overline{A}$ in $\mathbb{P}^n(\Bbbk)$ is said to be the **projective closure** of A. □

Remark 5.1.12. In Section 6.2, considering the homogenization of ideals, we will show how to compute the projective closure. In the special case of a hypersurface, if $f \in \Bbbk[x_1, \dots, x_n]$ is any nonconstant polynomial, and f^h is its homogenization with respect to x_0, the argument will show that

$$\overline{\mathrm{V}_a(f)} = \mathrm{V}(f^h) \subset \mathbb{P}^n(\Bbbk).$$

□

In accordance with our provisional definition of $\mathbb{P}^2(\mathbb{R})$, we have:

Example 5.1.13. In $\mathbb{P}^2(\Bbbk)$, the projective closure of a line in $\mathbb{A}^2(\Bbbk) \cong \mathrm{D}(x_0)$ with equation $x_2 = mx_1 + b$ is defined by the equation $x_2 = mx_1 + bx_0$. It intersects the line $\mathrm{V}(x_0)$ at infinity in the point $[0 : 1 : m]$. A line with equation $x_1 = c$ is completed by adding the point $[0 : 1 : 0]$. □

In the discussion above, there is nothing special with x_0: For $0 \leq i \leq n$, we define U_i, H_i and φ_i by using x_i instead of x_0. Then the U_i, which are known as the **(affine) coordinate charts** of $\mathbb{P}^n(\Bbbk)$, cover $\mathbb{P}^n(\Bbbk)$:

$$\mathbb{P}^n(\Bbbk) = \bigcup_{i=0}^{n} U_i.$$

Hence, $\mathbb{P}^n(\Bbbk)$ looks locally like $\mathbb{A}^n(\Bbbk)$, and we may study a projective algebraic set $A \subset \mathbb{P}^n(\Bbbk)$ by examining the different intersections $A \cap U_i$.

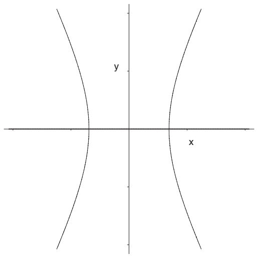
Example 5.1.14. Let $\Bbbk = \mathbb{R}$.

1. The projective closure C of the affine conic

$$\mathrm{V}(x^2 - 1/4y^2 - 1) \subset \mathbb{A}^2(\mathbb{R}) \cong \mathrm{D}(z)$$

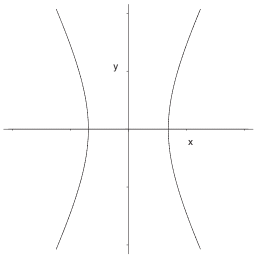
is defined by the quadratic form $x^2 - 1/4y^2 - z^2 = 0$. We show C in all three coordinate charts:

$z = 1$



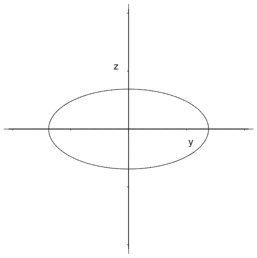
$x^2 - 1/4y^2 = 1$

$y = 1$



$x^2 - z^2 = 1/4$

$x = 1$

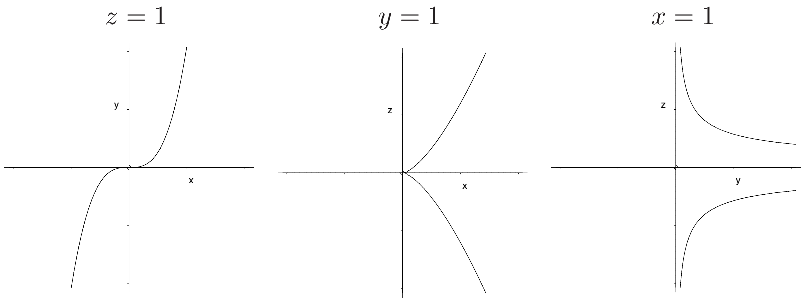


$1/4y^2 + z^2 = 1$

2. Similarly, starting from the affine curve

$$\mathrm{V}(y - x^3) \subset \mathbb{A}^2(\mathbb{R}) \cong \mathrm{D}(z),$$

we get the pictures below:


$z = 1$ $y = 1$ $x = 1$

$y = x^3$ $z^2 = x^3$ $yz^2 = 1$

□

Exercise 5.1.15. Draw the curve $\mathrm{V}(zy^2 - x^2z + x^3) \subset \mathbb{P}^2(\mathbb{R})$ in all three coordinate charts. For each chart, determine the points of the curve which lie on the line at infinity. Similarly for the curves in the previous example. □

Exercise 5.1.16. A **conic** in $\mathbb{P}^2(\mathbb{k})$ is defined by a nonzero quadratic form.

1. Show: A conic in $\mathbb{P}^2(\mathbb{R})$ is projectively equivalent to one of the following:

a)	$\mathrm{V}(x^2 + y^2 - z^2)$	(nondegenerate conic)
b)	$\mathrm{V}(x^2 + y^2 + z^2)$	(empty set)
c)	$\mathrm{V}(x^2 - y^2)$	(pair of lines)
d)	$\mathrm{V}(x^2 + y^2)$	(single point)
e)	$\mathrm{V}(x^2)$	("double" line).

2. Similarly, show that there are three classes of conics in $\mathbb{P}^2(\mathbb{C})$:

a)	$\mathrm{V}(x^2 + y^2 + z^2)$	(nondegenerate conic)
b)	$\mathrm{V}(x^2 + y^2)$	(pair of lines)
c)	$\mathrm{V}(x^2)$	("double" line).

3. More generally, show that quadric hypersurfaces in $\mathbb{P}^n(\mathbb{C})$ are classified by their rank. For this, recall that every quadratic form $f \in \mathbb{C}[x_0, \dots, x_n]$ may be written as
$$f(\boldsymbol{x}) = \boldsymbol{x}^t \cdot A \cdot \boldsymbol{x},$$
where $\boldsymbol{x}$ is the column vector with entries $x_0, \dots, x_n$, and where $A = (a_{ij})$ is a symmetric $(n+1) \times (n+1)$ matrix of scalars $a_{ij} \in \mathbb{C}$. The **rank** of the corresponding quadric $Q = \mathrm{V}(f) \subset \mathbb{P}^2(\mathbb{C})$ is defined to be the rank of A. Now show that Q has rank r iff it is projectively equivalent to a quadric with defining equation
$$\sum_{i=0}^{r} x_i^2 = 0.$$
If $r = n + 1$, then Q is **nondegenerate**.

4. Exactly, what invariants classify quadratic forms over $\mathbb{R}$?

By comparing the projective classification of conics with the classification of conics in the respective affine planes (work this out), you will find another example of how geometric statements become simpler if we pass from affine to projective geometry. In particular, as should be already clear from Example 5.1.14 and Figure 5.5, the difference between ellipses, parabolas, and hyperbolas disappears in the projective setting. □

Remark 5.1.17. In parts 2 and 3 of Exercise 5.1.16, we may replace $\mathbb{C}$ by any algebraically closed field of characteristic $\neq 2$. □

Before we go further, we adopt a convention which extends Convention 2.7.2:

Convention 5.1.18. *From now on,* $\mathbb{K}$ *will be an algebraically closed extension field of* $\Bbbk$*. We will write* $\mathbb{P}^n := \mathbb{P}^n(\mathbb{K})$*. If* $T \subset \Bbbk[x_0, \dots, x_n]$ *is a set of homogeneous polynomials, then* $A = \mathrm{V}(T)$ *will be its vanishing locus in* $\mathbb{P}^n$*. We will, then, say that* $\Bbbk$ *is a* ***field of definition*** *of* A*, or that* A *is* **defined over** $\Bbbk$*. A* $\Bbbk$***-rational point*** *of* A *is a point of the intersection*

$$A(\Bbbk) := A \cap \mathbb{P}^n(\Bbbk).$$

Furthermore, an element of $\mathrm{PGL}(n+1, \Bbbk) \subset \mathrm{PGL}(n+1, \mathbb{K})$ *will be called an* ***automorphism*** *of* $\mathbb{P}^n$ ***defined over*** $\Bbbk$*.* □

Remark 5.1.19. Convention 5.2.1 is justified by the projective Nullstellensatz which will be proved in the next chapter. The Nullstellensatz says, in particular, that hypersurfaces in $\mathbb{P}^n$ correspond to nonconstant square-free forms in $\mathbb{K}[x_0, \dots, x_n]$, where the form f is uniquely determined by the hypersurface H up to multiplication by a nonzero scalar. Then H is irreducible iff f is irreducible, and the degree of f is also called the **degree** of H. A hypersurface is a **quadric**, **cubic**, **quartic**, **quintic** if its degree is 2,3,4,5, respectively. □

As for the elements of the polynomial ring $\mathbb{K}[x_0, \dots, x_n]$, most elements of the rational function field $\mathbb{K}(x_0, \dots, x_n)$ cannot be regarded as functions in the projective context. However, if $g, h \in \mathbb{K}[x_0, \dots, x_n]$ are *forms* of the *same degree* d, then $f = g/h$ defines a function on $\mathrm{D}(h) \subset \mathbb{P}^n$. Indeed, in this case, substituting the homogeneous coordinates of a point $p \in \mathrm{D}(h)$ for the x_i in g and h gives a well-defined value $f(p)$:

$$\frac{g(\lambda x_0, \dots, \lambda x_n)}{h(\lambda x_0, \dots, \lambda x_n)} = \frac{\lambda^d g(x_0, \dots, x_n)}{\lambda^d h(x_0, \dots, x_n)} = \frac{g(x_0, \dots, x_n)}{h(x_0, \dots, x_n)}.$$

Specific examples are the affine coordinate functions x_j/x_i on $U_i = \mathrm{D}(x_i)$. The **rational function field** of $\mathbb{P}^n$ is the subfield

$$\mathbb{K}(\mathbb{P}^n) = \{g/h \in \mathbb{K}(x_0, \dots, x_n) \mid g, h \text{ forms of the same degree}\}$$

of $\mathbb{K}(x_0, \dots, x_n)$. Equivalently, $\mathbb{K}(\mathbb{P}^n)$ is the rational function field of any coordinate chart $U_i \cong \mathbb{A}^n$ (dehomogenize respectively homogenize to show that the two definitions give isomorphic fields). Similarly, we may define the **local ring of $\mathbb{P}^n$ at a point** $p \in \mathbb{P}^n$ either as the subring

$$\mathcal{O}_{\mathbb{P}^n,p} = \{g/h \in \mathbb{K}(\mathbb{P}^n) \mid \mathrm{D}(h) \ni p\} \subset \mathbb{K}(\mathbb{P}^n),$$

or as the local ring at p of any coordinate chart containing p. Concepts formulated in terms of the local ring can, then, be directly extended from the affine to the projective case. For instance, if $f \in \Bbbk[x_0, \dots, x_n]$ is a nonconstant form, and $p \in \mathbb{P}^n$ is a point, the **multiplicity of f at p**, written $\mathrm{mult}(f, p)$, is well-defined as the multiplicity at p of the dehomogenization of f in any chart U_i containing p. Similarly for the **intersection multiplicity** $i(f, g; p)$ of two nonconstant forms $f, g \in \Bbbk[x, y, z]$.

More generally, the local ring $\mathcal{O}_{A,p}$ of any projective algebraic set $A \subset \mathbb{P}^n$ at a point $p \in A$ can be defined in an analogous way, and in accordance with what is happening in the affine charts (we will treat this more systematically in Chapter 6). It makes, then, sense to say that p is a **smooth point** of A if $\mathcal{O}_{A,p}$ is a regular local ring. Equivalently, if U_i is any coordinate chart containing p, the affine algebraic set $A \cap U_i$ is smooth at p. The notions **singular point** and A_{sing} are as before.

Recall that $A \cap U_i$ is singular at p if the dimension of the tangent space to $A \cap U_i$ at p is strictly larger than the local dimension of $A \cap U_i$ at p. Though this can be checked in the chart U_i, it is occassionally useful to have a projective version of the tangent space:

Definition 5.1.20. Let $A \subset \mathbb{P}^n$ be a projective algebraic set, and let $p = [a_0 : \dots : a_n] \in A$ be a point. The **projective tangent space to A at p** is the linear subspace $T_pA \subset \mathbb{P}^n$ defined as follows: If A is a hypersurface, and $f \in \mathbb{K}[x_0, \dots, x_n]$ is a square-free form such that $A = \mathrm{V}(f)$, set

$$T_pA = \mathrm{V}\left(\sum_{i=0}^{n} \frac{\partial f}{\partial x_i}(a_0, \dots, a_n) \cdot x_i\right) \subset \mathbb{P}^n.$$

In the general case, let T_pA be the intersection of all projective tangent spaces at p to hypersurfaces containing A. □

Exercise* 5.1.21. If $f \in \mathbb{K}[x_0, \dots, x_n]$ is a square-free form, and $A = \mathrm{V}(f) \subset \mathbb{P}^n$, use Euler's rule to show:

1. If U_i is any coordinte chart containing p, then T_pA is the projective closure of the tangent space to the affine algebraic set $A \cap U_i$ at p.
2. If $C(A) \subset \mathbb{A}^{n+1}$ is the affine cone over A, and $q \in C(A)$ is any point representing p, the tangent space to $C(A)$ at q passes through the origin. It is, thus, a linear subspace W of $\mathbb{K}^{n+1}$. Furthermore, W is independent of the choice of q, and $T_pA = \mathbb{P}(W)$. □

For a hypersurface $A = \mathrm{V}(f)$ as in the exercise, p is a smooth point of A iff T_pA is a hyperplane. That is,

$$A_{\mathrm{sing}} = \mathrm{V}\left(f, \frac{\partial f}{\partial x_0}, \ldots, \frac{\partial f}{\partial x_n}\right).$$

If $\operatorname{char} \Bbbk$ does not divide $\deg f$, it is clear from Euler's rule that only the partial derivatives need to be considered.

Exercise 5.1.22. Determine the singular points of the curves in Example 5.1.14 and Exercise 5.1.15. □

In the discussion above, there is no need to restrict ourselves to coordinate charts: We may take any hyperplane H to be the hyperplane at infinity, regarding its complement U as affine n-space, and calling U an **affine chart**. Explicitly, if $H = \mathrm{V}(\sum \lambda_i x_i)$, where at least one λ_j is nonzero, identify

$$U \cong \mathbb{A}^n(\Bbbk)$$

via

$$[a_0 : \cdots : a_n] \mapsto \left(\frac{a_0}{\sum \lambda_i a_i}, \ldots, \widehat{\frac{a_j}{\sum \lambda_i a_i}}, \ldots, \frac{a_n}{\sum \lambda_i a_i}\right).$$

This is useful since a convenient choice of chart may ease explicit computations. Given any collection $y_0, \ldots, y_n$ of linearly independent linear forms, the $\mathrm{D}(y_i)$ form a covering of $\mathbb{P}^n(\Bbbk)$ which is obtained from the one given by the $\mathrm{D}(x_i)$ by a projective change of coordinates.

In Renaissance texts on perspective, the idea of considering different affine charts is a central theme. We illustrate this in Figure 5.5, where the reader may think of one chart as the floor in a medieval palace, of the other chart as the canvas of a painter, and of the origin $o \in \mathbb{A}^3(\mathbb{R})$ as the artist's eye.

In case $\Bbbk = \mathbb{R}$ respectively $\Bbbk = \mathbb{C}$, the projective space $\mathbb{P}^n(\Bbbk)$ also carries an Euclidean topology, namely the quotient topology induced from the Euclidean topology on $\Bbbk^{n+1} \setminus \{0\}$ via the canonical projection $\Bbbk^{n+1} \setminus \{0\} \to \mathbb{P}^n(\Bbbk)$.

Remark 5.1.23. 1. Let $\Bbbk = \mathbb{R}$ respectively $\Bbbk = \mathbb{C}$. Then $\mathbb{P}^n(\Bbbk)$ carries an Euclidean topology, namely the quotient topology induced by the canonical projection $\Bbbk^{n+1} \setminus \{0\} \to \mathbb{P}^n(\Bbbk)$. With respect to this topology, the coordinate charts exhibit $\mathbb{P}^n(\Bbbk)$ as a real respectively complex manifold, which is, in fact, compact. Indeed, we may regard $\mathbb{P}^n(\Bbbk)$ as the quotient

$$\mathbb{P}^n(\Bbbk) \cong S^n / \sim,$$

where

$$S = \{x \in \Bbbk^{n+1} \mid ||x|| = 1\}$$

is the (compact) unit sphere, and where $\sim$ refers to identifying antipodal points.

Compactness follows since $\mathbb{P}^n(\mathbb{R})$ is the image of the compact unit sphere

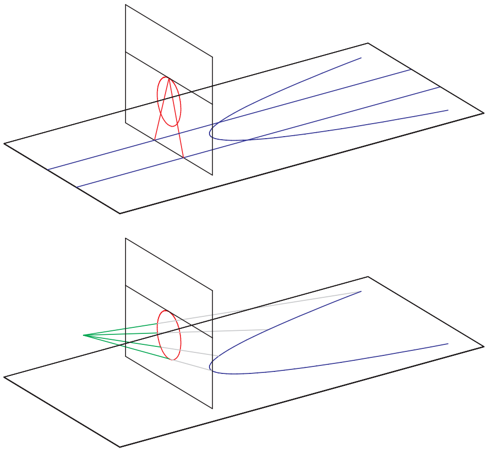


Fig. 5.4. *Different charts in perspective drawing.*

$$S = \{x \in \Bbbk^{n+1} \mid ||x|| = 1\} \subset \mathbb{R}^{n+1}.$$

In case $\Bbbk = \mathbb{C}$, $\mathbb{P}^n(\mathbb{C})$ is the image of the $2n+1$-dimensional unit sphere in $\mathbb{C}^{n+1}$.

2. Let $\Bbbk = \mathbb{C}$. If $f \in \mathbb{C}[x_1, \ldots, x_n\}$ ia a polynomial, then $\mathrm{V}(f^h)$ is not only the Zariski closure of $\mathrm{V}(f) \subset U_0 \subset \mathbb{P}^n(\mathbb{C})$, but also the closure of $\mathrm{V}(f)$ with respect to the Euclidian topology. To see this we may assume that f is irreducible. Then all affine hypersurface $\mathrm{V}(f^h) \cap U_i$ are irreducible hence path connected by Theorem 6.7.13 in Chapter

. The claim follows since the $\bigcup_i U_i$ covers $\mathbb{P}^n$. In particular we see that a projective hypersurface equipped with the Euclidean topology is compact as a closed subset of the compact manifold $\mathbb{P}^n(\mathbb{C})$. □

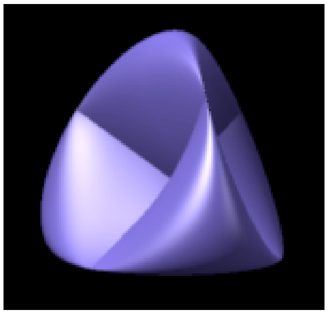
We will discuss the structure of the differentiable maps $S^n \to \mathbb{P}^n(\mathbb{R})$ and $S^{2n+1} \to \mathbb{P}^n(\mathbb{C})$ for small $n = 2$ respectively $n = 1$.

In the last two remarks of this section, we discuss the construction of $\mathbb{P}^2(\mathbb{R})$ and $\mathbb{P}^1(\mathbb{C})$ by focusing on their Euclidean topology. These considerations will not play a role in subsequent parts of the book.

Remark 5.1.24. The real projective plane $\mathbb{P}^2(\mathbb{R})$ has an interesting structure as a 2-dimensional real manifold. Every line through the origin in $\mathbb{R}^3$ intersects the unit sphere $S^2 = \{(x_0, x_1, x_2 \in \mathbb{R}^3 \mid x_0^2 + x_1^2 + x_2^2 = 1\}$ in two points. Thus

$$\mathbb{P}^2(\mathbb{R}) = S^2 / \sim,$$

where the equivalence relation $\sim$ identifies antipodal points. Thus as real manifold we obtain $\mathbb{P}^2(\mathbb{R})$ by gluing the Moebius strip, which is the image of a belt around the equator in S^2, with a disc, which is the image of the cap around the north (or south) pole. Hence, the manifold $\mathbb{P}^2(\mathbb{R})$ is not orientable. In particular, we cannot embed $\mathbb{P}^2(\mathbb{R})$ into $\mathbb{R}^3$. There are however models of $\mathbb{P}^2(\mathbb{R})$ in $\mathbb{R}^3$, if we allow self-intersections. The Steiner roman surface discussed in Example 2.6.6 ,

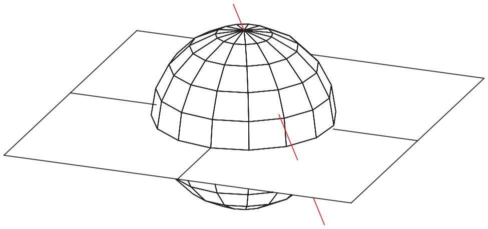


is such an example, because

$$S^2 \to \mathbb{R}^3,\ (x_0, x_1, x_2) \mapsto (x_1x_2, x_0x_2, x_0x_1)$$

factors over $\mathbb{P}^2(\mathbb{R})$. The map $\varphi : \mathbb{P}^2(\mathbb{R}) \to \mathbb{R}^3$ is an **immersion** at all points $p \in \mathbb{P}^2(\mathbb{R})$ except at the 6 pinch points on the coordinate axes. (A map between $\varphi : M \to N$ differential manifolds is a immersion at $p \in M$, if the induced map on the tangent spaces $d_p\varphi : T_pM \to T_pN$ is an inclusion. An immersion is a map which is an immersion everywhere). An immersion of $\mathbb{P}^2(\mathbb{R}) \to \mathbb{R}^3$ is given by the Boy surface.

Remark 5.1.25. For the complex projective line we have established two points of view. We can regard $\mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \infty \cong S^2$ via the projection from the north pole onto the Gaussian number plane.

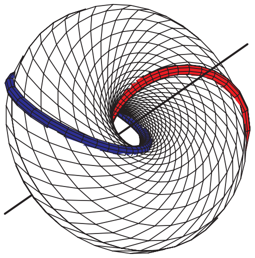


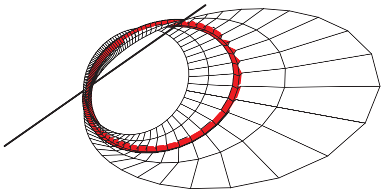
The other description realizes $\mathbb{P}^1(\mathbb{C})$ as the complex lines in $\mathbb{C}^2$. On the unit sphere $S^3 \subset \mathbb{C}^2$ every point in $\mathbb{P}^1$ has an $S^1 \cong \{z \in \mathbb{C} \mid |z| = 1\}$ of representatives. Combining both descriptions, we find a map

$$h : S^3 \to S^2,$$

whose fibers are S^1's, the **Hopf fibration**.

Identifying $S^3 = \mathbb{R}^3 \cup \{(1+i0, 0+i0)\}$ via stereographic projection, we find that $\mathbb{R}^3$ is fibered into an $\mathbb{R}^2$ of circles and a line.





Exercise 5.1.26. Prove that there is no continues section $\sigma\colon S^2 \to S^3$ of h, but that there exists a continues section of $h\colon S^3 \setminus h^{-1}(\infty) \to \mathbb{C}$. What is the closure of the graph in your example? □

5.2 The Extension of Basic Concepts

Coordinate charts allow us to extend concepts such as function fields, local rings, smoothness, tangent spaces, and dimension with almost no extra effort to the projective case. In this section, we will give some examples of how this works. First of all, we adopt a convention which adds to Convention 2.7.2:

Convention 5.2.1. *From now on, $\mathbb{K}$ will be an algebraically closed extension field of $\Bbbk$. We will write $\mathbb{P}^n := \mathbb{P}^n(\mathbb{K})$. If $T \subset \Bbbk[x_0, \dots, x_n]$ is a set of homogeneous polynomials, then $A = \mathrm{V}(T)$ will be its vanishing locus in $\mathbb{P}^n$. We will, then, say that $\Bbbk$ is a* ***field of definition*** *of A, or that A is* **defined over** *$\Bbbk$. A* ***$\Bbbk$-rational point*** *of A is a point of the intersection*

$$A(\Bbbk) := A \cap \mathbb{P}^n(\Bbbk).$$

Furthermore, an element of $\mathrm{PGL}(n+1, \Bbbk) \subset \mathrm{PGL}(n+1, \mathbb{K})$ will be called an ***automorphism*** *of $\mathbb{P}^n$* ***defined over*** *$\Bbbk$.* □

Remark 5.2.2. Convention 5.2.1 is justified by the projective Nullstellensatz which will be proved in the next chapter. The Nullstellensatz says, in particular, that hypersurfaces in $\mathbb{P}^n$ correspond to nonconstant square-free forms in $\mathbb{K}[x_0, \dots, x_n]$, where the form f is uniquely determined by the hypersurface H up to multiplication by a nonzero scalar. Then H is irreducible iff f is irreducible, and the degree of f is also called the **degree** of H. A hypersurface is a **quadric**, **cubic**, **quartic**, **quintic** if its degree is 2,3,4,5, respectively. □

As for the elements of the polynomial ring $\mathbb{K}[x_0, \dots, x_n]$, most elements of the rational function field $\mathbb{K}(x_0, \dots, x_n)$ cannot be regarded as functions in the projective context. However, if $g, h \in \mathbb{K}[x_0, \dots, x_n]$ are *forms* of the *same degree d*, then $f = g/h$ defines a function on $\mathrm{D}(h) \subset \mathbb{P}^n$. Indeed, in this case, substituting the homogeneous coordinates of a point $p \in \mathrm{D}(h)$ for the x_i in g and h gives a well-defined value $f(p)$:

$$\frac{g(\lambda x_0, \ldots, \lambda x_n)}{h(\lambda x_0, \ldots, \lambda x_n)} = \frac{\lambda^d g(x_0, \ldots, x_n)}{\lambda^d h(x_0, \ldots, x_n)} = \frac{g(x_0, \ldots, x_n)}{h(x_0, \ldots, x_n)}.$$

Specific examples are the affine coordinate functions x_j/x_i on $U_i = \mathrm{D}(x_i)$.

Definition 5.2.3. The **rational function field** of $\mathbb{P}^n$ is the subfield

$$\begin{aligned}\mathbb{K}(\mathbb{P}^n) &= \{g/h \in \mathbb{K}(x_0, \ldots, x_n) \mid g, h \text{ forms of the same degree}\} \\ &\subset \mathbb{K}(x_0, \ldots, x_n).\end{aligned}$$

The **local ring of $\mathbb{P}^n$ at a point** $p \in \mathbb{P}^n$ is the subring

$$\mathcal{O}_{\mathbb{P}^n,p} = \{g/h \in \mathbb{K}(\mathbb{P}^n) \mid \mathrm{D}(h) \ni p\} \subset \mathbb{K}(\mathbb{P}^n). \qquad \square$$

Note that $\mathcal{O}_{\mathbb{P}^n,p}$ is indeed a local ring. Note also that the definition of $\mathcal{O}_{\mathbb{P}^n,p}$ is consistent with our definition in the affine case: If U_i is a coordinate chart containing p, then $\mathcal{O}_{\mathbb{P}^n,p}$ is isomorphic to the local ring of $\mathbb{A}^n \cong U_i$ at p (dehomogenize; for the inverse map, homogenize).

Concepts formulated in terms of the local ring can, thus, be directly extended from the affine to the projective case. For instance, if $f \in \Bbbk[x_0, \ldots, x_n]$ is a nonconstant form, and $p \in \mathbb{P}^n$ is a point, the **multiplicity of f at p**, written $\mathrm{mult}(f, p)$, is well-defined as the multiplicity at p of the dehomogenization of f in any chart U_i containing p. In the same way, given two nonconstant forms $f, g \in \Bbbk[x, y, z]$ and a point $p \in \mathbb{P}^2$, we define the **intersection multiplicity of f and g at p**, written $i(f, g; p)$. As in Chapter 4, these notions carry over to hypersurfaces (plane curves) by considering square-free forms defining the hypersurfaces (plane curves).

More generally, the local ring $\mathcal{O}_{A,p}$ of any projective algebraic set $A \subset \mathbb{P}^n$ at a point $p \in A$ can be defined in an analogous way, and such that the definition is consistent with that in the affine case (we will treat this more systematically in Chapter 6). It makes, then, sense to say that p is a **smooth point** of A if $\mathcal{O}_{A,p}$ is a regular local ring. Equivalently, if U_i ia any coordinate chart containing p, the affine algebraic set $A \cap U_i$ is smooth at p. Otherwise, p is a **singular point** of A. As before, we write A_{sing} for the set of these points.

Recall that $A \cap U_i$ is singular at p if the dimension of the tangent space to $A \cap U_i$ at p is strictly larger than the local dimension of $A \cap U_i$ at p. Though this can be checked in the chart U_i, it is occassionally useful to have a projective version of the tangent space. Here is the definition in the hypersurface case (see Chapter 6 for the general case):

Definition 5.2.4. Let $A \subset \mathbb{P}^n$ be a hypersurface, let $p = [a_0 : \cdots : a_n] \in A$ be a point, and let $f \in \mathbb{K}[x_0, \ldots, x_n]$ be a square-free form such that $A = \mathrm{V}(f)$. The **projective tangent space T_pA to A at p** is the linear subspace

$$T_pA = \mathrm{V}\left(\sum_{i=0}^{n} \frac{\partial f}{\partial x_i}(a_0, \ldots, a_n) \cdot x_i\right) \subset \mathbb{P}^n. \qquad \square$$

Exercise* 5.2.5. In the situation of the definition, use Euler's rule to show:

1. If U_i is any coordinte chart containing p, then T_pA is the projective closure of the tangent space to the affine algebraic set $A \cap U_i$ at p.
2. If $C(A) \subset \mathbb{A}^{n+1}$ is the affine cone over A, and $q \in C(A)$ is any point representing p, the tangent space to $C(A)$ at q passes through the origin. It is, thus, a linear subspace W of $\mathbb{K}^{n+1}$. Furthermore, W is independent of the choice of q, and $T_pA = \mathbb{P}(W)$. □

With notation as in the definition, p is a smooth point of A iff T_pA is a hyperplane. That is,

$$A_{\text{sing}} = \mathrm{V}\left(f, \frac{\partial f}{\partial x_0}, \dots, \frac{\partial f}{\partial x_n}\right).$$

If $\operatorname{char} \Bbbk$ does not divide $\deg f$, it is clear from Euler's rule that only the partial derivatives need to be considered.

Exercise 5.2.6. Determine the singular points of the curves in Example 5.1.14 and Exercise 5.1.15. □

With regard to local studies, there is no need to restrict ourselves to coordinate charts: We may take any hyperplane H to be the hyperplane at infinity, regarding its complement U as affine n-space, and calling U an **affine chart**. Explicitly, if $H = \mathrm{V}(\sum \lambda_i x_i)$, where at least one λ_j is nonzero, identify

$$U \cong \mathbb{A}^n(\Bbbk)$$

via

$$[a_0 : \dots : a_n] \mapsto \left(\frac{a_0}{\sum \lambda_i a_i}, \dots, \widehat{\frac{a_j}{\sum \lambda_i a_i}}, \dots, \frac{a_n}{\sum \lambda_i a_i}\right).$$

This is useful since a convenient choice of chart may ease explicit computations. Given any collection $y_0, \dots, y_n$ of linearly independent linear forms, the $\mathrm{D}(y_i)$ form a covering of $\mathbb{P}^n(\Bbbk)$ which is obtained from the one given by the $\mathrm{D}(x_i)$ by a projective change of coordinates.

In Renaissance texts on perspective, the idea of considering different affine charts is a central theme. We illustrate this in Figure 5.5, where the reader may think of one chart as the floor in a medieval palace, of the other chart as the canvas of a painter, and of the origin $o \in \mathbb{A}^3(\mathbb{R})$ as the artist's eye.

5.3 Linear Systems of Plane Curves

The concept of linear systems is a classical concept of algebraic geometry. In this section, we study the special case of linear systems of plane curves. As motivation for this, we consider the following question:

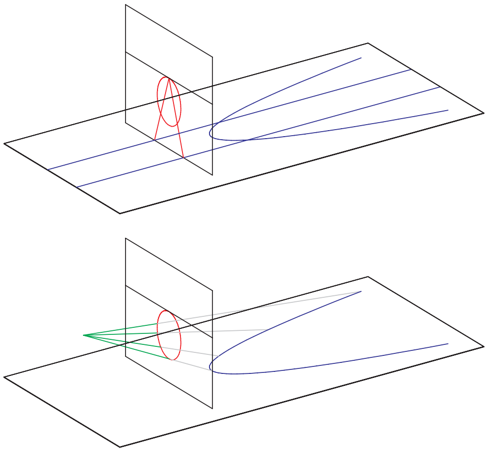


Fig. 5.5. *Different charts in perspective drawing.*

$$\begin{gathered}\text{Given } d \geq 1 \text{ and finitely many points in the projective plane,}\\ \text{how many curves of degree } d \text{ pass through these points?}\end{gathered} \tag{5.1}$$

To give the question a precise meaning, we describe the curves with the help of equations. We denote the coordinates by x, y, z. Recall from Remark 5.2.2 that each curve $C \subset \mathbb{P}^2$ of degree d is defined by a square-free form $f \in \mathbb{K}[x, y, z]$ of degree d, where f is determined up to multiplication by a nonzero scalar. In other words, C defines a point in the projective space $\mathbb{P}(L(d))$, where

$$L(d) = \mathbb{K}[x, y, z]_d = \{f \in \mathbb{K}[x, y, z] \mid f \ \text{ is homogenous of degree } \ d\}.$$

In $\mathbb{P}(L(d))$, there are also points corrsponding to polynomials with multiple factors. Nevertheless, we prefer to work with this space since the subset defined by the square-free polynomials is difficult to handle. By abuse of notation, we refer to every point of $\mathbb{P}(L(d))$ as a **projective plane curve of degree *d***, and to $\mathbb{P}(L(d))$ itself as a **parameter space for the plane curves of degree *d***. In speaking of components, of curves passing through a point, and of curves intersecting at a point, we extend the terminology introduced in Section 4.3 from the affine to the projective case.

Note that $\mathbb{P}(L(d))$ is a projective space of dimension

$$\binom{d+2}{2} - 1 = \frac{d(d+3)}{2}.$$

In fact, since the monomials of degree d form a $\mathbb{K}$-basis for $L(d)$, we may regard the coefficients of the polynomials in $L(d)$ as homogeneous coordinates on $\mathbb{P}(L(d))$ (listed in some order). Note that every change of coordinates of $\mathbb{P}^2$ induces a change of coordinates of $\mathbb{P}(L(d))$ (in the obvious way).

We can, now, illustrate question (5.1) by an example:

Example 5.3.1. Consider the four points

$$p_1 = [0 : 0 : 1],\ p_2 = [1 : 0 : 1],\ p_3 = [0 : 1 : 1],\ p_4 = [1 : 1 : 1] \in \mathbb{P}^2.$$

To describe the conics passing through these points, note that a quadratic polynomial

$$f = f_{20}x^2 + f_{11}xy + f_{10}xz + f_{02}y^2 + f_{01}yz + f_{00}z^2$$

vanishes at p_1, p_2, p_3, p_4 iff

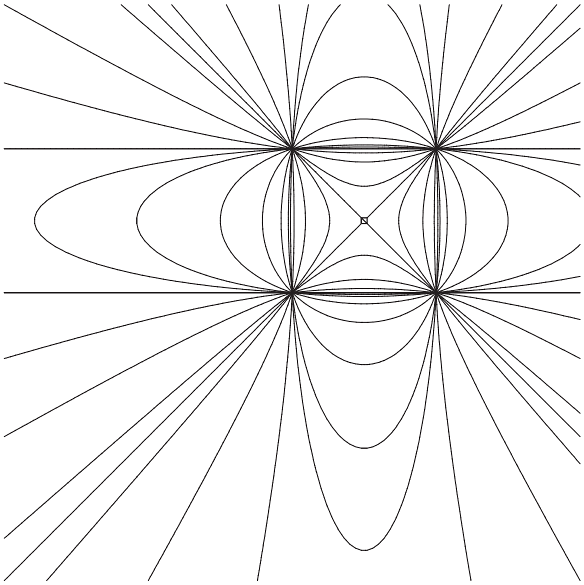
$$f_{00} = 0,\ f_{20} + f_{10} = 0,\ f_{02} + f_{01} = 0,\ f_{20} + f_{11} + f_{10} + f_{02} + f_{01} = 0.$$

This gives four linear conditions on the coefficients of f which are, in fact, independent – the conditions determine the two-dimensional linear subspace

$$L = \{\lambda x(x - z) + \mu y(y - z) \mid \lambda, \mu \in \mathbb{K}\} \subset \mathbb{K}[x, y, z]_2.$$

Geometrically, the generators $x(x - z)$ and $y(y - z)$ of L define two pairs of lines in $\mathbb{P}^2$ which, considered as points of $L(2)$, span the line $\mathbb{P}(L) \subset L(2)$. This line parametrizes the conics passing through p_1, p_2, p_3, p_4 – there is a $\mathbb{P}^1$ of such conics.

The following real picture shows the conics in the affine chart $\mathrm{D}(z)$:



□

It is clear from the example that "passing through a point $p \in \mathbb{P}^2$" imposes one linear condition on the curves of degree d – the curves passing through p form a hyperplane in $\mathbb{P}(L(d))$. More generally, as we will see in Proposition 5.3.3 below, we impose linear conditions by asking that the multiplicities of the curves at p exceed a given value r.

Definition 5.3.2. Let $d \geq 1$ be an integer.

1. A **linear system** of curves of degree d in $\mathbb{P}^2$ is a linear subspace $\mathbb{P}(L) \subset \mathbb{P}(L(d))$. A point $p \in \mathbb{P}^2$ is a **base point** of $\mathbb{P}(L)$ if every curve in $\mathbb{P}(L)$ passes through p. The words **pencil**, **net**, and **web** refer to a linear system of dimension 1,2, and 3, respectively.
2. If $p_1, \dots, p_s \in \mathbb{P}^2$ are distinct points, and $r_1, \dots, r_s \geq 1$ are integers, we write
$$L(d; r_1p_1, \dots, r_sp_s) := \{f \in L(d) \mid \operatorname{mult}(f, p_i) \geq r_i \text{ for all } i\},$$
and call
$$\mathbb{P}(L(d; r_1p_1, \dots, r_sp_s)) \subset \mathbb{P}(L(d))$$
the **linear system of curves of degree $\boldsymbol{d}$ with multiplicity at least $\boldsymbol{r_i}$ at $\boldsymbol{p_i}$, for all $\boldsymbol{i}$**. Moreover, we say that $p_1, \dots, p_s$ are the **assigned base points** of $\mathbb{P}(L(d; p_1, \dots, p_s))$. □

Proposition 5.3.3. *Let $p_1, \dots, p_s \in \mathbb{P}^2$ be distinct points, and let $r_1, \dots, r_s \geq 1$ be integers. Then $L(d; r_1p_1, \dots, r_sp_s)$ is a linear subspace of $L(d)$ of dimension*
$$\dim_{\mathbb{K}} L(d; r_1p_1, \dots, r_sp_s) \geq \binom{d+2}{2} - \sum_i \binom{r_i+1}{2}. \tag{5.2}$$

Proof. Since $L(d; r_1p_1, \dots, r_sp_s) = \bigcap_i L(d; r_ip_i)$, it suffices to treat the points seperately. After a change of coordinates, we may suppose that the given point is the point $p = [0:0:1]$. Then, a polynomial $f = \sum f_{\alpha\beta} x^\alpha y^\beta z^{d-\alpha-\beta} \in L(d)$ vanishes at p with multiplicity at least r iff $f_{\alpha\beta} = 0$ for all α, β with $\alpha+\beta < r$. The result follows since there are $\binom{r+1}{2}$ monomials $x^\alpha y^\beta$ with $\alpha + \beta < r$. □

Whether equality or strict inequality holds in (5.2) depends on whether the conditions imposed by the different points are linearly independent or not. Both cases do occur. In the example below, which illustrates this, we say that three or more points $p_1, \dots, p_s \in \mathbb{P}^2$ are **collinear** if the points lie on a line.

Example 5.3.4. For four distinct points $p_1, \dots, p_4 \in \mathbb{P}^2$, (5.2) gives
$$\dim_{\mathbb{K}} L(2; p_1, \dots, p_4) \geq 2.$$

If no three of these points are collinear, equality holds (make use of a suitable change of coordinates to reduce to the case treated in Example 5.3.1). If three of the points are collinear, say $p_1, p_2, p_3 \in L$, where $L \subset \mathbb{P}^2$ is a line, then L must be a component of any conic containing p_1, p_2, p_3 (one way of seeing this is to apply Bézout's theorem which will be proved in the next section). Hence, a conic through p_1, p_2, p_3 is determined by the component residual to L, which may be any line. If we require that the conic also contains p_4, and if $p_4 \notin L$, the residual line must pass through p_4 which imposes one extra linear condition. If $p_4 \in L$, there is no extra condition. We conclude that $\dim_{\mathbb{K}} L(2; p_1, \ldots, p_4) = 2$ iff $p_1, \ldots, p_4$ are not collinear, and that $\dim_{\mathbb{K}} L(2; p_1, \ldots, p_4) = 3$ if the four points lie on a line. □

In the example, the dimension of the linear system under consideration depends on the position of the points in the plane – for "almost all" collections of four points, the dimension is 2, and it is 3 only in the special case where the four points are collinear. To give "almost all" a more precise meaning, we say that a condition on a collection of points is satisfied for points $p_1, \ldots, p_s$ in general position if the points for which the condition is satisfied can be chosen in the following way: if $p_1, \ldots, p_r$, $r < s$, are already given, there is a nonempty Zariski open (hence dense) subset $U \subset \mathbb{P}^2$ such that we can choose p_{r+1} from among the points in U (in the example, if p_1, p_2, p_3 are not collinear, take $U = \mathbb{P}^2$; if p_1, p_2, p_3 lie on a line L, take $U = \mathbb{P}^2 \setminus L$). With this notation, we have:

Proposition 5.3.5. *Let $p_1, \ldots, p_s \in \mathbb{P}^2$ be distinct points in general position. If $\binom{d+2}{2} \geq s$, then*

$$\dim_{\mathbb{K}} L(d; p_1, \ldots, p_s) = \binom{d+2}{2} - s.$$

Proof. The result follows from the lemma below by induction on s. □

Lemma 5.3.6. *Let $\mathbb{P}(L) \subset \mathbb{P}(L(d))$ be a nonempty linear system. Then there is a nonempty Zariski open subset $U \subset \mathbb{P}^2$ such that $L \cap L(d;p) \subset L$ has codimension 1 for all $p \in U$.*

Proof. Since $L(d;p)$ is a hyperplane in $L(d)$, the linear subspace $L \cap L(d;p)$ of L has codimension one iff $L \not\subset L(d;p)$. But if f is any nonzero polynomial in L, then $f \notin L(d;p)$ for any point $p \in U := \mathbb{P}^2 \setminus \mathrm{V}(f)$. □

In the case where not each $r_i = 1$, it is an open problem to determine the tupels $(d, r_1, \ldots, r_n)$ for which the analogue of Proposition 5.3.5 holds (see Ciliberto and Miranda (2000) for some recent research).

Example 5.3.7. For five distinct points $p_1, \ldots, p_5 \in \mathbb{P}^2$, inequality (5.2) gives

$$\dim_{\mathbb{K}} L(4; 2p_1, \ldots, 2p_5) \geq 0.$$

However, we always have the sharper estimate

$$\dim_{\mathbb{K}} L(4; 2p_1, \ldots, 2p_5) \geq 1.$$

Indeed, since $\dim_{\mathbb{K}} L(2; p_1, \ldots, p_5) \geq 1$, there is a conic $\mathrm{V}(f)$ passing through all 5 points, and $f^2 \in L(4; 2p_1, \ldots, 2p_5)$. □

The remark below contains a simple example which prepares for the subsequent exercises:

Remark 5.3.8. If $p = [a_0 : a_1 : a_2],\ q = [b_0 : b_1 : b_2] \in \mathbb{P}^2(\Bbbk)$ are two distinct points, the unique line passing through p and q is defined by the determinantal equation

$$\det \begin{pmatrix} x_0 & x_1 & x_2 \\ a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \end{pmatrix} = 0.$$

□

Exercise 5.3.9. If $p_1, \dots, p_5 \in \mathbb{P}^2$ are five distinct points such that no three are collinear, show that there is a unique conic passing through the five points, show that this conic is nondegenerate, and give a determinantal equation for the conic. What happens if we only suppose that no four of the points are collinear? □

Exercise 5.3.10. If $p_1, p_2, p_3 \in \mathbb{R}^2$ are three points not lying on a line, show that there is a unique circle passing through these points, and give a determinantal equation for the circle.

Hint. Note that the set of all circles is an affine chart of a 3-dimensional linear system $L \subset \mathbb{P}(\mathbb{R}[x,y]_{\leq 2}) = \mathbb{P}(\mathbb{R}[x,y,z]_2)$. The base points of this system in $\mathbb{P}^2(\mathbb{C})$ are known as the **circle points**. Where do these points lie? □

If d is large enough, we always get the dimension expected from (5.2):

Proposition 5.3.11. *Let $p_1, \dots, p_s \in \mathbb{P}^2$ be distinct points, and let $r_1, \dots, r_s$ be integers ≥ 1. If $d \geq (\sum_i r_i) - 1$, then*

$$\dim_{\mathbb{K}} L(d; r_1 p_1, \dots, r_s p_s) = \binom{d+2}{2} - \sum_i \binom{r_i + 1}{2}.$$

Proof. We do induction on $m := (\sum_i r_i) - 1$. If $m \leq 1$, then either $s = 1$, or $s = 2$ and $r_1 = r_2 = 1$. In both cases, the result is clear. We may, hence, suppose that $d \geq m > 1$. In the induction step, we distinguish two cases.

Case 1. Suppose that each $r_i = 1$. Choose a linear form l_0 not vanishing at any p_i (this is possible since "not vanishing at a point" imposes a Zariski open (dense) condition on lines). In addition, for $i = 1, \dots, s-1$, choose linear forms l_i such that $p_i \in \mathrm{V}(l_i)$, but $p_j \notin \mathrm{V}(l_i)$ for $j \neq i$. Then $f := l_1 \cdots l_{s-1} \cdot l_0^{d-s+1} \in L(d; p_1, \dots, p_{s-1}) \setminus L(d; p_1, \dots, p_s)$. This shows that $L(d; p_1, \dots, p_s) \subsetneq L(d; p_1, \dots, p_{s-1})$, and we are done by applying the induction hypothesis.

Case 2. Now, suppose that not all $r_i = 1$, say $r := r_1 > 1$. Assume that $p_1 = [0 : 0 : 1]$, and set $L_0 = L(d; (r_1 - 1)p_1, r_2 p_2, \dots, r_s p_s)$. Then, for any $f \in L_0$, the dehomogenization $f(x, y, 1)$ is of type

$$f(x, y, 1) = \sum_{i=0}^{r-1} f_i x^i y^{r-1-i} + \text{ terms of higher degree.}$$

Setting $L_i = \{f \in L_0 \mid f_j = 0 \text{ for all } j < i\}$, we get a filtration

$$L_0 \supset L_1 \supset \ldots \supset L_r = L(d; r_1p_1, \ldots, r_sp_s).$$

Since the induction hypothesis applies to L_0, it suffices to show that $L_i \supsetneqq L_{i+1}$, $i = 0, \ldots, r-1$. For this, set $W_0 = L(d-1; (r-2)p_1, r_2p_2, \ldots, r_sp_s)$. Following the recipe above, define a filtration

$$W_0 \supset W_1 \supset \ldots \supset W_{r-1} = L(d-1; (r-1)p_1, r_2p_2, \ldots, r_sp_s).$$

By the induction hypothesis, $W_i \supsetneqq W_{i+1}$, $i = 0, \ldots, r-2$. Choosing polynomials $f_i \in W_i \setminus W_{i+1}$, we have $yf_i \in L_i \setminus L_{i+1}$, $i = 0, \ldots, r-2$, and $xf_{r-2} \in L_{r-1} \setminus L_r$. This concludes the proof. □

Exercise 5.3.12. For each set of integers $r_1, \ldots, r_s \geq 1$, show by example that the conclusion of the proposition may be wrong if $d = (\sum r_i) - 2$. □

5.4 Bézout's Theorem and Applications

The projective plane has been constructed such that any two distinct lines intersect in a unique point. The theorem of Bézout says that much more is true: given two curves in $\mathbb{P}^2$ of arbitrary degrees $d, e \geq 1$, the curves intersect in $d \cdot e$ points, counted with multiplicity. The proof of the theorem, which will be given in this section, is an application of elimination: we proceed by projecting the intersection points to a line. For this, we will consider the resultant which is a classical tool in elimination theory (its use can be traced back to work of Leibniz, Newton, Euler, and others – see the accounts in Kline (1972) and von zur Gathen and Gerhard (1999)).

Given two univariate polynomials f, g, the resultant of f and g is a polynomial expression in the coefficients of f and g which vanishes iff f and g have a nontrivial common factor (see Theorem 5.4.3 below). In the classical papers, the authors obtained the resultant by different ways of eliminating the variable from the system $f = g = 0$. Accordingly, there are different ways of representing the resultant. We will define it, here, as the determinant of the Sylvester matrix which provides one natural way of introducing linear algebra into the common factor problem.

Let R be a ring, and let

$$\begin{aligned} f = \quad & a_0x^d + a_1x^{d-1} + \ldots + a_d, \\ g = \quad & b_0x^e + b_1x^{e-1} + \ldots + b_e \quad \in R[x] \end{aligned} \tag{5.3}$$

be two polynomials of degrees $d, e \geq 1$. Then the **Sylvester matrix** of f and g is the matrix

$$\mathrm{Syl}(f,g) = \begin{pmatrix} a_0 & 0 & \dots & 0 & b_0 & 0 & \dots & 0 \\ a_1 & a_0 & & \vdots & b_1 & b_0 & & \vdots \\ \vdots & a_1 & \ddots & \vdots & \vdots & b_1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & a_0 & \vdots & \vdots & \ddots & b_0 \\ a_d & & & a_1 & b_e & & & b_1 \\ 0 & a_d & & \vdots & 0 & b_e & & \vdots \\ \vdots & & \ddots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a_d & 0 & 0 & \dots & b_e \end{pmatrix}.$$

Note that $\mathrm{Syl}(f,g)$ is a square matrix of size $d+e$: there are e colums containing a_i's, and d columns containing b_j's.

Definition 5.4.1. With notation as above, the **resultant** of f and g is the determinant

$$\mathrm{Res}(f,g) = \det \mathrm{Syl}(f,g) \in R. \qquad \square$$

Remark 5.4.2. No matter what ring we are considering, the resultant as a determinant can always be computed using the same recipe. We conclude that the construction of the resultant is universal in the following sense: If S is the polynomial ring

$$S = \mathbb{Z}[u_i, v_j \mid i = 0, \dots, d,\ j = 0, \dots, e]$$

in $d+e+2$ variables with integer coefficients, and

$$\begin{aligned} F = &\quad u_0x^d + u_1x^{d-1} + \dots + u_d, \\ G = &\quad v_0x^e + v_1x^{e-1} + \dots + v_e \quad \in S[x] \end{aligned}$$

are the "generic" polynomials in x of degrees d, e, then for any ring R and any two polynomials f, g as in (5.3), $\mathrm{Res}(f,g)$ is obtained from $\mathrm{Res}(F,G)$ by substituting the a_i, b_j for the u_i, v_j. $\square$

Theorem 5.4.3. *Let R be a UFD, and let $f, g \in R[x]$ be polynomials of degrees $d, e \geq 1$. Then f and g have a common factor of degree ≥ 1 iff $\mathrm{Res}(f,g) = 0$.*

Proof. Consider the "linear combination map"

$$\phi : R[x]_{<e} \oplus R[x]_{<d} \to R[x]_{<d+e},\ (A, B) \mapsto Af + Bg.$$

This is a map between two free R-modules of rank $d+e$ which, with respect to the R-bases

$$(x^{e-1}, 0), (x^{e-2}, 0), \dots, (1, 0), (0, x^{d-1}), \dots, (0, 1)$$

and

$$x^{d+e-1}, \dots, x, 1,$$

is represented by the Sylvester matrix $\operatorname{Syl}(f,g)$. We conclude that ϕ is injective iff $\operatorname{Res}(f,g) \neq 0$. On the other hand, since R is a UFD, ϕ is injective iff $\operatorname{GCD}(f,g) = 1$. Indeed, if $h := \operatorname{GCD}(f,g) \neq 1$, then $(-g/h, f/h) \in \ker\phi$. Conversely, suppose that $\operatorname{GCD}(f,g) = 1$, and let $(A,B)^t \in \ker\phi$ be a syzygy on f, g. Then $Af = -Bg$, which implies that B is a multiple of f. By degree reasoning, B and, thus, A are zero. □

Note that if $R = \Bbbk$ is a field, then f and g have a nontrivial common factor iff they have a common root in some algebraically closed extension field of $\Bbbk$. It was precisely the search for common (complex) roots which led the classical authors to consider the resultant.

Example 5.4.4. Computing the resultant of the two polynomials

$$f = 3x^2 + 5x - 2, \quad g = 7x^3 + x + 4 \in \mathbb{Q}[x],$$

we get

$$\operatorname{Res}(f,g) = \det \operatorname{Syl}(f,g) = \det \begin{pmatrix} 3 & 0 & 0 & 7 & 0 \\ 5 & 3 & 0 & 0 & 7 \\ 2 & 5 & 3 & 1 & 0 \\ 0 & 2 & 5 & 4 & 1 \\ 0 & 0 & 2 & 0 & 4 \end{pmatrix} = 1142792 \neq 0.$$

Hence, f and g do not have a common root in $\mathbb{C}$. □

Exercise 5.4.5. Let R be an integral domain, and let $f, g \in R[x]$ be two polynomials of degrees ≥ 1. Then show that

$$\operatorname{Res}(f,g) \in \langle f, g \rangle \cap R. \tag{5.4}$$

More precisely, show that there are polynomials $A, B \in R[x]$ such that $Af + Bg = \operatorname{Res}(f,g)$, $\deg A < \deg g$, and $\deg B < \deg f$. □

It is property (5.4) which links the resultant to elimination. Here are the details: Given two polynomials $f, g \in \Bbbk[x_1, \dots, x_n]$ of degree ≥ 1 in x_1, we may associate a resultant to f, g and the variable $x = x_1$ by regarding f, g as univariate polynomials in x_1. To indicate the distinguished variable, we, then, write $\operatorname{Syl}(f, g, x_1)$ for the Sylvester matrix and $\operatorname{Res}(f, g, x_1)$ for the resultant. Note that $\operatorname{Res}(f, g, x_1)$ is a polynomial in $R = \Bbbk[x_2, \dots, x_n]$ which, by (5.4), is contained in the first elimination ideal of $\langle f, g \rangle \subset \Bbbk[x_1, \dots, x_n]$. Moreover, if $(a_2, \dots, a_n) \in \mathbb{A}^{n-1}(\Bbbk)$ is a point such that neither of the leading coefficients of $f, g \in R[x_1]$ vanishes at $(a_2, \dots, a_n)$, then, by Remark 5.4.2,

$$\operatorname{Res}(f, g, x_1)(a_2, \dots, a_n) = \operatorname{Res}(f(x_1, a_2, \dots, a_n), g(x_1, a_2, \dots, a_n)). \tag{5.5}$$

The following exercise illustrates the use of this by an example which, at the same time, shows that $\operatorname{Res}(f, g, x_1)$ may fail to generate the elimination ideal.

Exercise 5.4.6. Consider the polynomials

$$f = xy^2 - xy - y^3 + 1, \quad g = x^2y^2 - x^2y + xy - 1 \in \mathbb{Q}[x, y].$$

1. Compute that

$$\begin{aligned}\operatorname{Res}(f, g, x) &= \det \begin{pmatrix} y^2 - y & 0 & y2 - y \\ -y^3 + 1 & y^2 - y & y \\ 0 & -y^3 + 1 & -1 \end{pmatrix} \\ &= y^8 - y^7 + y^6 - 3y^5 + y^4 + y^3 + y^2 - y \\ &= y(y-1)^2(y^5 + y^4 + 2y^3 - y - 1).\end{aligned}$$

   Since the resultant is contained in the elimination ideal $\langle f, g\rangle \cap \mathbb{Q}[y]$, the y-values of the complex solutions of $f = g = 0$ must be among its roots. This gives eight candidates for the y-values.
2. If $\pi : \mathbb{A}^2 \to \mathbb{A}^1$ is projection onto the y-component, show that

$$\pi(\mathrm{V}(f, g)) \subsetneq \mathrm{V}(\operatorname{Res}(f, g)).$$

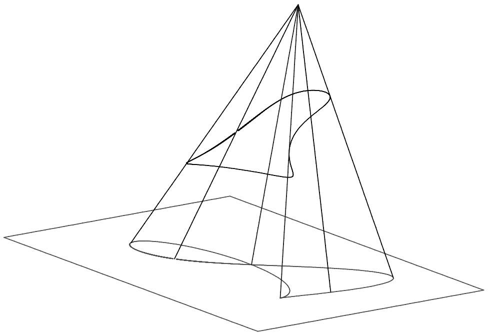
   Exactly, what y-value does not have a preimage point?
3. Use Gröbner bases to compute that the elimination ideal $\langle f, g\rangle \cap \mathbb{Q}[y]$ is generated by the polynomial $(y-1)^2(y^5 + y^4 + 2y^3 - y - 1)$. Compare this with the result of the previous part. □

Exercise* 5.4.7. Let $f, g \in \Bbbk[x_1, \dots, x_n]$ be forms of degrees $d, e \geq 1$. Suppose that both $f(1, 0, \dots, 0)$ and $g(1, 0, \dots, 0)$ are nonzero. That is, the leading coefficients of f and g – regarded as polynomials in x_1 – are nonzero constants. Then show that $\operatorname{Res}(f, g, x_1)$ is homogeneous of degree $d \cdot e$. □

In the projective setting, there is no value for the point $p = [1 : 0 : \cdots : 0] \in \mathbb{P}^n$ under projection onto the last n components. We are, thus, led to consider the projection map

$$\pi : \mathbb{P}^n \setminus \{p\} \to \mathbb{P}^{n-1}, \ [a_0 : \cdots : a_n] \mapsto [a_1 : \cdots : a_n].$$

More geometrically, think of $\mathbb{P}^{n-1}$ as the hyperplane $H_0 = \mathrm{V}(x_0) \subset \mathbb{P}^n$. Then the image of a point $q \in \mathbb{P}^n \setminus \{p\}$ under π is the intersection of the line spanned by p and q with H_0. More generally, if $H \subset \mathbb{P}^n$ is any hyperplane, and $p \in \mathbb{P}^n \setminus H$ is any point, the same recipe gives a map from $\mathbb{P}^n \setminus \{p\}$ to $H \cong \mathbb{P}^{n-1}$. This map is called **projection from $\boldsymbol{p}$ to $\boldsymbol{H}$**.



We can, now, prove Bézout's theorem:

Theorem 5.4.8 (Bézout). *Let $f, g \in \Bbbk[x, y, z]$ be forms of degrees $d, e \geq 1$. Assume that f and g have no common component. Then*

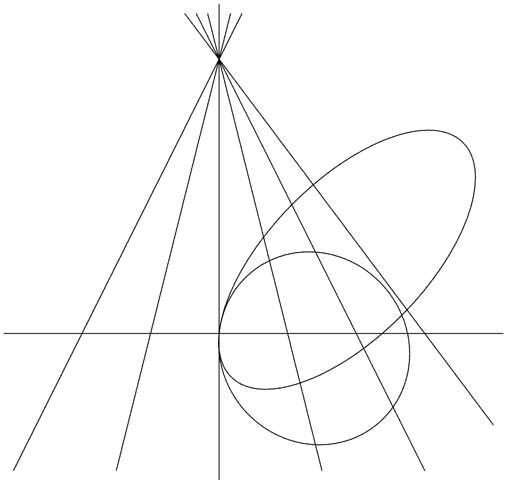
$$\sum_{p \in \mathbb{P}^2} i(f, g; p) = d \cdot e. \tag{5.6}$$

Proof. Step 1. It follows from the assumption on f, g and property 2 of intersection multiplicities (see Theorem 4.3.18) that $i(f, g; p) < \infty$ for each point p. Using the assumption again, we find that there are only finitely many intersection points of f and g (apply Exercise 1.7.13 in each coordinate chart). Since $i(f, g; p) \neq 0$ iff $p \in \mathrm{V}(f) \cap \mathrm{V}(g)$ (this is property 1 of intersection multiplicities), we conclude that the sum on the left hand side of (5.6) makes sense.

Step 2. Since $\mathrm{V}(f) \cup \mathrm{V}(g)$ is strictly contained in $\mathbb{P}^2$ by the Nullstellensatz, we may choose the coordinates such that the point $[0 : 1 : 0] \notin \mathrm{V}(f) \cup \mathrm{V}(g)$. That is, we may assume the leading coefficients of the forms f, g – regarded as polynomials in y – are nonzero constants. Let

$$\pi : \mathbb{P}^2 \setminus [0 : 1 : 0] \to \mathbb{P}^1,\ [a : b : c] \mapsto [a : c],$$

be projection from $[0 : 1 : 0]$ to the line $\mathrm{V}(y) \cong \mathbb{P}^1$:



Then a point $q = [a : c] \in \mathbb{P}^1$ is the image of a point $p \in \mathrm{V}(f) \cap \mathrm{V}(g)$ iff $f(a, y, c)$ and $g(a, y, c)$ have a common factor. Equivalently, by Theorem 5.4.3, the resultant

$$F := \mathrm{Res}(f, g, y) \in \Bbbk[x, z]$$

vanishes at $[a : c]$. By Exercise 5.4.7, F is a form of degree $d \cdot e$ which by Theorem 5.4.3 and the assumption on f, g, is nonzero. It follows that

$$\sum_{q \in \mathbb{P}^1} \mathrm{mult}(F, q) = d \cdot e$$

(counted with multiplicity, there are $\deg F(x, 1)$ zeros of F in the affine chart $\mathrm{D}(z)$ of $\mathbb{P}^1$, whereas $\mathrm{mult}(F, [1 : 0]) = \deg F - \deg F(x, 1)$).

To prove (5.6), it remains to show that the relevant multiplicities match: We claim that

$$\mathrm{mult}(F, q) = \sum_{\substack{p \in \mathrm{V}(f) \cap \mathrm{V}(g) \\ \pi(p) = q}} i(f, g; p),$$

for all points $q \in \mathbb{P}^1$ with $F(q) = 0$.

Step 3. Given a point $q \in \mathbb{P}^1$ as above, we may suppose after a further projective change of coordinates that $q = [0 : 1]$. Then all intersection points mapped to q lie in the affine chart $U = \mathrm{D}(z) \cong \mathbb{A}^2$ of $\mathbb{P}^2$. Thus, writing $f^a = f(x, y, 1)$, $g^a = g(x, y, 1)$, and $\mathcal{O}_p = \mathcal{O}_{\mathbb{A}^2, p}$, the claim from Step 2 reads

$$\mathrm{mult}(F(x, 1), 0) = \dim_{\mathbb{K}} \prod_{\substack{p \in \mathrm{V}(f) \cap \mathrm{V}(g) \\ \pi(p) = q}} \mathcal{O}_p / \langle f^a, g^a \rangle \mathcal{O}_p.$$

Step 4. Since there are only finitely many intersection points, Corollary 4.6.17 gives us an isomorphism of $\mathbb{K}$-algebras

$$M := \mathbb{K}[x,y]/\langle f^a, g^a\rangle \cong \prod_{p\in \mathrm{V}(f)\cap \mathrm{V}(g)\cap U} \mathcal{O}_p/\langle f^a, g^a\rangle \mathcal{O}_p. \qquad (5.7)$$

Step 5. To relate (5.7) to the claim in Step 3, we have to get rid of the intersection points which are not mapped to q. For this, we localize: Let $h \in \Bbbk[x]$ be a generator for $\langle F(x,1)\rangle : x^\infty$. Then h vanishes precisely at the points of $\mathrm{V}(f)\cap\mathrm{V}(g)\cap U \setminus \pi^{-1}(q)$. In algebraic terms, for $p \in \mathrm{V}(f)\cap\mathrm{V}(g)\cap U$, the residue class of h in $\mathcal{O}_p/\langle f^a, g^a\rangle\mathcal{O}_p$ is a unit if $p \in \pi^{-1}(q)$, and is nilpotent otherwise (recall from Step 1 that $\dim_{\mathbb{K}} \mathcal{O}_p/\langle f^a, g^a\rangle\mathcal{O}_p < \infty$). Hence, after inverting h on both sides of (5.7), we have

$$M[h^{-1}] := \mathbb{K}[x,y,h^{-1}]/\langle f^a, g^a\rangle \cong \prod_{\substack{p\in \mathrm{V}(f)\cap \mathrm{V}(g)\\ \pi(p)=q}} \mathcal{O}_p/\langle f^a, g^a\rangle \mathcal{O}_p. \qquad (5.8)$$

Step 6. Since M is generated by (the residue class of) y as a $\mathbb{K}[x]$-algebra, and since the leading coefficients of the forms f, g – regarded as polynomials in y – are nonzero constants, the powers $1, y, \dots, y^{\min(d,e)-1}$ generate M as a $\mathbb{K}[x]$-module. Working with the larger set of generators $y^{d+e-1}, \dots, y, 1$, and writing $R = \mathbb{K}[x]$, we get the free presentation

$$R[y]_{<e} \oplus R[y]_{<d} \xrightarrow{\phi} R[y]_{<d+e} \to M \to 0,$$

where ϕ is the linear combination map

$$(A,B) \mapsto Af + Bg.$$

This map is represented by the Sylvester matrix $\mathrm{Syl}(f^a, g^a, y)$ which, then, is also a representation matrix for $M[h^{-1}]$ considered as an $R[h^{-1}] = \mathbb{K}[x,h^{-1}]$-module. Since $R[h^{-1}]$ is a PID, and $M[h^{-1}]$ is annihilated by a power of x (this is clear from the right hand side of (5.8)), the structure theorem for modules over PID's gives that $\mathrm{Syl}(f^a, g^a, y)$ has a Smith normal form of type

$$\mathrm{Syl}(f^a, g^a, y) \underset{R}{\sim} \begin{pmatrix} x^{m_1} & 0 & \dots & 0 \\ 0 & x^{m_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x^{m_{d+e}} \end{pmatrix}$$

(see Example 2.8.8, Definition 2.8.9, and Exercise 2.8.10). We conclude that

$$\mathrm{mult}(F(x,1),0) = \mathrm{mult}(\det \mathrm{Syl}(f^a, g^a), 0) = \sum_{i=1}^{e+d} m_i = \dim_{\mathbb{K}} M[h^{-1}].$$

This finishes the proof. □

Example 5.4.9. Consider the quadratic forms

$$f = x^2 + y^2 - xz, \quad g = (x-y)^2 + 2(y+x)^2 - 3xz.$$

Then $[0:1:0] \notin \mathrm{V}(f) \cup \mathrm{V}(g)$. With notation as in the proof above, we have

$$\mathrm{Syl}(f^a, g^a, y) = \begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2x & 3 \\ x^2 - x & 0 & 3x^2 - 3x & 2x \\ 0 & x^2 - x & 0 & 3x^2 - 3x \end{pmatrix} \underset{\mathbb{K}[x]}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x^2(x-1) \end{pmatrix},$$

so that, as $\mathbb{K}[x]$-modules,

$$M = \mathrm{coker}\ \mathrm{Syl}(f^a, g^a, y) \cong \mathbb{K}[x]/\langle x \rangle \oplus \mathbb{K}[x]/\langle x^2 \rangle \oplus \mathbb{K}[x]/\langle x - 1 \rangle$$

(see Exercise 2.8.10). From this decomposition, it is clear that f and g intersect with multiplicity one at a point p_1 of type $p_1 = [1 : \beta_1 : 1]$, and one might be tempted to believe that there are two *distinct* intersection points $p_{2/3}$ of type $p_j = [0 : \beta_j : 1]$. This naive guess, however, is not true. One way of seeing this is to interchange the role of x and y in the proof of Bézout's theorem (note that $[1:0:0] \notin \mathrm{V}(f) \cup \mathrm{V}(g)$):

$$\mathrm{Syl}(f^a, g^a, x) = \begin{pmatrix} 1 & 0 & 3 & 0 \\ -1 & 1 & 2y-3 & 3 \\ y^2 & -1 & 3y^2 & 2y-3 \\ 0 & y^2 & 0 & 3y^2 \end{pmatrix} \underset{\mathbb{K}[y]}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & y & 0 \\ 0 & 0 & 0 & y^3 \end{pmatrix}.$$

Now, we conclude, that there is an intersection point p_2 of multiplicity at least three of type $p_2 = [\alpha_2 : 0 : 1]$ (and possibly another intersection point of the same type). Taking Bézout's Theorem into account and comparing with what we got above, we find that

$$\mathrm{V}(f) \cap \mathrm{V}(g) = \{[1:0:1], [0:0:1]\},$$

with intersection multiplicities

$$i(f, g; p_1) + i(f, g; p_2) = 1 + 3 = 4 = \deg C \cdot \deg D. \qquad \square$$

Exercise 5.4.10. Let $f, g \in \Bbbk[x, y, z]$ be nonconstant forms. Show that f and g intersect transversally at each point of $\mathrm{V}(f) \cap \mathrm{V}(g) \cap \mathrm{D}(z)$ iff, with notation as in the proof of Bézout's theorem, every elementary divisor of $\mathrm{Syl}(f^a, g^a, y)$ over $\mathbb{K}[x]$ is square-free. □

Exercise 5.4.11. Let $p_1, \dots, p_4 \in \mathbb{A}^2$ be four points in the affine plane such that no three are collinear. Then show that there is a parabola passing through these points iff $p_1, \dots, p_4$ do not form a parallelogram.
Hint: A parabola in $\mathbb{A}^2$ is the affine part of a nondegenerate conic in $\mathbb{P}^2$ which intersects the line at infinity in a single point with multiplicity 2. □

As an application of Bézout's Theorem, we give a bound on the number of singular points of a plane curve:

Theorem 5.4.12. *Let $C \subset \mathbb{P}^2$ be a curve of degree $d \geq 1$. If r_p denotes the multiplicity of C at a point $p \in \mathbb{P}^2$, then:*

1. C has at most $\binom{d}{2}$ singular points. More precisely,

$$\sum_{p \in C} \binom{r_p}{2} \leq \binom{d}{2}.$$

2. If C is irreducible, then C has at most $\binom{d-1}{2}$ singular points. More precisely,

$$\sum_{p \in C} \binom{r_p}{2} \leq \binom{d-1}{2}.$$

Proof. If $d = 1$, then C is a line, and there is nothing to show. We may, hence, assume that $d \geq 2$. Let $p_1, \dots, p_s$ be the distinct singular points of C, and write $r_i = r_{p_i}$. Moreover, let $f \in \mathbb{K}[x, y, z]$ be a square-free form defining C.

1. Since f is square-free, not all formal partial derivatives of f vanish, and we may suppose that $\frac{\partial f}{\partial x} \neq 0$. Then f and $\frac{\partial f}{\partial x}$ have no common component. Applying Bézout's theorem, we conclude that f and $\frac{\partial f}{\partial x}$ intersect in $d(d-1)$ points, counted with multiplicity. On the other hand, $\operatorname{mult}(\frac{\partial f}{\partial x}, p_i) \geq \operatorname{mult}(f, p_i) - 1 = r_i - 1$ for all i. Taking property 3 of intersection multiplicities into account (see Theorem 4.3.18), we get, as desired:

$$d(d-1) = \sum_i i(f, \frac{\partial f}{\partial x}, p_i) \geq \sum_i \operatorname{mult}(f, p_i) \cdot \operatorname{mult}(\frac{\partial f}{\partial x}, p_i) \geq \sum_i r_i(r_i - 1).$$

2. By Proposition 5.3.3 and part 1,

$$\dim_{\mathbb{K}} L(d-1; (r_1 - 1)p_1, \dots, (r_s - 1)p_s) \geq \binom{d+1}{2} - \sum_i \binom{r_i}{2} \geq d.$$

In particular, $t := \binom{d+1}{2} - \sum \binom{r_i}{2} - 1 \geq 1$, and we may choose smooth points $q_1, \dots, q_t \in C$. Once more applying Proposition 5.3.3, we see that we can find a nonzero form $g \in L(d-1; (r_1 - 1)p_1, \dots, (r_s - 1)p_s, q_1, \dots, q_t)$. Since, by assumption, f is irreducible, the forms f and g have no component in common. Making use of Bézout's theorem and arguing as in part 1, we get

$$d(d-1) \geq \sum_i r_i(r_i - 1) + t = \sum_i r_i(r_i - 1) + \frac{d^2 + d - 2}{2} - \sum_i \binom{r_i}{2}.$$

The desired bound follows. □

Theorem 5.4.13. *Let $C \subset \mathbb{P}^2$ be an irreducible curve of degree $d \geq 1$ such that $\binom{d-1}{2} = \sum_{p \in C} \binom{r_p}{2}$. Then C admits a rational parametrization.*

Proof. The basis idea is the same as in the proof of part 2 of Theorem 5.4.12. If we choose $t-1 = \binom{d+1}{2} - \sum \binom{r_i}{2} - 2$ addititional points $q_1, \dots, q_{t-1}$ on C then

$$\mathbb{P}(L(d-1;(r_1-1)p_1,\dots,(r_s-1)p_s,q_1,\dots,q_{t-1})) = \mathbb{P}(\langle g_0, g_1\rangle)$$

is a pencil of curves, whose intersection points except one with C are known to us. Thus, if $p(t_0,t_1)$ denotes the moving intersection point of $C \cap \mathrm{V}(t_0g_0+t_1g_1)$ then

$$\mathbb{P}^1 \to C,\ [t_0:t_1] \mapsto p(t_0,t_1)$$

is the desired parmetrization. This proves the Theorem for algebraically closed fields. Before we complete the proof for arbitrary fields, we discuss the resulting algorithm.

Remark 5.4.14. Suppose that C contains a smooth $\Bbbk$-rational point. Then C can be parametrized by rational functions defined over $\Bbbk$. □

Remark 5.4.15. Using the concept of the bihomogeneous coordinate ring $R = \Bbbk[x_0,x_1,x_2,t_0,t_1]$ of $\mathbb{P}^2 \times \mathbb{P}^1$, which we will introduce properly in Section 6, we can compute the parametrization explicitely as follows.

Let f be the equation of C. The zero locus of the ideal $J = \langle f, t_0g_0+t_1g_1\rangle \subset R$ decompose into

$$\mathrm{V}(J) = (B \times \mathbb{P}^1) \cup C' \subset \mathbb{P}^2 \times \mathbb{P}^1,$$

where $B = \mathrm{V}(g_1,g_2) \cap C \subset \mathbb{P}^2$ is the base loci of the pencil on C. The component C' is the graph of the desired parmetrization. Note that the two hypersufaces $C \times \mathbb{P}^1$ and $\mathrm{V}(t_0g_0+t_1g_1)$ intersect transversally along $(C' \setminus B) \times \mathbb{P}^1$, because the additional intersection is simple. Thus, if we saturate J in $\langle g_0,g_1\rangle$ and $\langle t_0,t_1\rangle$, we obtain the bihomogeneous ideal of $C' \subset \mathbb{P}^2 \times \mathbb{P}^1$:

$$\mathrm{I}(C') = (\langle f, t_0g_0+t_1g_1\rangle : \langle g_0,g_1\rangle^N) : \langle t_0,t_1\rangle^N \subset R$$

for N large enough. On the other hand, the rational map

$$\mathbb{P}^1 \to C,\ [t_0:t_1] \mapsto p(t_0,t_1) = [\varphi_0(t_0,t_1) : \varphi_1(t_0,t_1) : \varphi_2(t_0,t_1)],$$

is defined by three forms $\varphi_0,\varphi_1,\varphi_2 \in \Bbbk[t_0,t_1]$ of degree d. So $\mathrm{I}(C')$ contains the minors of the matrix

$$\begin{pmatrix} x_0 & x_1 & x_2 \\ \varphi_0 & \varphi_1 & \varphi_2 \end{pmatrix}.$$

Since there cannot be more than 3 equations of bi-degree $(1,d)$ in $\mathrm{I}(C')$, we can get the bi-graded piece $\mathrm{I}(C')_{(1,d)}$ spanned by these minors from $\mathrm{I}(C')$. Finally, to compute ${}^t(\varphi_0,\varphi_1,\varphi_2)$ from the space of minors $\langle m_0,m_1,m_2\rangle$, we calculate the syzygy of the matrix $(\frac{\partial m_i}{\partial x_j})_{i,j=0,1,2}$.

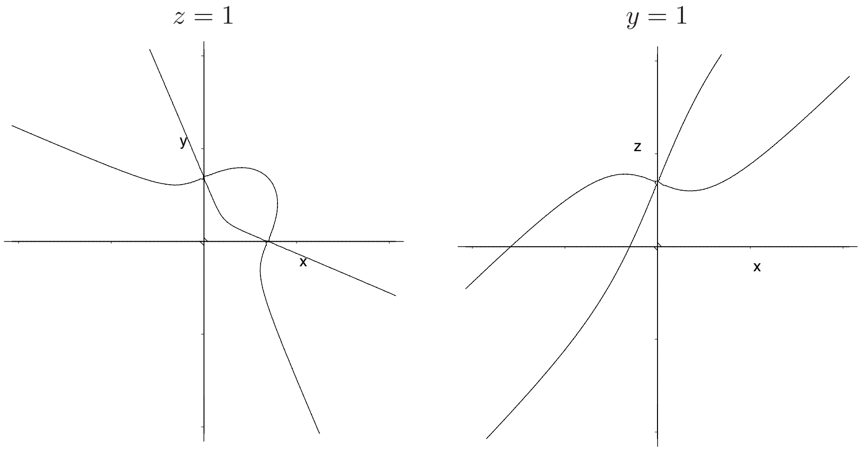
Returning to the proof of the theorem, suppose that the field of definition of C is $\mathbb{k} = \mathbb{Q}$. Then we would like the polynomials $\varphi_0, \varphi_1, \varphi_2 \in \mathbb{Q}[t_0, t_1]$ such that $\mathbb{P}^1(\mathbb{Q})$ parametrizes $C(\mathbb{Q})$ with perhaps of the exception of a few singular points. For this we need that $L(d-1; (r_1-1)p_1, \ldots, (r_s-1)p_s, q_1, \ldots, q_{t-1})$ is defined over $\mathbb{Q}$. For the singular points this is no problem: They might not be defined individually over $\mathbb{Q}$, but the collection $Sing_r = \{p \in C \mid \operatorname{mult}(C, p) = r\}$ is defined over $\mathbb{Q}$. So we need to find additional points $q_1, \ldots, q_{t-1}$ in C which are defined over $\mathbb{Q}$. A single point suffices if we alter the pencil.

Let $q \in C$ be a smooth point defined over $\mathbb{Q}$. Consider

$$L = \{g \in L(d-1; (r_1-1)p_1, \ldots, (r_s-1)p_s) \mid v_q(g) \geq t-1\}.$$

Then L has codimension at most $t-1$ in $L(d-1; (r_1-1)p_1, \ldots, (r_s-1)p_s))$ and $i(g, f; q) \geq t-1$ for every $g \in L$. Thus L is a pencil, again there is only one free intersection point and we obtain a parametrization defined over $\mathbb{Q}$. The same argument works for arbitrary fields of definition. □

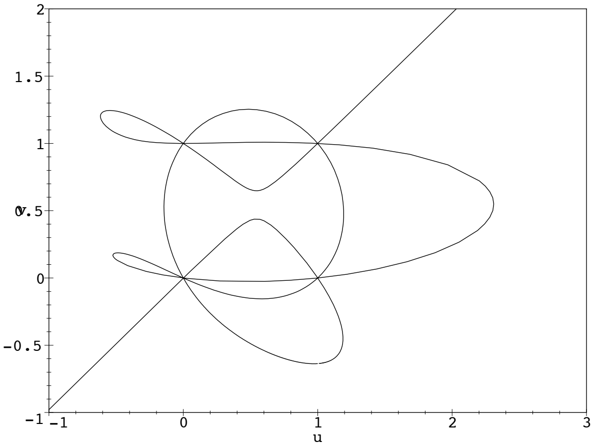
Exercise 5.4.16. Parametrize $V(2x^2y^2 - y^2(z-x-y)^2 - (z-x-y)^2x^2)$ over $\mathbb{Q}$.



□

Remark 5.4.17. In view of the application it is inconvenient, that we need a smooth rational point. Indeed this can be avoided as proved by Hilbert and Hurwitz [19xx]. A computer implementation of this algorithm was given in Maple packages [Winkler,19xx] and [?] and in Singular []. In general curves of odd degree defined over $\mathbb{Q}$ with $\binom{d-1}{2} = \sum_{p \in C} \binom{r_p}{2}$ always allow a $\mathbb{Q}$-rational parametrization. For even degree a quadratic field extension might be necessary, as we can see from the example of the conic $V(x^2 + y^2 + z^2)$, which has no real point, hence also no rational point.

Exercise 5.4.18. Compute a rational parametrization of the curve from Example 1.4.4,



$$\begin{aligned}11\,y^7 + 7\,y^6x + 8\,y^5x^2 - 3\,y^4x^3 - 10\,y^3x^4 - 10\,y^2x^5 - x^7 - 33\,y^6 - 29\,y^5x\\ -13\,y^4x^2 + 26\,y^3x^3 + 30\,y^2x^4 + 10\,yx^5 + 3\,x^6 + 33\,y^5 + 37\,y^4x - 8\,y^3x^2\\ -33\,y^2x^3 - 20\,yx^4 - 3\,x^5 - 11\,y^4 - 15\,y^3x + 13\,y^2x^2 + 10\,yx^3 + x^4 = 0\end{aligned}$$

without using additional rational points except the 4 singular points.

Hint: Use a suitable pencil of curves of degrees $\leq d-1$. □

5.5 Max Noether's Fundamental Theorem

Let $f, g \in \Bbbk[x, y, z]$ be two forms of degrees ≥ 1 without common components. Then f and g intersect in finitely many points, and we could ask: which other forms pass through these points? Of course, there are the obvious forms of type $h = Af + Bg$. In the special case where f and g intersect in $\deg f \cdot \deg g$ *distinct* points, it follows from Max Noether's theorem that there are no other possibilities. More generally, if we allow arbitrary intersection multiplicities, the theorem tells us that a form h is contained in the image of the linear combination map $(A, B) \mapsto Af + Bg$ iff this containment condition is fulfilled locally at each intersection point of f and g.

In formulating a precise statement, we use the following notation. Given a form $f \in \Bbbk[x, y, z]$ and a point $p \in \mathbb{P}^2$, choose a coordinate chart U containing p and set $f_p = f^a \in \mathcal{O}_p$, where f^a is the dehomogenization of f in U. Then f_p depends on the choice of U, but only up to multiplication by a unit in $\mathcal{O}_p$. Hence, the local conditions in Max Noether's theorem below make sense.

Theorem 5.5.1 (Max Noether's Fundamental Theorem). *Let f, g, h be forms of degrees ≥ 1 in $\Bbbk[x, y, z]$. Assume that f and g have no common component. Then there is an expression*

$$h = Af + Bg,$$

with forms $A, B \in \mathbb{K}[x, y, z]$ of degrees $\deg h - \deg f$*,* $\deg h - \deg g$*, iff*

$$h_p \in \langle f_p, g_p \rangle \subset \mathcal{O}_p$$

for every point $p \in \mathrm{V}(f) \cap \mathrm{V}(g)$.

Proof. Clearly, the global condition in the theorem implies the local ones. For the converse, arguing as is in the proof of Proposition 5.3.11, we can find a linear form not vanishing at any of the finitely many intersection points of f ang g. We may, hence, choose the coordinates such that $\mathrm{V}(f) \cap \mathrm{V}(g) \cap \mathrm{V}(z) = \emptyset$. That is, to work with the local conditions, we may dehomogenize with respect to z. We give the remaining part of the proof in two steps, consisting of an affine and projective argument, respectively.

Step 1. We write $f^a = f(x, y, 1)$, $g^a = g(x, y, 1) \in \mathbb{K}[x, y]$ and consider the composite map

$$\phi : \mathbb{K}[x, y, z] \to \mathbb{K}[x, y] \to \bigoplus_{p \in \mathrm{V}(f) \cap \mathrm{V}(g)} \mathcal{O}_p / \langle f_p, g_p \rangle$$

defined by

$$h \mapsto h^a = h(x, y, 1) \mapsto (h_p + \langle f_p, g_p \rangle)_{p \in \mathrm{V}(f) \cap \mathrm{V}(g)}.$$

The local conditions in the theorem imply $\phi(h) = 0$, so that $h^a \in \langle f^a, g^a \rangle$ by Corollary 4.6.17. Homogenizing, we get an equation of type

$$z^k h = A' f + B' g,$$

for some k and some forms $A', B' \in \mathbb{K}[x, y, z]$. The theorem will follow once we show that multiplication by z is injective on $\mathbb{K}[x, y, z] / \langle f, g \rangle$.

Step 2. Let an equation of type $zh' = A' f + B' g$ in $\mathbb{K}[x, y, z]$ be given. We show that $h' \in \langle f, g \rangle$. For this, if $E \in \mathbb{K}[x, y, z]$ is any polynomial, we write $E_0 = E(x, y, 0)$. We, then, have $A'_0 f_0 + B'_0 g_0 = 0$. On the other hand, since f and g have no common zero on the line $\mathrm{V}(z)$, the polynomials f_0 and g_0 have no common factor. It follows that $(A'_0, B'_0) = c \cdot (-g_0, f_0)$ for some $c \in \mathbb{K}[x, y]$. Setting $A'' = A' + cg$ and $B'' = B' - cf$, we have $A''_0 = B''_0 = 0$, so that $A'' = zA$ and $B'' = zB$ for some forms A and B. Since $zh' = A' f + B' g = A'' f + B'' g = z(Af + Bg)$, we conclude that $h' = Af + Bg$, as desired. □

Remark 5.5.2. Nowadays, Max Noether's theorem is usually not treated in textbooks on algebraic curves since it can be easily deduced from the cohomological vanishing result

$$\mathrm{H}^1(\mathbb{P}^2, \mathcal{O}(h-d-e)) = 0.$$

In this first course on algebraic curves, we will not develop the machinery of sheaves and cohomology. In a second course, Max Noether's theorem may serve as a motivation for the interest in vanishing theorems. □

Corollary 5.5.3. *Let* $f, h \in \Bbbk[x,y,z]$ *be forms of degrees* ≥ 1 *which intersect in* $\deg f \cdot \deg h$ *distinct points. Let* $g \in \Bbbk[x,y,z]$ *be a form of degree* ≥ 1 *passing through* $\deg f \cdot \deg g$ *of these points. Then there is a form of degree* $h - e$ *in x,y,z passing through the residual* $\deg f \cdot (\deg h - \deg g)$ *points.*

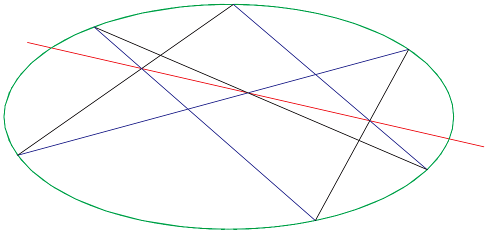
Proof. The conditions $h_p \in \langle g_p, f_p \rangle$ are satisfied, because g and f intersect transversally by Bézout's Theorem. Thus

$$h = af + bg$$

by Noether's Theorem. The polynomial b defines the curve of degree $h - e$, which contains the remaining $d \cdot (h-e)$ intersection points. □

A special case of the Corollary is Pascal's Theorem.

Example 5.5.4 (Pascal's Theorem). Consider a hexagon with vertices $p_1, \ldots, p_6 \in \mathbb{P}^2$ and the three intersection points $q_1 = \overline{p_1p_2} \cap \overline{p_4p_5}, q_2 = \overline{p_2p_3} \cap \overline{p_5p_6}, q_3 = \overline{p_3p_4} \cap \overline{p_6p_1}$ of the opposite lines. Then $p_1, \ldots, p_6$ lie on a conic iff q_1, q_2, q_3 lie on a line.



To prove this, we consider the cubic curves $C = \overline{p_1p_2} \cup \overline{p_3p_4} \cup \overline{p_5p_6}$ and $H = \overline{p_2p_3} \cup \overline{p_4p_5} \cup \overline{p_6p_1}$, which intersect in $\{p_1, \ldots, p_6\} \cup \{q_1, q_2, q_3\}$.
The statement for hexagons with vertices on a reducible conic is known as Pappus' Theorem.

5.6 Cubic Curves

Let $C = \mathrm{V}(f) \subset \mathbb{P}^2$ be an absolutely irreducible cubic. Given two points $p, q \in C$, we can construct another point on C as the third intersection point of the line $\overline{p,q}$ with C. We denote this point momentarily by $\neg(p \vee q)$. Similarly for a single smooth point $p \in C$, the third intersection point of the projective tangent line $T_pC \subset \mathbb{P}^2$ with C gives another point, momentarily called $\neg(p \vee p)$. With this secant-tangent construction, we can give C the structure of an abelian group as follows:

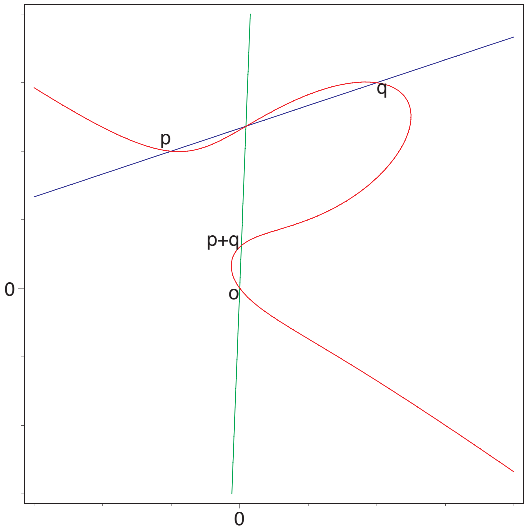
Fix a smooth point $o \in C$, which will serve as the "zero" element in the group. The group law is defined as

$$p + q := \neg((\neg(p \vee q)) \vee o),$$

that is the third intersection point of C with the line $\overline{\neg(p \vee q), o}$. We illustrate the group law on the curve given by the affine equation

$$11x^3 - 4xy^2 + 23y^3 - 6x^2 - 32xy - 67y^2 + 43x + 32y = 0$$

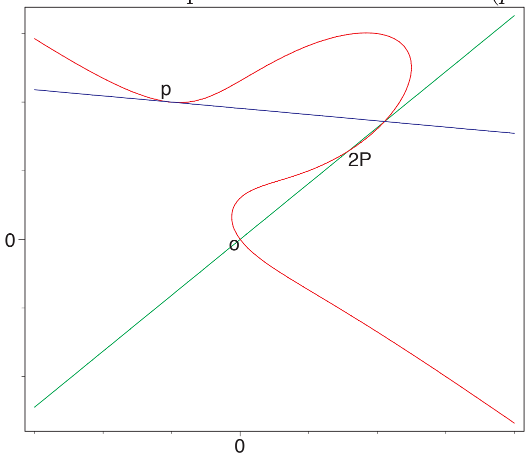
and points o, p, q with affine integral coordinates.



Similarly, replacing the secant by the tangent, we define

$$2p := \neg((\neg(p \vee p)) \vee o),$$

that is the third intersection point of C with the line $\overline{\neg(p \vee p), o}$.

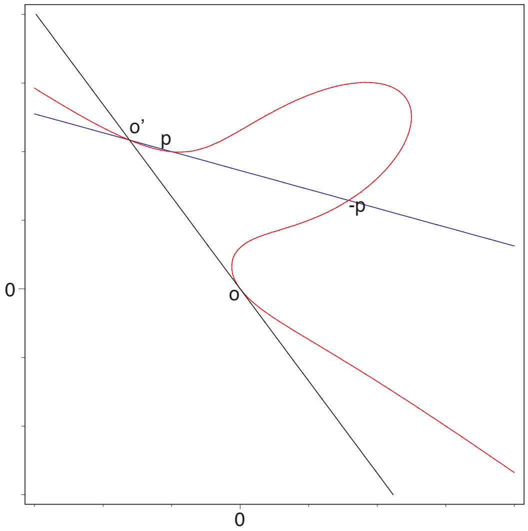


Theorem 5.6.1. *Let $\mathbb{k}$ be a not necessarily algebraically closed field. Let C be an absolutely irreducible cubic, let $C^0 = C \setminus \operatorname{Sing} C$ denote the set of non-singular points, and let $o \in C^0$ be a fixed point. The binary operation*

$$C^0 \times C^0 \to C^0,\ (p, q) \mapsto p + q$$

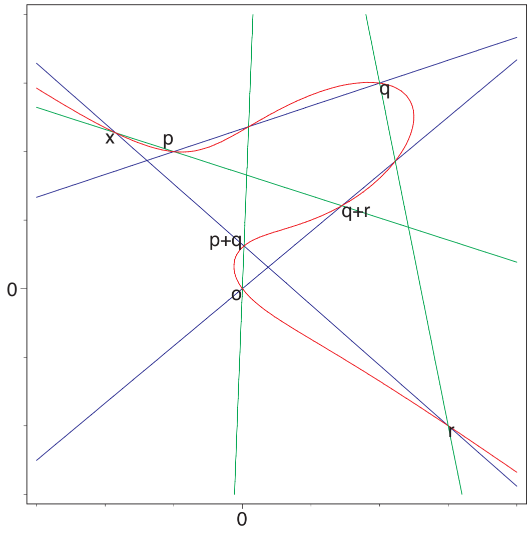
defined above, gives C^0 the structure of an abelian group with $o \in C$ as zero element. If $o \in C(\mathbb{k})$ then $C^0(\mathbb{k}) \subset C^0$ forms a subgroup.

Proof. All is clear except the associativity law. For example, the statement about the subgroup follows, because a cubic polynomial in one variable with two $\mathbb{k}$-rational roots has all three roots $\mathbb{k}$-rational. For the negative of a point $p \in C^0$, we consider the third intersection o' of C with T_oC. Then $-p$ is the third intersection of C with $\overline{o', p}$.



Note that we get only smooth points, because a secant or tangent through smooth points cannot intersect C in a singular point by Bézout's Theorem and Proposition 4.3.18.3.

To prove $(p + q) + r = p + (q + r)$, we consider all lines involved in the construction. We have to prove that the lines $\overline{p + q, r}$ and $\overline{p, q + r}$ intersect in a point of C.



For this we consider the cubics C and $\overline{p,q} \cup \overline{p+q,r} \cup \overline{o,q+r}$, which intersect in the nine points

$$p, q, \neg(p \vee q), p+q, r, x = \neg((p+q) \vee r), o, q+r, \text{ and } \neg(q \vee r),$$

which we assume to be different. Out of these, the following six $q, r, \neg(q \vee r), o, p+q, \neg(p \vee q)$ lie on the quadric $\overline{q,r} \cup \overline{o,q+q}$. Thus, the remaining three $p, q+r, \neg((p+q) \vee r)$ lie on the line $\overline{p,q+r}$ by Corollary 5.5.3, and the points

$$\neg((p+q) \vee r) \text{ and } \neg(p \vee (q+r))$$

coincide. To prove $2p + q = p + (p+q)$ or other cases, where some of the points coincide, we argue with continuity. So far, we have proved that $\neg((p+q) \vee r) = \neg(p \vee (q+r))$ holds for an non-empty Zariski open subset of triples $(p,q,r) \in C^0 \times C^0 \times C^0$. We will define the Zariski toplogy on $C \times C \times C$ in Chapter 6 precisely. It is clear, that iff some of the points in the construction come together, some secant lines might become tangent lines, and that some lines might coincide as well. The condition $\neg((p+q) \vee r) = \neg(p \vee (q+r))$ is an algebraic condition on the irreducible algebraic set $C^0 \times C^0 \times C^0$. Since it holds on a non-empty Zariski open subset, it holds everywhere. □

The negative in the group law becomes particularly simple, if we can choose a flex as the origin o. In that case o and o' coincide, and $-p$ is the third intersection of $\overline{o,p}$ with C.

Definition 5.6.2. *Let $p \in C \subset \mathbb{P}^2$ be a smooth point on a curve. The point $p \in C$ is a* **flex** *of C, if $i(T_pC, C; p) \geq 3$. The multiplicity of the flex is $i(T_pC, C; p) - 2$.*

Thus, every point on a line is a flex. A smooth conic has no flexes at all by Bézout's Theorem.

To determine the flexes of a curve $C(f)$ defined by a square-free polynomial $f \in \Bbbk[x,y,z]$, we consider the **Hessian** and the Hessian matrix. We abbreviate $f_x = \frac{\partial f}{\partial x}$ and so on. Then

$$Hess(f) = \det \begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{xy} & f_{yy} & f_{yz} \\ f_{xz} & f_{yz} & f_{zz} \end{pmatrix}.$$

Note, that the Hessian curve $H = \mathrm{V}(Hess(f))$ is independent of the choice of the coordinate system, because a change coordinates $(x,y,z)^t = A(u,v,w)^t$ amounts to the multiplication with the matrices A^t and A.

Proposition 5.6.3. *Assume* $\operatorname{char} \Bbbk = 0$ *and that* $f \in \Bbbk[x,y,z]$ *is square-free. Then* $C = \mathrm{V}(f)$ *and* $H = \mathrm{V}(Hess(f))$ *intersect in the singular points of* C *and in the flexes. More precisely,*

$$i(C, T_pC; p) - 2 = i(C, H; p)$$

for smooth points of $p \in C$.

Proof. We may assume that $d = \deg C \geq 2$. That H and C intersect in singular points of C follows with the **Euler relation:**

$$xg_x + yg_y + zg_z = \deg g \cdot g,$$

for g homogeneous. Thus

$$(d-1)\begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix} = \begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{xy} & f_{yy} & f_{yz} \\ f_{xz} & f_{yz} & f_{zz} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Since f_x, f_y, f_z vanish at singular points of C, we conclude that at a singular point $[\alpha : \beta : \gamma]$ of C the Hessian matrix has a nonzero kernel, hence determinant zero. For a smooth point $p \in C$, we consider appropriate coordinates. Suppose $p = [0:0:1]$ and $T_pC = \mathrm{V}(y)$. Then the affine equation of C is of the form

$$f(x,y,1) = yu(x,y) + x^k g(x) \textbf{ with } k \geq 2$$

with $u(0,0) \neq 0$, $g(0) \neq 0$. Homogenization gives

$$f(x,y,z) = yu(x,y,z) + x^k g(x,z)$$

for suitable homogeneous polynomials u, g with $u(0,0,1), u_z(0,0,1)$ and $g(0,1) \neq 0$. We evaluate the vanishing order $v_p(Hess(f))$ at p on C. Since $v_p(y) = k$ and $v_p(x) = 1$, we find that $v_p(f_{xx}) = v_p(yu_{xx} + k(k-1)x^{k-2}g + 2kx^{k-1}g_x + x^k g_{xx}) = k-2$ and $v_p(f_{yz}) = v_p(u_z) = 0$. It follows that

$$i(C,H;p) = v_p(Hess(f)) = v_p(-f_{xx}f_{yz}^2) = k-2 = i(C,T_pC;p)-2,$$

since all other terms in the Laplace expansion of the Hessian have higher vanishing order. □

Corollary 5.6.4. *Let* $\operatorname{char}\mathbb{k} = 0$*. A smooth curve* C *of degree* d *has* $3d(d-2)$ *flexes counted with multiplicity.*

Proof. The degree of the Hessian is 3(d-2). □

Exercise 5.6.5. Suppose $\operatorname{char}\mathbb{k} = 0$. Let $C \subset \mathbb{P}^2$ be a curve with singularities. Prove:

1. $i(C, Hess(C), p) = 6$, for $p \in C$ an ordinary node,
2. $i(C, Hess(C), p) = 8$, for $p \in C$ an ordinary cusp.

Conclude that a curve with δ ordinary nodes and κ ordinary cusps as its only singularities has

$$f = 3d(d-2) - 6\delta - 8\kappa$$

flexes counted with multiplicities. □

A smooth cubic curve can have only simple flexes by Bézout. Analysing in the proof the assumption $\operatorname{char}\mathbb{k} = 0$, we find for cubic curves

Corollary 5.6.6. *If* $\operatorname{char}\mathbb{k} \neq 2,3$ *then a smooth cubic curve has precisely 9 flexes.*

Corollary 5.6.7. *Suppose that* $\operatorname{char}\mathbb{k} \neq 2,3$*. Then, after a change of coordinates, any smooth cubic curve* $C \subset \mathbb{P}^2$ *can be defined by an equation*

$$y^2z = x^3 + axz^2 + bz^3$$

with coefficients a,b*. Conversely, the cubic defined by such an equation is smooth iff the disriminant* $27a^3 + 4b^2 \neq 0$*.*

Proof. We may change coordinates such, that $o = [0:1:0]$ is a flex, and that $T_oC = \mathrm{V}(z)$. Then the affine equation of C has shape

$$a_0'y^2 + a_1'xy + a_2' = x^3 + b_1'x^2 + b_2'x^2 + b_3'$$

Taking a_0' into z, we arrive at

$$y^2 + a_1xy + a_2y = x^3 + b_1x^2 + b_2x + b_3.$$

Finally, substituting first $y = y - a_1/2x - a_2/2$ and then $x = x - b_1''/3$, we arrive at

$$y^2 = x^3 + ax + b.$$

The curve defined by such an equation is singular iff x^3+ax+b has a multiple root iff $27a^3 + 4b^2 = 0$. Note that this change of coordinates is defined over the ground field iff the flex is a $\mathbb{k}$-rational point. □

Definition 5.6.8. *An* **elliptic curve in Weierstrass normal form** *is a smooth cubic curve E defined by an affine* **Weierstrass equation**

$$y^2 + a_1xy + a_2y = x^3 + b_1x^2 + b_2x + b_3.$$

The curve E carries a group structure with the single intersection point $o = [0:1:0]$ of E and the line at infinity as Null in the group. If $\operatorname{char} \Bbbk \neq 2, 3$ *then the equation can be simplified to*

$$y^2 = x^3 + ax + b.$$

The main difference between a smooth cubic and an elliptic curve is, that an elliptic curve has a $\Bbbk$-rational point over its field of definition. We will see in Chapter 8, that indeed any smooth cubic curve C with a $\Bbbk$-rational point is isomorphic to a cubic in Weierstrass normal form. However, in general the isomorphism is not induced by a linear automorphism of $\mathbb{P}^2$.

Exercise 5.6.9. Suppose that $\operatorname{char} \Bbbk \neq 2, 3$. Prove that the secant line through two flexes of an irreducible cubic curve intersects the curve in a further flex. *Hint:* Choose one of the flexes as the Null in the group, and consider the 3-torsion elements of the group. □

Exercise 5.6.10. Proove that the incidence correspondence between flexes and secant lines joining them, coincides with the incidence configuration of $\mathbb{F}_3$-rational points and lines in $\mathbb{A}^2(\mathbb{F}_3)$. □

Exercise 5.6.11. Prove that for an irreducible cubic defined over $\mathbb{R}$, at most three of the flexes can be real. □

Exercise 5.6.12. Let $C = \mathrm{V}(f)$ be a cubic defined by an affine Weierstrass equation

$$y^2 = x^3 + ax + b.$$

Choose as Null the single intersection point $o = [0:1:0]$ of C with the line at infinity. Prove the following formulas for the group law on C^0:

1. $-(x, y) = (x, -y)$
2. $(x_1, y_1) + (x_2, y_2) = (x_3, y_3)$ with

$$x_3 = (\frac{y_2 - y_1}{x_2 - x_1})^2 - x_1 - x_2 \text{ and } y_3 = (\frac{y_2 - y_1}{x_2 - x_1})(x_3 - x_1) + y_1$$

3. $2(x_1, y_1) = (x_3, y_3)$ with

$$x_3 = (\frac{3x_1 + a}{2y_1})^2 - 2x_1 \text{ and } y_3 = (\frac{3x_1 + a}{2y_1})(x_3 - x_1) + y_1$$

□

Exercise 5.6.13. Let C be the projective closure of $\mathrm{V}(y^2 - x^3)$ and the Null $o \in C$ as in Exercise 5.6.12. Prove that

$$C^0(\Bbbk) \cong (\Bbbk, +).$$

Let C be the projective closure of $\mathrm{V}(y^2 - x^3 - x^2)$ and the Null $o \in C$ as in Exercise 5.6.12. Prove that

$$C^0(\Bbbk) \cong (\Bbbk^*, *).$$

□

Remark 5.6.14. Elliptic curves E defined over the finite field $\mathbb{F}_q$ with q elements recently found applications in cryptography, see Koblitz [1994]. Choosing an elliptic curve over $\mathbb{F}_q$ at random, is like choosing a random abelian group of size $\approx q + 1$ by the famous Hasse-Weil Theorem. Let $\sharp E(\mathbb{F}_q)$ denote the number of $\mathbb{F}_q$-rational points.

Theorem 5.6.15 (Hasse-Weil Theorem). *Let E be a smooth elliptic curve defined over $\mathbb{F}_q$. Then the number of $\mathbb{F}_q$-rational points is estimated by*

$$|\sharp E(\mathbb{F}_q) - q - 1| \leq 2\sqrt{q}.$$

We will prove a more general formula for arbitrary smooth curves in Chapter 8, Theorem 8.8.24. A plausibility argument runs as follows: $E(\mathbb{F}_q)$ contains the point o at infinity. All other points project onto a point of $\mathbb{A}^1(\mathbb{F}_q) \subset \mathbb{P}^1(\mathbb{F}_q)$. There are q points in $\mathbb{A}^1(\mathbb{F}_q)$. Over the possible three roots α of $x^3 + ax + b$ in $\mathbb{A}^1(\mathbb{F}_q)$ we have precisely one point $[\alpha : 0 : 1]$ in $E(\mathbb{F}_q)$. Over the other points $\alpha \in \mathbb{A}^1(\mathbb{F}_q)$, we find either two or no point depending on whether $\alpha^3 + a\alpha + b \in (\mathbb{F}_q^*)^2$ or not. If we assume that the map

$$D(x^3 + ax + b)(\mathbb{F}_q) \to \mathbb{F}_q^*/(\mathbb{F}_q^*)^2,\ \alpha \mapsto \alpha^3 + a\alpha + b$$

behaves like a random function then we can model $\sharp E(\mathbb{F}_q) - q - 1$ with a random path with steps ± 1 of length q. Then the expectation value of $\sharp E(\mathbb{F}_q)$ is $q + 1$ and the expectation value of $|\sharp E(\mathbb{F}_q) - q - 1|$ is $\approx \sqrt{q}$.

A much more precise statemant about the distribution of the orders $\sharp E(\mathbb{F}_q)$ of elliptic curves over $\mathbb{F}_q$, when E runs through the finite set of elliptic curves over $\mathbb{F}_q$, can be found in [Gekeler, 2003].

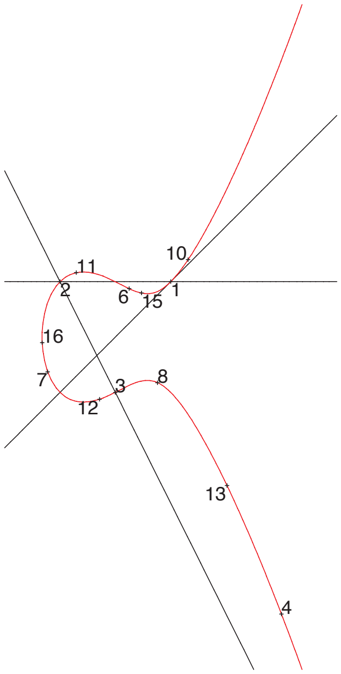
A different application of elliptic curves over finite fields due to Lenstra and Lenstra concerns integer factorization and primality tests, see **?** and **?**.

Elliptic curves over number fields are an intense area of current research. To start, we have Mordell's Theorem:

Theorem 5.6.16 (Mordell, 1922). *Let E be an elliptic curve defined over $\mathbb{Q}$. Then $E(\mathbb{Q})$ is a finitely generated group.*

Thus, every $\mathbb{Q}$-rational point on E can be constructed via the tangent-secant construction from finitely many points. For a proof we refer to Silverman [1986].

Example 5.6.17. The point $p = (1, 1)$ on the elliptic curve E defined by $y^2 = x^3 - x + 1$ generates an infinite subgroup of $E(\mathbb{Q})$.



The torsion part of $E(\mathbb{Q})$ was clarified by the celebrated Theorem of Mazur.

Theorem 5.6.18 (Mazur, 1976). *Let E be an elliptic curve over $\mathbb{Q}$. Then $E(\mathbb{Q})_{tors}$ is one of the following groups*

$$\mathbb{Z}/n \text{ with } 1 \leq n \leq 10 \text{ or } n = 12$$

or

$$\mathbb{Z}/2 \times \mathbb{Z}/2n \text{ with } 1 \leq n \leq 4.$$

On the other hand, the rank of $E(\mathbb{Q})$ is the topic of one of most famous conjectures in Mathematics. Let $E(\mathbb{Q})$ be a smooth elliptic curve with defining equation in $\mathbb{Z}[x, y]$. Then for almost all prime numbers p, we obtain a smooth cubic curve $E \mod p$ over the finite field $\mathbb{F}_p = \mathbb{Z}/p$. Write its number of points in the form

$$E(\mathbb{F}_p) = 1 - a_p + p.$$

A more precise version of the Hasse-Weil Theorem (Theorem 8.8.25) says that the reciprocal roots $\alpha, \overline{\alpha}$ of

$$1 - a_p t + pt^2 = (1 - \alpha t)(1 - \overline{\alpha} t)$$

are integral algebraic numbers of absolute value $|\alpha| = \sqrt{p}$.

We collect the local information of $E \mod \mathbb{F}_p$ with an Euler product to an analytic function: The **Hasse-Weil L-function** of E is defined by

$$L(E/\mathbb{Q}; s) = \frac{\zeta(s)\zeta(1-s)}{\prod_p (1 - a_p p^{-s} + p^{1-2s})}$$

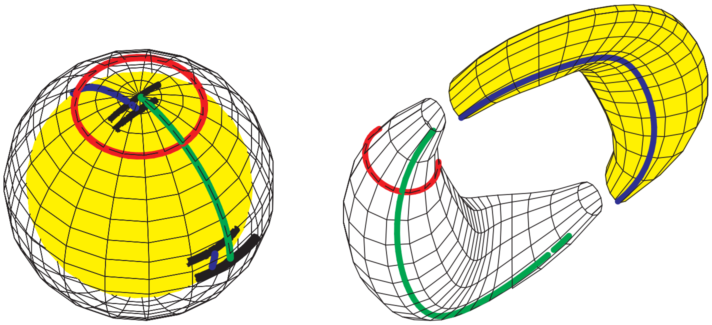
where $\zeta(s) = \prod_p (1 - p^{-s})^{-1} = \sum_n n^{-s}$ denotes the Riemann zeta function. The product of the denominators of $L(E/\mathbb{Q}, s)$ converges to an holomorphic function of s for s with real part $Re\, s > 1$. As the Riemann zeta function, the function $L(E/\mathbb{Q}, s)$ should have an analytic continuation.

Conjecture 5.6.19 (Birch and Swinnerton-Dyer, 1963, 1965). The Hasse-Weil L-function has an analytic continuation to the whole complex plane, and $\operatorname{rank} E(\mathbb{Q})$ equals the vanishing order of $L(E/\mathbb{Q}, s)$ at the critical point $s = 1$.

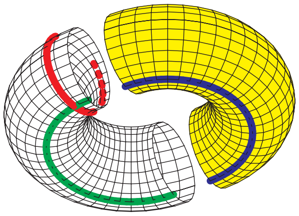
They also conjecture a precise statement about the leading coefficient. For reading on this fascinating topic in number theory we recommend Silverman [1986] or Husemöller [1986].

We now turn to the complex analytic side of the story about elliptic curves. One way to think about the elliptic curve $E \subset \mathbb{P}^2(\mathbb{C})$ defined by $y^2 = x^3 + ax + b$ over $\mathbb{C}$ is as the Riemann surface of the 2-valued analytic function $\sqrt{x^3 + ax + a}$. This amount to study E via the projection from $o = [0 : 1 : 0]$. The image of o will be the point at infinity $\infty = [1 : 0] \in \mathrm{V}(y) \cong \mathbb{P}^1$. Moreover, this is a ramification point, because o is a flex. The other ramification points lie on the line $\mathrm{V}(y)$. The three roots ρ_1, ρ_2, ρ_3 of $x^3 + ax + a$ give us three further ramification points $p_j = [\rho_j : 0 : 1] \in E$.

We can make $\sqrt{x^3 + ax + a}$ to a single valued function, if we restrict the domain of definition appropriately. Consider the line segment S_1 joining p_1 and p_2 and a half-line S_2, disjoint from S_1, which connects p_3 with ∞ on the Riemann number sphere $\mathbb{P}^1(\mathbb{C})$. Then $\sqrt{x^3 + ax + a}$ is single valued on $\mathbb{P}^1(\mathbb{C}) \setminus (S_1 \cup S_2)$, and the Riemann surface E is obtained by gluing two copies of $\mathbb{P}^1(\mathbb{C}) \setminus (S_1 \cup S_2)$ crosswise along the cuts.



It is easier to understand the Euclidean topology globally, if we draw the spheres not in each other and deform them a little bit. Note that the angle of two arcs ending at one of the branch point get divided by 2. Thus the angle of 360^o of the cut gives an angle of 180^o and thus a smooth arc. We conclude that the Riemann surface E is homeomorphic to a torus.



The universal covering space $\widetilde{E}$ of E is $\mathbb{C}$ as Riemann surface and

$$E = \mathbb{C}/\Lambda,$$

where $\Lambda \subset \mathbb{C}$ is a lattice. We see the group structure on E very clearly from this: $(E, +)$ is the quotiont group of $(\mathbb{C}, +)$ by the subgroup $(\Lambda, +)$. To prove $\widetilde{E} \cong \mathbb{C}$, one considers the elliptic integral

$$\int \frac{dx}{\sqrt{x^3 + ax + b}}.$$

$$\omega = \frac{dx}{\sqrt{x^3 + ax + b}} = \frac{dx}{y} = \frac{2dy}{\sqrt{3x^2 + a}}$$

is a nowhere vanishing holomorphic 1-form, because $y = \sqrt{x^3 + ax + b}$ and $x^3 + ax + b$ has no multiple roots. Thus we can define the integral

$$\int_o^p \frac{dx}{\sqrt{x^3 + ax + b}}$$

by choosing an arbitrary path from o to p, and the result is well defined up to a period, that is the integral of ω along a closed path. The first homology group $H_1(E, \mathbb{Z})$ has as basis represented by the red and blue/green paths γ_1, γ_2 indicated above.

One can prove that the periods

$$\lambda_j = \int_{\gamma_j} \omega$$

are $\mathbb{R}$-linearly independent. Thus $\Lambda = \mathbb{Z}\lambda_1 \oplus \mathbb{Z}\lambda_2$ is a lattice, and integration defines an unramified holomorphic map

$$\int_o : E \to \mathbb{C}/\Lambda,\, p \mapsto \int_o^p \omega \mod \Lambda$$

The inverse is given by the Weierstraß $\wp$-function and its derivative. Recall from the theory of complex function in one variable that

$$\wp(z) = \frac{1}{z^2} + \sum_{\lambda \in \Lambda \setminus \{0\}} \frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}$$

defines a meromorphic function with poles of order 2 at the lattice points. Moreover, the $\wp$-function and its derivative

$$\wp'(z) = \sum_{\lambda \in \Lambda} \frac{1}{(z-\lambda)^3}$$

are double periodic and satisfy the functional equation

$$(\wp')^2 = 4\wp^3 + g_2\wp + g_3$$

with $g_2 = \frac{1}{60} \sum'_{\lambda \in \Lambda} \frac{1}{\lambda^4}$ and $g_3 = \frac{1}{140} \sum'_{\lambda \in \Lambda} \frac{1}{\lambda^6}$. The inverse of $\int_o : E \to \mathbb{C}/\Lambda$ is given by

$$\mathbb{C}/\Lambda \to E \subset \mathbb{P}^2,\ z \mapsto [\wp(z] : \wp'(z)/2 : 1]$$

In particular, we claim that $a = g_2/4$ and $b = g_3/4$ holds. We do not prove this fact, but refer to Silverman [1986] and Husemoeller [1986] for further reading.

Chapter 6

Projective Algebraic Sets and Morphisms

In this Chapter we study arbitrary subvarieties of $\mathbb{P}^n$. In the first section we develop the algebra geometry dictionary for the projective setting and settle the question, how to compute the projective closure of arbitrary algebraic sets.

The second section is devoted to the definitions of products and morphisms. The main result of this section is the fundamtental theorem of elimination theory, which says that the image of an algebraic set under a projective morphism is an algebraic set. As consequence we get that regular functions on absolutely irreducible algebraic varieties are constant.

In Section 6.4 we introduce the Hilbert polynomial, which allows to define the degree of algebraic sets of higher codimension. Using the Hilbert polynomial we prove another version of Bézout's Theorem for the intersection of projective varieties of arbitrary codimension with hypersurface

In Section 6.5 we prove the dimension bound for intersections and the semi-continuity of the fiber dimension in a projective morphism. Section 6.6 deals with Bertini's Theorem and projective duality. An appendix contains the monodromy arguments for the uniform position of a general hyperplane section of curves and the irreducible of general hyperplane sections of higher dimensional varieties over fields of characteristic zero.

6.1 The Projective Nullstellensatz

In this section, we will explain how to link algebraic sets to ideals in the projective case. Since projective algebraic sets are defined by *homogeneous* polynomials, the ideals under consideration will have *homogeneous* generators. The general context for such ideals is that of graded rings.

Definition 6.1.1. A **graded ring** is a ring S with a decomposition $S = \bigoplus_{d\geq 0} S_d$ as Abelian groups such that $S_d S_e \subset S_{d+e}$ for all d, e. A **homogeneous element** of S is an element f of some graded piece S_d, and d is then called the **degree** of f. If $f = f_0 + f_1 + f_2 + \dots$ is the unique decomposition of an element $f \in S$ into homogeneous summands f_i of degree i, the f_i are called the **homogeneous components** of f. A **homogeneous ideal** of S is an ideal generated by homogeneous elements. □

If $S = \bigoplus_{d\geq 0} S_d$ is a graded ring, then S_0 is a ring with $1 \in S_0$, and S is an S_0-algebra. Furthermore, $S_+ := \bigoplus_{d\geq 1} S_d$ is a homogeneous ideal. In the case where $S_0 = \Bbbk$ is a field, this ideal is maximal and contains all other homogenous ideals of S.

Proposition 6.1.2. *Let I be an ideal of a graded ring $S = \bigoplus_{d\geq 0} S_d$. Then the following are equivalent:*

1. *I is homogeneous.*
2. *For each $f \in I$, the homogeneous components of f are in I as well:*

$$I = \bigoplus_{d\geq 0} (I \cap S_d)$$

Proof. 1 $\Longrightarrow$ 2: Let $\{f^{(\lambda)}\}$ be a set of homogenous generators for I, with $d_\lambda := \deg f^{(\lambda)}$ for all λ. Moreover, let $f \in I$, and let $f_m \neq 0$ be the homogenous component of f of least degree. The result will follow by induction once we show that $f - f_m \in I$. For this, we write f as a sum $f = g^{(\lambda_1)} f^{(\lambda_1)} + \dots + g^{(\lambda_r)} f^{(\lambda_r)}$. Then, with the obvious notation, $f_m = g^{(\lambda_1)}_{m-d_{\lambda_1}} f^{(\lambda_1)} + \dots + g^{(\lambda_r)}_{m-d_{\lambda_r}} f^{(\lambda_r)} \in I$.

2 $\Longrightarrow$ 1: If condition 2 is satisfied, the homogeneous components of the elements of any given set of generators for I generate I, too. □

Exercise* 6.1.3. Let S be a graded ring.

1. Show that the sum, product, intersection, ideal quotient, and radical of homogeneous ideals are homogeneous.
2. Show that a homogeneous ideal $\mathfrak{p} \subset S$ is prime iff for any two *homogeneous* elements $f, g \in S$ with $fg \in \mathfrak{p}$ we must have $f \in \mathfrak{p}$ or $g \in \mathfrak{p}$. □

It is clear from the proof of Proposition 6.1.2 that every homogeneous ideal of a *Noetherian* graded ring is generated by *finitely many* homogeneous elements. The polynomial ring $\Bbbk[x_0, \dots, x_n]$ with its natural grading by the degree of polynomials is our basic example of a Noetherian graded ring.

Exercise* 6.1.4 (Characterization of Noetherian Graded Rings). Let $S = \bigoplus_{d \geq 0} S_d$ be a graded ring. Show that the following are equivalent:

1. S is Noetherian.
2. S_0 is Noetherian and S_+ is a finitely generated ideal.
3. S_0 is Noetherian and S is a finitely generated S_0-algebra. □

At this point, setting up an I–V-correspondence between algebraic subsets of $\mathbb{P}^n(\Bbbk)$ and homogeneous ideals of $\Bbbk[x_0. \ldots, x_n]$, the reader will have no difficulty in verifying results analogous to those proved in Chapter 1. In particular, each algebraic set $A \subset \mathbb{P}^n(\Bbbk)$ is defined by finitely many homogeneous polynomials; it has finitely many **irreducible components**; and, it is irreducible iff $\mathrm{I}(A)$ is a prime ideal. Moreover, the Zariski closure of the difference of two projective algebraic sets is obtained as in Theorem 1.9.1. For the sake of brevity, we will only treat the projective version of the Nullstellensatz in some detail. In doing so, we will use I and V in accordance with Convention 5.2.1:

Definition 6.1.5. 1. If $I \subset \Bbbk[x_0, \ldots, x_n]$ is a homogeneous ideal, its **locus of zeros** (or **vanishing locus**) in $\mathbb{P}^n$ is the projective algebraic set

$$\mathrm{V}(I) := \{p \in \mathbb{P}^n \mid f(p) = 0 \text{ for all homogeneous } f \in I\}.$$

2. Let $S := \mathbb{K}[x_0, \ldots, x_n]$. If $A \subset \mathbb{P}^n$ is any subset, its **vanishing ideal** is the homogeneous ideal

$$\mathrm{I}(A) := \langle f \in S \mid f \text{ is homogeneous and } f(p) = 0 \text{ for all } p \in A\rangle.$$

□

Remark 6.1.6. Note that

$$\mathrm{I}(A) = \{f \in S \mid f(a_0, \ldots, a_n) = 0 \text{ for any } p \in A \text{ and any set } a_0, \ldots, a_n \text{ of homogeneous coordinates for } p\}.$$

Indeed, if $f = f_m + \ldots + f_d$ is an element of the ideal on the right hand side, where the f_i are homogenous of degree i, and $p = [a_0 : \cdots : a_n] \in A$, then

$$0 = f(\lambda a_0, \ldots, \lambda a_n) = \lambda^m f_m(a_0, \ldots, a_n) + \cdots + \lambda^d f_d(a_0, \ldots, a_n)$$

for all $\lambda \in \mathbb{K}$. Since $\mathbb{K}$ is infinite, this is only possible iff $f_i(a_0, \ldots, a_n) = 0$ for all i. It follows that $f \in \mathrm{I}(A)$. The reverse inclusion is clear. □

Theorem 6.1.7 (Projective Nullstellensatz). *Let $I \subset \Bbbk[x_0, \ldots, x_n]$ be a homogeneous ideal. Then:*

1. $\mathrm{V}(I) = \emptyset \iff I \supset \langle x_0, \ldots, x_n\rangle^d$ *for some* $d \geq 1$.
2. If $\mathrm{V}(I)$ *is nonempty, then*

$$\mathrm{I}(\mathrm{V}(I)) = \mathrm{rad}\,(I\,\mathbb{K}[x_0, \ldots, x_n]).$$

Proof. The theorem follows by applying the affine version of the Nullstellensatz to the affine cone $C(\mathrm{V}(I))$:

1. We have

$$\mathrm{V}(I) = \emptyset \iff C(\mathrm{V}(I)) \subset \{0\} \iff \mathrm{rad}\,(I) \supset \langle x_0, \dots, x_n \rangle.$$

2. If $\mathrm{V}(I)$ is nonempty, we have

$$f \in \mathrm{I}(\mathrm{V}(I)) \iff f \in \mathrm{I}(C(\mathrm{V}(I))) \iff f \in \mathrm{rad}\,(I\,\mathbb{K}[x_0, \dots, x_n]).$$

□

Corollary 6.1.8. *There is an inclusion-reversing one-to-one correspondence*

$$\{\textit{algebraic subsets of}\ \mathbb{P}^n\} \underset{\mathrm{V}}{\overset{\mathrm{I}}{\rightleftarrows}} \left\{\begin{matrix}\textit{homogeneous radical ideals}\\ \textit{of}\ \mathbb{K}[x_0, \dots, x_n]\\ \textit{not equal to}\ \langle x_0, \dots, x_n\rangle\end{matrix}\right\}.$$

Under this correspondence, subvarieties of $\mathbb{P}^n$ *correspond to homogeneous prime ideals of* $\mathbb{K}[x_0, \dots, x_n]$ *not equal to* $\langle x_0, \dots, x_n \rangle$. □

Since the ideal $\mathbb{K}[x_0, \dots, x_n]_+ = \langle x_0, \dots, x_n \rangle$ is missing in this correspondence, it is often called the **irrelevant ideal**.

Definition 6.1.9. The **homogeneous coordinate ring** of an algebraic set $A \subset \mathbb{P}^n$ is the quotient ring

$$\mathbb{K}[A] = \mathbb{K}[x_0, \dots, x_n]/\mathrm{I}(A).$$

□

In terms of affine algebraic sets, $\mathbb{K}[A]$ is the coordinate ring of the affine cone $C(A) \subset \mathbb{A}^{n+1}$. Note that $\mathbb{K}[A]$ has a natural grading. In fact, if $S = \bigoplus_{d \geq 0} S_d$ is any graded ring, and $I = \bigoplus_{d \geq 0}(I \cap S_d)$ is any homogeneous ideal of S, then

$$S/I = \bigoplus_{d \geq 0} S_d/(I \cap S_d).$$

The relationship between algebraic subsets of A and homogeneous ideals of $\mathbb{K}[A]$ is analogous to Exercise 1.11.7.

Remark 6.1.10 (Buchberger's Algorithm and Homogeneous Ideals). With respect to computational aspects, we note that Buchberger's algorithm applied to homogeneous polynomials yields Gröbner basis elements which are homogeneous, too. In particular, given any global monomial order on $S = \Bbbk[x_0, \dots, x_n]$, the elements of the reduced Gröbner basis for a homogeneous ideal I of S are homogeneous. Hence, the computational recipes given in Chapter 2 are valid in the projective case as well. □

We finish this section by defining the dimension of a projective algebraic set. One way of doing this is to extend the affine notion of dimension via coordinate charts (alternative ways will be discussed in subsequent sections):

Definition 6.1.11. The **dimension** of an algebraic subset $A \subset \mathbb{P}^n$, written $\dim A$, is defined to be the number

$$\dim A = \max\{A \cap U_i \mid i = 0, \dots, n\}. \qquad \square$$

We will use the words **codimension**, **equidimensional**, **curve**, and **surface** exactly as in the affine case. It follows from that case that $\dim A$ is the maximum dimension of the irreducible components of A, and that A is a hypersurface iff it is equidimensional of dimension $n-1$. In algebraic terms, we will see in Corollary 6.4.19 that if $A \subset \mathbb{P}^n$ is any projective algebraic set, then

$$\dim A = \dim C(A) - 1 = \dim \mathbb{K}[A] - 1.$$

6.2 Computing the Projective Closure

To describe the projective closure of an affine algebraic set in algebraic terms, we introduce the following notation: The **homogenization** of an ideal $I \subset \Bbbk[x_1, \dots, x_n]$ with respect to an extra variable x_0 is the ideal

$$I^h = \langle f^h \mid f \in I \rangle \subset \Bbbk[x_0, \dots, x_n].$$

Theorem 6.2.1. *Let $I \subset \Bbbk[x_1, \dots, x_n]$ be an ideal, and let I^h be its homogenization with respect to x_0. Then $\mathrm{V}(I^h) \subset \mathbb{P}^n$ is the projective closure of the affine algebraic set $\mathrm{V}_a(I) \subset \mathbb{A}^n \cong U_0 \subset \mathbb{P}^n$.*

Proof. First, it is clear that $\mathrm{V}(I^h)$ is an algebraic subset of $\mathbb{P}^n$ which contains $\mathrm{V}_a(I)$. To show that $\mathrm{V}(I^h)$ is the smallest such set, let $B \subset \mathbb{P}^n$ be any algebraic set containing $\mathrm{V}_a(I)$, and let $F \in \mathrm{I}(B) \subset \mathbb{K}[x_0, \dots, x_n]$ be any form. Then the dehomogenization $f = F(1, x_1, \dots, x_n)$ is contained in $\mathrm{I}_a(\mathrm{V}_a(I))$ (with obvious notation). Hence, by the affine Nullstellensatz, $f^m \in \operatorname{rad}(I\,\mathbb{K}[x_1, \dots, x_n])$ for some m. This shows

$$(f^h)^m = (f^m)^h \in (I\,\mathbb{K}[x_1, \dots, x_n])^h = I^h\,\mathbb{K}[x_1, \dots, x_n] \subset \mathrm{I}(\mathrm{V}(I^h)).$$

Since $F = x_0^s f^h$ for some $s \geq 0$, it follows that $F \in \mathrm{I}(\mathrm{V}(I^h))$, as desired. $\square$

Exercise* 6.2.2. Let $A \subset \mathbb{A}^n \cong U_0$ be an affine algebraic set, and let $\overline{A}$ be its projective closure in $\mathbb{P}^n$. Show:

1. A is irreducible iff $\overline{A}$ is irreducible.
2. If $A = V_1 \cup \dots \cup V_r$ is the decomposition into irreducible components, then $\overline{A} = \overline{V}_1 \cup \dots \cup \overline{V}_r$ is the decomposition into irreducible components.

In particular, no irreducible component of $\overline{A}$ is contained in the hyperplane at infinity. $\square$

With respect to computing I^h, we note that the naive approach of just homogenizing the given generators for I may lead to the wrong ideal:

Example 6.2.3. Consider the ideal $I = \langle y - x^2, z - x^3 \rangle \subset \Bbbk[x, y, z]$, which defines the twisted cubic curve C in $\mathbb{A}^3$. Homogenizing the generators, we get the ideal $J = \langle wy - x^2, w^2 z - x^3 \rangle \subset \Bbbk[w, x, y, z]$, which decomposes as

$$J = \langle x^2 - wy, xy - wz, y^2 - xz \rangle \cap \langle x^2 - yw, xw, w^2 \rangle.$$

This shows that the line $\mathrm{V}(w, x)$, which is contained in the hyperplane at infinity, is an irreducible component of $\mathrm{V}(J) \subset \mathbb{P}^3$. Hence, J cannot be the homogenization of I (of course, this can also be seen directly by specifaying an element of I^h not contained in J). The projective closure of C, which is called the **twisted cubic curve in projective 3-space** $\mathbb{P}^3$, is defined by the ideal

$$J : \langle w, x \rangle = \langle x^2 - wy, xy - wz, y^2 - xz \rangle.$$

Note that the generators for this ideal are obtained by homogenizing the elements of the (reduced) Gröbner basis for I with respect to $>_{\mathrm{drlex}}$. □

In general, we have:

Proposition 6.2.4. *Let $I \subset \Bbbk[x_1, \dots, x_n]$ be an ideal. Pick a degree-compatible (global) monomial order $>$ on $\Bbbk[x_1, \dots, x_n]$, and set*

$$x^\alpha x_0^d >_h x^\beta x_0^e \iff x^\alpha > x^\beta \text{ or } (x^\alpha = x^\beta \text{ and } d > e).$$

Then $>_h$ is a global monomial order on $\Bbbk[x_0, \dots, x_n]$. Moreover, when homogenizing with respect to x_0, the following holds: If $f_1, \dots, f_r$ form a Gröbner basis for I with respect to $>$, then the homogenized polynomials $f_1^h, \dots, f_r^h$ form a Gröbner basis for the homogenized ideal I^h with respect to $>_h$.

Proof. That $>_h$ is a global monomial order is immediate from the definitions. For the second statement, note that if $f \in \Bbbk[x_1, \dots, x_n]$ is any nonzero polynomial, then $\deg \mathbf{L}_>(f) = \deg f$ since $>$ is degree-compatible. Hence, $\mathbf{L}_>(f)$ remains unchanged when we homogenize. According to how we defined $>_h$, it follows that $\mathbf{L}_{>_h}(f^h) = \mathbf{L}_>(f)$.

We use this to show that $\mathbf{L}(I^h) \subset \langle f_1^h, \dots, f_r^h \rangle$ (the reverse inclusion is clear). Let $F \in I^h$. Since I^h is a homogeneous ideal, any homogeneous component of F is contained in I^h, and we may suppose that F itself is homeogeneous. Writing F as a $\Bbbk[x_0, \dots, x_n]$-linear combination of polynomials g_j^h, with all $g_j \in I$, we find that the dehomogenization $f = F(1, x_1, \dots, x_n)$ is a $\Bbbk[x_1, \dots, x_n]$-linear combination of the g_j. In particular, $f \in I$. On the other hand, since F is homogeneous, we have $F = x_0^s f$ for some $s \geq 0$. Hence,

$$\mathbf{L}_{>h}(F) = x_0^s \cdot \mathbf{L}_{>h}(f^h) = x_0^s \cdot \mathbf{L}_>(f).$$

Since $\mathbf{L}_>(f)$ is a multiple of one of the $\mathbf{L}_>(f_i)$ by assumption, we conclude that $\mathbf{L}_{>h}(F)$ is a multiple of $\mathbf{L}_>(f_i) = \mathbf{L}_{>_h}(f_i^h)$, as required. □

Exercise 6.2.5. Let $d \geq 2$, and consider the image C of the parametrization

$$\mathbb{A}^1 \to \mathbb{A}^d,\ t \mapsto (t, t^2, \dots, t^d).$$

The projectice closure $\overline{C} \subset \mathbb{P}^d$ is known as the **rational normal curve** in $\mathbb{P}^d$. Note that for $d = 2, 3$, we get a nondegenerate conic respectively the twisted cubic curve. In general, show that $\mathrm{I}(\overline{C})$ is generated by $\binom{d}{2}$ quadrics, and that there is no set of generators with fewer elements. Note that for $d \geq 3$, the number of generators is strictly larger than the codimension $d - 1$. □

6.3 Products and Morphisms

We have seen in Exercise 1.11.5 that the product $A \times B$ of two affine algebraic sets $A \subset \mathbb{A}^n$ and $B \subset \mathbb{A}^m$ is an algebraic subset of $\mathbb{A}^n \times \mathbb{A}^m = \mathbb{A}^{n+m}$. In the projective setting, it is initially not even clear that $\mathbb{P}^n \times \mathbb{P}^m$ can be viewed as an algebraic set. There is, however, a natural way of doing this. The basic idea is to embed $\mathbb{P}^n \times \mathbb{P}^m$ in some $\mathbb{P}^N$ such that the image is a projective variety which locally, in the coordinate charts of $\mathbb{P}^N$, is isomorphic to the product $\mathbb{A}^n \times \mathbb{A}^m$. To make this precise, we note that sending $([a_0 : \dots : a_n], [b_0 : \dots : b_m])$ to $[a_0b_0 : \dots : a_0b_m : a_1b_0 : \dots : a_nb_m]$ gives a well-defined map

$$\sigma_{m,n} : \mathbb{P}^n \times \mathbb{P}^m \to \mathbb{P}^N, \quad \text{where} \quad N = (n+1)(m+1) - 1$$

(the map does not depend on the choice of homogeneous coordinates a_i, b_j, and at least one of the a_ib_j is nonzero). In studying $\sigma_{m,n}$, we denote the homogeneous coordinates on $\mathbb{P}^n$, $\mathbb{P}^m$, and $\mathbb{P}^N$ by $\boldsymbol{x} = x_0, \dots, x_n$, $\boldsymbol{y} = y_0, \dots, y_m$, and $\boldsymbol{z} = z_{00}, \dots, z_{0m}, z_{10}, \dots, z_{nm}$. Moreover, we say that a polynomial of type

$$f = \sum_{|\alpha|=d,\, |\beta|=e} c_\alpha x^\alpha y^\beta \in \Bbbk[\boldsymbol{x}, \boldsymbol{y}]$$

is **bihomogeneous (in $\boldsymbol{x}$ and $\boldsymbol{y}$, of bidegree $(\boldsymbol{d}, \boldsymbol{e})$)**.

Proposition 6.3.1. *The map $\sigma_{m,n}$ is injective, and its image $\Sigma_{m,n}$ is a subvariety of $\mathbb{P}^N$. The vanishing ideal $\mathrm{I}(\Sigma_{m,n})$ is generated by the 2×2 minors of the $(n+1) \times (m+1)$ matrix of coordinates (z_{ij}). In terms of coordinate charts, we have*

$$U_i \times U_j \cong \Sigma_{m,n} \cap U_{ij}.$$

Proof. It is clear that the minors vanish on $\Sigma_{m,n}$:

$$\det \begin{pmatrix} x_{i_1}y_{j_1} & x_{i_1}y_{j_2} \\ x_{i_2}y_{j_1} & x_{i_2}y_{j_2} \end{pmatrix} = 0. \tag{6.1}$$

Hence, if $A \subset \mathbb{P}^N$ denotes the algebraic set defined by the minors, then $\Sigma_{m,n} \subset A$. To show equality, we first intersect with the coordinate chart U_{00}. If $r =$

$[1 : c_{01} : \dots : c_{ij} : \dots] \in A \cap U_{00}$ is a point, then $c_{ij} = c_{i0}c_{0j}$. Hence, $([1 : c_{10} : \dots : c_{n0}], [1 : c_{01} : \dots : c_{0m}])$ is the unique pair of points $(p, q) \in U_0 \times U_0$ such that $\sigma_{m,n}\big((p, q)\big) = r$. We conclude that $\sigma_{m,n}$ restricts to an isomorphism $U_0 \times U_0 \cong A \cap U_{00}$ of affine varieties. Since the corresponding statement holds for the other coordinate charts, we have $\Sigma_{m,n} = A$, as desired. At the same time, the argument shows that $\sigma_{m,n}$ is injective.

The proposition will follow once we show that the ideal $I \subset \mathbb{K}[\boldsymbol{z}]$ generated by the minors is prime. For this, we show that I coincides with the kernel of the ring homomorphism

$$\phi : \mathbb{K}[\boldsymbol{z}] \to \mathbb{K}[\boldsymbol{x}, \boldsymbol{y}],\ z_{ij} \mapsto x_i y_j.$$

It is clear from (6.1) that $I \subset \ker \phi$. For the reverse inclusion, we use a counting argument which actually gives that the minors form a Gröbner basis for $\ker \phi$.

On $\mathbb{K}[\boldsymbol{z}]$, consider a global monomial order $>$ refining the partial order on the variables defined as follows:

$$\begin{array}{ccccccc} z_{00} & > & z_{01} & > & \dots & > & z_{0m} \\ \vee & & \vee & & & & \vee \\ z_{10} & > & z_{11} & > & \dots & > & z_{1m} \\ \vee & & \vee & & & & \vee \\ \vdots & & \vdots & & & & \vdots \\ \vee & & \vee & & & & \vee \\ z_{n0} & > & z_{n1} & > & \dots & > & z_{nm} \end{array}.$$

Then

$$\mathbf{L}(\det \begin{pmatrix} z_{i_1 j_1} & z_{i_1 j_2} \\ z_{i_2 j_1} & z_{i_2 j_2} \end{pmatrix}) = -z_{i_1 j_2} z_{i_2 j_1}$$

whenever $i_1 < i_2$ and $j_1 < j_2$. Hence, if $f \in \mathbb{K}[\boldsymbol{z}]$ is any polynomial, division with remainder yields a representation

$$f = g + h,$$

where g is a $\mathbb{K}[\boldsymbol{z}]$-linear combination of the minors, and such that h is a $\mathbb{K}$-linear combination of monomials of type

$$z_{i_1 j_1} z_{i_2 j_2} \cdot \ldots \cdot z_{i_d j_d}, \text{ where } i_1 \le i_2 \le \ldots \le i_d \text{ and } j_1 \le j_2 \le \ldots \le j_d.$$

Then $\phi(g) = 0$. Since ϕ restricts to a bijection between the set of ordered monomials of degree d as above and the set of bihomogeneos monomials in $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$ of bidegree (d, d), we conclude that $\phi(f) = 0$ iff $\phi(h) = 0$ iff $h = 0$. Hence, as claimed, the minors form a Gröbner basis for $\ker \phi$. In particular, $I = \ker \phi$. Moreover, I is a prime ideal since $\mathbb{K}[\boldsymbol{z}]/\ker \phi$ is isomorpic to a subring of the integral domain $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$. □

Being defined by by the 2×2 minors of the matrix (z_{ij}), the Segre variety $\Sigma_{m,n}$ is sometimes called an example of a **determinantal variety**.

Exercise 6.3.2. In the situation of the proof above, describe the syzygies on the 2×2 minors arising from Buchberger's test. □

Definition 6.3.3. The map $\sigma_{m,n}$ is called the **Segre embedding** of $\mathbb{P}^n \times \mathbb{P}^m$ into $\mathbb{P}^N$. Its image $\Sigma_{m,n}$ is called the **Segre variety**. We give $\mathbb{P}^n \times \mathbb{P}^m$ the **structure of a projective variety** by identifying it with $\Sigma_{m,n}$. □

Example 6.3.4. The Segre variety $\Sigma_{1,1}$ is the image of the map

$$\sigma_{1,1} : \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^3,\ ([a_0 : a_1], [b_0 : b_1]) \mapsto [a_0b_0 : a_0b_1 : a_1b_0 : a_1b_1].$$

It is a quadric defined by the equation $z_{00}z_{11} - z_{01}z_{10} = 0$. Note that the fibers of either projection of $\mathbb{P}^1 \times \mathbb{P}^1$ onto $\mathbb{P}^1$ form a pencil of lines on $\Sigma_{1,1}$ such two different lines in the same pencil do not meet, and such that two lines from different pencils intersect in one point.

□

Now that we have given $\mathbb{P}^n \times \mathbb{P}^m$ the structure of a projective algebraic set, we wish to describe its algebraic subsets. In terms of the Segre embedding, a subset $A \subset \mathbb{P}^n \times \mathbb{P}^m \cong \Sigma_{n,m} \subset \mathbb{P}^N$ is closed iff it is the vanishing locus of finitely many polynomials $f_k \in \mathbb{K}[\boldsymbol{z}]$, where each f_k is homogenous of some degree d_k. For a characterization just in terms of $\mathbb{P}^n \times \mathbb{P}^m$, substitute the x_iy_j for the z_{ij} in the f_k as in the proof of Proposition 6.3.1. The resulting polynomials are bihomogeneous in $\boldsymbol{x}$ and $\boldsymbol{y}$, of bidegrees (d_k, d_k), and their common vanishing locus in $\mathbb{P}^n \times \mathbb{P}^m$ is A. In fact, *every* bihomogeneous polynomial $f \in \mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$ has a well-defined vanishing locus $\mathrm{V}(f)$ in $\mathbb{P}^n \times \mathbb{P}^m$, and we have:

Proposition 6.3.5. *A subset of $\mathbb{P}^n \times \mathbb{P}^m$ is algebraic iff it is the common vanishing locus of finitely many bihomogeneous polynomials in $\boldsymbol{x}$ and $\boldsymbol{y}$.*

Proof. The implication from left to right is clear from the discusion above. For the converse implication, let $f \in \mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$ be any bihomogeneous polynomial of any bidegree (d, e). We show that $\mathrm{V}(f)$ is an algebraic subset of $\mathbb{P}^n \times \mathbb{P}^m$. This is obvious if $d = e$ since, then, we may rewrite f as a homogeneous polynomial in the x_iy_j and, thus, in the z_{ij}. If $d \neq e$, say $e < d$, we get $\binom{n+d-e}{n}$ bihomogeneous polynomials of bidegree (d, d) by multiplying f with each of the monomials in $\boldsymbol{y}$ of degree $d - e$. Since the common vanishing locus of these polynomials equals $\mathrm{V}(f)$, we are done. □

If $f \in \mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$ is a nonconstant polynomial of bidegree (d, e), then its vanishing locus $\mathrm{V}(f)$ in $\mathbb{P}^n \times \mathbb{P}^m$ is called a **hypersurface of bidegree (d, e)**.

Example 6.3.6. The equation $z_{00}z_{11} - z_{01}z_{10} = 0$ of the quadric $\Sigma_{1,1} \subset \mathbb{P}^3$ is one of the equations of the twisted cubic curve C in $\mathbb{P}^3$ which is, thus, contained in $\Sigma_{1,1}$. Taking the other two defining quadrics of C as in Example 6.2.3 and substituting, we get the bihomogeneous polynomials $x_0(x_0y_1^2 - x_1y_0^2)$ and $x_1(x_1y_0^2 - x_0y_1^2)$. Hence, $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ is defined by the single equation $x_0y_1^2 - x_1y_0^2 = 0$. It is a hypersurface of bidegree (1,2). □

Given algebraic subsets $A \subset \mathbb{P}^n$ and $B \subset \mathbb{P}^m$, it is, now, clear that the product $A \times B \subset \mathbb{P}^n \times \mathbb{P}^m$ is an algebraic subset as well: If $A = \mathrm{V}(f_1, \dots, f_r)$ and $B = \mathrm{V}(g_1, \dots, g_s)$, with homogeneous f_k and g_ℓ, then the f_k and g_ℓ considered as bihomogeneous polynomials in $\boldsymbol{x}$ and $\boldsymbol{y}$ of bidegrees $(\deg f_k, 0)$ and $(0, \deg g_\ell)$ define $A \times B$. We call

$$\mathrm{I}(A \times B) = \langle f \in \mathbb{K}[\boldsymbol{x}, \boldsymbol{y}] \text{ bihomogeneous } \mid f(p) = 0 \text{ for all } p \in A \times B \rangle$$

the **bihomogenous ideal** and $\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]/\mathrm{I}(A \times B)$ the **bihomogeneous coordinate ring** of $A \times B$.

Exercise* 6.3.7. In the situation above, show:

1. $\mathrm{I}(A \times B) = \big((\mathrm{I}(A)\,\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}] + \mathrm{I}(B)\,\mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]) : \langle \boldsymbol{x} \rangle^\infty\big) : \langle \boldsymbol{y} \rangle^\infty \subset \mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$.
2. The Zariski topology on $A \times B$ is not the product of the Zariski topologies on A and B, except when one of A and B is a finite set of points. □

Identifying $\mathbb{A}^m$ with the affine chart U_0 of $\mathbb{P}^m$, the product $\mathbb{P}^n \times \mathbb{A}^m$ inherits a Zariski topology from $\mathbb{P}^n \times \mathbb{P}^m$. With respect to this topology, a subset $A \subset \mathbb{P}^n \times \mathbb{A}^m$ is closed iff there are finitely many polynomials in $\mathbb{K}[x_0, \dots, x_n, y_1, \dots, y_m]$ which are homogeneous in $x_0, \dots, x_n$, and such that their common vanishing locus is A. Here, any polynomial of type

$$f = \sum_{|\alpha| = d} x^\alpha h_\alpha(y_1, \dots, y_m) \in \mathbb{K}[x_0, \dots, x_n, y_1, \dots, y_m],$$

with polynomials $h_\alpha(y_1, \dots, y_m) \in \mathbb{K}[y_1, \dots, y_m]$, is called **homogeneous in $\boldsymbol{x_0, \dots, x_n}$ (of degree $\boldsymbol{d}$)**. Note that every such polynomial f has a well-defined vanishing locus $\mathrm{V}(f)$ in $\mathbb{P}^n \times \mathbb{A}^m$.

Our next objective is to define morphisms between projective algebraic sets. Among the maps introduced so far in the projective setting are the coordinate maps $\varphi_i : U_i \to \mathbb{A}^n$, the canonical projection $\mathbb{A}^{n+1} \setminus \{o\} \to \mathbb{P}^n$, and the projection maps $\mathbb{P}^n \setminus \{p\} \to \mathbb{P}^{n-1}$. To include these and other natural maps in our treatment of morphisms, we work with a class of sets which embraces the affine and projective algebraic sets, and all open subsets of these.

Definition 6.3.8. An open subset of an affine algebraic set is called a **quasi-affine algebraic set**. Similarly, we have the notion of a **quasi-projective algebraic set**. □

Remark 6.3.9. The product of two quasi-affine (quasi-projective) algebraic sets is quasi-affine (quasi-projective) as well:

$$(A_1 \setminus A_2) \times (B_1 \setminus B_2) = (A_1 \times B_1) \setminus ((A_1 \times B_2) \cup (A_2 \times B_1)) . \qquad \square$$

As in Section 1.11, our discussion of morphisms begins with the study of admissible functions. For quasi-affine algebraic sets, these have been introduced in Definition 4.2.25. Adapting this definition, we get well-defined functions in the quasi-projective case:

Remark-Definition 6.3.10. Let $A \subset \mathbb{P}^n$ be a quasi-projective algebraic set. A function $f : A \to \mathbb{K}$ is called **regular at a point** $p \in A$ if there are *homogeneous* polynomials $g, h \in \mathbb{K}[x_0, \dots, x_n]$ *of the same degree* such that $h(p) \neq 0$ and f agrees with the function g/h on some open neighborhood of p in A. We say that f is **regular on $\boldsymbol{A}$** if it is regular at every point of A. The set $\mathcal{O}(A)$ of all regular functions on A becomes a ring, with pointwise defined algebraic operations. □

The definition is natural in that locally, in the coordinate charts of $\mathbb{P}^n$, we get the notion already familiar to us:

Exercise 6.3.11. Let $f : A \to \mathbb{K}$ be a function on a quasi-projective algebraic set $A \subset \mathbb{P}^n$. Show that the following are equivalent:

1. f is regular.
2. If $\pi : \mathbb{A}^{n+1} \setminus \{o\} \to \mathbb{P}^n$ is the canonical projection, then $f \circ \pi : \pi^{-1}(A) \to \mathbb{K}$ is regular in the sense of Definition 4.2.25.
3. For each coordinate chart U_i, the composition $f \circ \varphi_i^{-1} : \varphi_i(A \cap U_i) \to \mathbb{K}$ is regular in the sense of Definition 4.2.25. □

We use the regular functions to define morphisms:

Definition 6.3.12. Let A be a quasi-affine or quasi-projective algebraic set.

1. Let $B \subset \mathbb{A}^m$ be a quasi-affine algebraic set. A map $\varphi : A \to B$ is called a **morphism** if it is given by a tuple of regular functions: There exist functions $f_1, \dots, f_m \in \mathcal{O}(A)$ such that
$$\varphi(q) = (f_1(q), \dots, f_m(q)) \text{ for all } q \in A.$$
2. Let $B \subset \mathbb{P}^m$ be a quasi-projective algebraic set. A map $\varphi : A \to B$ is called a **morphism** if it is *locally* given by a tuple of regular functions: For any $p \in A$, there exist an open neighborhood U of p in A and functions $f_0, \dots, f_m \in \mathcal{O}(U)$ such that
$$\varphi(q) = [f_0(q) : \dots : f_m(q)] \text{ for all } q \in U. \qquad \square$$

As we will see in Example 6.3.19 below, the neighborhood U and the f_j in part 2 of Definition 6.3.12 may well depend on the point p. That is, the functions giving φ may not exist *globally*.

Exercise* **6.3.13.** Let A and B be quasi-affine or quasi-projective algebraic sets, and let $\varphi : A \to B$ be a map. Show that φ is a morphism iff the following two conditions hold:

1. φ is continous.
2. For any open subset $U \subset B$ and any regular function f on U, the composition $f \circ \varphi$ is a regular function on the open subset $\varphi^{-1}(U) \subset A$. □

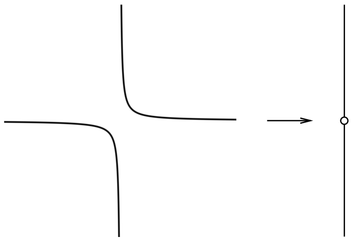
Clearly, the composition of two morphisms is a morphism. As usual, we have the notions of **isomorphism** and **isomorphic**. A morphism $\varphi : A \to B$ is said to be a **closed embedding** if $\varphi(A) \subset B$ is closed, and φ is an isomorphism of A onto $\varphi(A)$.

Example 6.3.14. 1. The canonical projection $\pi : \mathbb{A}^{n+1} \setminus \{o\} \to \mathbb{P}^n$ is a morphism.

2. The coordinate maps $\varphi_i : U_i \to \mathbb{A}^n$ are isomorphisms.

3. The Segre embedding $\sigma_{m,n}$ is a closed embedding.

4. Projecting onto the y-component, we get an isomorphism of the hyperbola $\mathrm{V}(xy-1) \subset \mathbb{A}^2$ with the punctured line $\mathbb{A}^1 \setminus \{0\}$.



Whereas the hyperbola is an affine algebraic set in the sense considered so far, the punctured line is not.

5. More generally, if $f \in \mathbb{K}[y_1, \dots, y_m]$ is any polynomial, then $\mathrm{V}(xf-1) \subset \mathbb{A}^{m+1}$ and $\mathrm{D}(f) \subset \mathbb{A}^m$ are isomorphic. □

To make the notion of affine and quasi-affine algebraic sets invariant under isomorphisms, we alter our definitions. For this, note that if $A \subset \mathbb{A}^n$ is a quasi-affine algebraic set, then $\varphi_0^{-1}(A) \subset U_0 \subset \mathbb{P}^n$ is quasi-projective, and φ_0 restricts to an isomorphism $\varphi_0^{-1}(A) \to A$. We may, thus, regard A as a quasi-projective algebraic set.

Definition 6.3.15. An **affine algebraic set** is a quasi-projective algebraic set which is isomorphic to an algebraic subset of some affine space. A **quasi-affine algebraic set** is defined similarly. □

A quasi-projective algebraic set $A \subset \mathbb{P}^n$ which is isomorphic to an algebraic subset of some $\mathbb{P}^m$ is necessarily a closed subset of $\mathbb{P}^n$ and, thus, a projective algebraic set in the sense of Definition 5.1.3 (see Theorem 6.3.26 below).

Exercise 6.3.16. Show that $A = \mathbb{A}^2 \setminus \{(0,0)\}$ is a quasi-projective algebraic set which is neither projective nor affine.
Hint. To exclude that A is affine, compute the ring $\mathcal{O}(A)$. □

The definition of a morphism says what conditions we require, but not how to create meaningful examples. Here is one possibility for the latter:

Remark 6.3.17. Let $A \subset \mathbb{P}^n$ be a quasi-projective algebraic set. Suppose that $f_0, \dots, f_m \in \mathbb{K}[x_0, \dots, x_n]$ are forms of the same degree, and such that $A \cap \mathrm{V}(f_0, \dots, f_m) = \emptyset$. Then we have a well-defined map

$$A \to \mathbb{P}^m,\ p \mapsto [f_0(p) : \dots : f_m(p)],$$

where $[f_0(p) : \dots : f_m(p)]$ is obtained by substituting the homogeneous coordinates of p for the x_i in the f_j. This map is a morphism: the open subsets $A \setminus \mathrm{V}(f_j)$ cover A, and on $A \setminus \mathrm{V}(f_j)$, the map is given by the tuple of regular functions $f_0/f_j, \dots, f_m/f_j$. □

Projection from a point gives an example. More generally, we have:

Example 6.3.18. Let $y_0, \dots, y_m \in \mathbb{K}[x_0, \dots, x_n]$ be linearly independent linear forms, and let $L = \mathrm{V}(y_0, \dots, y_m) \cong \mathbb{P}^{n-m-1}$ be the corresponding linear subspace of $\mathbb{P}^n$. Then the y_j define a morphism

$$\mathbb{P}^n \setminus L \to \mathbb{P}^m$$

which is called **projection from $\boldsymbol{L}$ to $\mathbb{P}^{\boldsymbol{m}}$**. □

Example 6.3.19. Let $n, d \geq 1$, let $N = \binom{d+n}{n} - 1$, and let $m_0, \dots, m_N$ be the monomials of degree d in $x_0, \dots, x_n$ (listed in some order). Then the m_j define a morphism

$$\rho_{n,d} : \mathbb{P}^n \to \mathbb{P}^N$$

which is called the ***d*-uple embedding** (or **Veronese embedding**) of $\mathbb{P}^n$ into $\mathbb{P}^N$. □

If $n = 1$ and d is arbitrary, we get the map

$$\rho_{1,d} : \mathbb{P}^1 \to \mathbb{P}^d,\ [s : t] \mapsto [s^d : s^{d-1}t : \dots : t^d],$$

whose image is the rational normal curve in $\mathbb{P}^d$ (see Exercise 6.2.5). Another special case is treated in the following example:

Example 6.3.20. If $n = d = 2$, we get the map

$$\rho_{2,2} : \mathbb{P}^2 \to \mathbb{P}^5,\ [a : b : c] \mapsto [a^2 : ab : b^2 : ac : bc : c^2].$$

Let V be the image of $\rho_{2,2}$, and let $w_0, \dots, w_5$ be the homogeneous coordinates on $\mathbb{P}^5$. Consider the symmetric matrix

$$\Delta = \begin{pmatrix} w_0 & w_1 & w_3 \\ w_1 & w_2 & w_4 \\ w_3 & w_4 & w_5 \end{pmatrix}.$$

Clearly, the 2 × 2 minors of Δ vanish on V. That is, if $I \subset \mathbb{K}[w_0, \dots, w_5]$ is the ideal generated by the minors, then $V \subset \mathrm{V}(I)$. We show that $V = \mathrm{V}(I)$, and that $\rho_{2,2}$ maps $\mathbb{P}^2$ isomorphically onto V. For this, we define a morphism $\varphi : \mathrm{V}(I) \to \mathbb{P}^2$ which, as the reader may easily check, is inverse to $\rho_{2,2}$. We consider a covering of $\mathrm{V}(I)$ by coordinate charts: $\mathrm{V}(I) \subset U_0 \cup U_2 \cup U_5$. On $\mathrm{V}(I) \cap U_0$, let φ be the map $p \mapsto [w_0(p) : w_1(p) : w_3(p)]$. On $\mathrm{V}(I) \cap U_2$ and $\mathrm{V}(I) \cap U_5$, define φ similarly by considering the second and third column of the matrix Δ. Since Δ has rank 1 one $\mathrm{V}(I)$, the respective local maps agree on the respective overlaps $U_i \cap U_j \cap \mathrm{V}(I)$, so that φ is well-defined. Note that φ is not a morphism of the type described in Example 6.3.17.

It turns out that I is in fact the vanishing ideal of V. To see this, we proceed as in the case of the Segre embedding, using a counting argument to show that I is prime ideal. This time, according to how we defined $\rho_{2,2}$, we consider the ring homomorphism

$$\phi : \mathbb{K}[w_0, \dots, w_5] \to \mathbb{K}[x, y, z],\ w_0 \mapsto x^2, w_1 \mapsto xy, \dots, w_5 \mapsto z^2,$$

whose kernel contains I. To show that $I = \ker \phi$, choose the degree reverse lexicographic order on $\mathbb{K}[w_0, \dots, w_5]$, where the variables are ordered such that $w_1, w_3, w_4 > w_0, w_2, w_5$. Then the leading monomials of the minors are

$$w_1^2, w_1w_3, w_3^2, w_1w_4, w_3w_4, w_4^2.$$

It follows that for each $d \geq 2$, there are precisely $3\binom{d+1}{2} + \binom{d+2}{2} = \binom{2d+2}{2}$ standard monomials of degree d. Hence, since the map

$$\mathbb{K}[w_0, \dots, w_5]_d / I_d \to \mathbb{K}[x, y, z]_{2d}$$

induced by ϕ is surjective, it must be an isomorphism. Thus, as in the case of the Segre embedding, a polynomial $f \in \mathbb{K}[w_0, \dots, w_5]$ is contained in $\ker \phi$ iff the remainder on division by the minors is zero. We conclude that the minors form a Gröbner basis for $\ker \phi$, and the result follows. □

The variety $V \subset \mathbb{P}^5$ in the example is known as the **Veronese surface**.

Exercise 6.3.21. Show that $\rho_{n,d}$ is a closed embedding for every n and d. Moreover, show that the vanishing ideal of the image is generated by quadrics which are binomials. How many quadrics do you get? □

In contrast to the affine case, the homogeneous coordinate ring of a projective algebraic set is not invariant under isomorphism:

Exercise 6.3.22. Let $A = \mathbb{P}^1$, and let $B \subset \mathbb{P}^2$ be the image of A under the 2-uple embedding. Then show that $S(A) \ncong S(B)$. □

Proposition 6.3.23. *Every quasi-projective algebraic set $A \subset \mathbb{P}^n$ has a finite open covering of affine algebraic sets.*

Proof. If $A = A_1 \setminus A_2$, where A_1 and $A_2 \subset A_1$ are closed subsets of $\mathbb{P}^n$, let $f_1, \dots, f_r \in \mathbb{K}[x_0, \dots, x_n]$ be forms such that $A_2 = \mathrm{V}(f_1, \dots, f_r)$ (if $A_2 = \emptyset$, take the linear forms $x_0, \dots, x_n$). Then $A = \bigcup_{i=1}^{r} \big(A_1 \setminus \mathrm{V}(f_i)\big)$. Hence, since $A_1 \setminus \mathrm{V}(f_i)$ is closed in $\mathbb{P}^n \setminus \mathrm{V}(f_i)$, it is enough to show that $\mathbb{P}^n \setminus \mathrm{V}(f)$ is an affine algebraic set for each form f. For this, we identify $\mathbb{P}^n$ with its image under the d-uple embedding of $\mathbb{P}^n$ into $\mathbb{P}^N$. Then $\mathrm{V}(f)$ is the intersection of $\mathbb{P}^n$ with a hyperplane H of $\mathbb{P}^N$. The result follows since $\mathbb{P}^n \setminus \mathrm{V}(f)$ is closed in $\mathbb{P}^N \setminus H \cong \mathbb{A}^N$. □

Remark 6.3.24. Let $A \subset \mathbb{P}^n$ be a quasi-projective algebraic set, and let

$$\phi = (f_{ij})$$

be a matrix of forms $f_{ij} \in \mathbb{K}[x_0 \dots, x_n]$, $1 \le i \le \ell$, $0 \le j \le m$. For all i, suppose that $\deg f_{ij}$ depends only on i. In addition, suppose:

1. $A \cap \mathrm{V}(f_{ij} \mid 1 \le i \le \ell,\ 0 \le j \le m) = \emptyset$;
2. All 2×2 minors of ϕ vanish on A.

Given a point $p \in A$, choose an index i such that $p \notin \mathrm{V}(f_{i0}, \dots, f_{im})$, and set $\varphi(p) = [f_{i0}(p) : \dots : f_{im}(p)]$. Then

$$\varphi : A \to \mathbb{P}^m,\ p \mapsto \varphi(p),$$

is a well-defined morphism. □

Exercise 6.3.25. If $A \subset \mathbb{P}^n$ is a quasi-projective algebraic set, show that every morphism $A \to \mathbb{P}^m$ is given by a matrix as in Remark 6.3.24 above. □

Morphisms between affine algebraic sets are easier to describe, but morphisms between projective algebraic sets are better behaved. For instance, as we already know, the image of an affine algebraic set under a morphism needs not be closed. In fact, the image may not even be a quasi-projective algebraic set: As an example, consider the map $\varphi : \mathbb{A}^2 \to \mathbb{A}^2$ corresponding to the substitution homomorphism

$$\mathbb{K}[x, y] \to \mathbb{K}[u, v],\ x \mapsto u,\ y \mapsto uv,$$

whose image is

$$(\mathbb{A}^2 \setminus \mathrm{V}(x)) \cup \{(0, 0)\}.$$

For the image of a projective algebraic set, however, we have:

Theorem 6.3.26. *Let A be a projective algebraic set, and let $\varphi : A \to B$ be a morphism of quasi-projective algebraic sets. Then $\varphi(A) \subset B$ is closed.*

Proof. The theorem follows from Lemma 6.3.27 and Theorem 6.3.28 below. □

Lemma 6.3.27. *If $\varphi : A \to B$ is a morphism of quasi-projective algebraic sets, then the graph of φ is a closed subset of $A \times B$.*

Proof. Closedness is a local property. Hence, by Corollary 6.3.23, we may replace B by an open affine subset U of B, and A by an open affine subset of $\varphi^{-1}(U) \subset A$. That is, we may suppose that A respectively B are algebraic subsets of some $\mathbb{A}^n$ respectively $\mathbb{A}^m$. Then φ is a polynomial map $(\overline{f}_1, \dots, \overline{f}_m)$, and its graph is defined by the ideal $\langle \overline{f}_1 - \overline{y}_1, \dots, \overline{f}_m - \overline{y}_m \rangle \subset \mathbb{K}[A \times B]$. □

Theorem 6.3.28 (Fundamental Theorem of Elimination Theory). *Let A be a projective algebraic set, and let B be any quasi-projective algebraic set. Then the projection $A \times B \to B$ is a closed map.*

Proof. As in the previous proof, we may suppose that B is an algebraic subset of some $\mathbb{A}^m$. Hence, if $\mathbb{P}^n$ is the ambient space of A, then $A \times B \subset \mathbb{P}^n \times \mathbb{A}^m$ is a closed subset, and it suffices to consider the case where $A \times B = \mathbb{P}^n \times \mathbb{A}^m$.

So let $X \subset \mathbb{P}^n \times \mathbb{A}^m$ be any closed subset. Then X is the common vanishing locus of polynomials $f_1, \dots, f_r \in \mathbb{K}[x_0, \dots, x_n, y_1, \dots, y_m]$, where each f_i is homogeneous in $x_0, \dots, x_n$ of some degree d_i. By the projective Nullstellensatz, a point $q \in \mathbb{A}^m$ is in the image Y of X iff the ideal

$$I(q) := \langle f_1(\boldsymbol{x}, q), \dots, f_r(\boldsymbol{x}, q) \rangle \subset \mathbb{K}[\boldsymbol{x}]$$

does not contain any of the ideals $\langle \boldsymbol{x} \rangle^d$, $d \geq 1$. Writing

$$Y_d = \{q \in \mathbb{A}^m \mid I(q) \supsetneq \langle \boldsymbol{x} \rangle^d\},$$

we have $Y = \bigcap_d Y_d$, and it suffices to show that Y_d is closed for any given d.

To obtain equations for Y_d, multiply each f_i with any monomial in $\boldsymbol{x}$ of degree $d - d_i$, and write T_d for the resulting set of polynomials. Then $q \in Y_d$ iff each monomial in $\mathbb{K}[\boldsymbol{x}]_d$ is a $\mathbb{K}$-linear combination of the polynomials $f(\boldsymbol{x}, q)$, $f \in T_d$. That is, the $f(\boldsymbol{x}, q)$, $f \in T_d$, span $\mathbb{K}[\boldsymbol{x}]_d$. Arranging the coefficients of the monomials $m \in \mathbb{K}[\boldsymbol{x}]_d$ appearing in the polynomials $f \in T_d$ as a $\binom{d+n}{n} \times \sum_i \binom{d-d_i+n}{n}$ matrix ϕ_d with entries in $\mathbb{K}[\boldsymbol{y}]$, the condition is that $\operatorname{rank} \phi_d(q) < \binom{d+n}{n}$. That is, the $\binom{d+n}{n} \times \binom{d+n}{n}$ minors of ϕ_d define Y_d. □

Remark 6.3.29. Theorem 6.3.26 is reminiscent of the fact that the image of a compact topological space under a continous map to an Hausdorff space is compact. Note that such a map is proper (that is, it is closed, and each fiber is compact). In complex analysis, Remmert's proper mapping theorem states that the image of a proper holomorphic map $f : X \to Y$ of complex analytic spaces is an analytic subset of Y (see **?**).

In algebraic geometry, the usual notion of properness is not suitable since the Zariski topology is not Hausdorff. There is, however, a corresponding notion of properness: A morphism $A \to B$ of quasi-projective algebraic sets is called **proper** if it can be factored as the composite of a closed embedding

$A \to \mathbb{P}^n \times B$ with the projection $\mathbb{P}^n \times B \to B$ (if A is projective, this condition is automatically fulfilled). It is clear from the proof of the fundamental theorem of elimination theory that Theorem 6.3.26 can be generalized to the following statement: If $\varphi : A \to B$ is a proper morphism of quasi-projective algebraic sets, then $\varphi(A) \subset B$ is closed. Moreover, it is easy to show that over the complex numbers, a morphism is proper in the sense of algebraic geometry iff it is proper in the usual sense with respect to the Euclidean topology. □

Corollary 6.3.30. *Let A be a projective variety. Then every regular function on A is constant. More generally, every morphism from A to an affine algebraic set is constant.*

Proof. Let $f \in \mathcal{O}(A)$. Then f defines a morphism $A \to \mathbb{A}^1 \subset \mathbb{P}^1$. The image is a closed, proper subset of $\mathbb{P}^1$ and consists, thus, of finitely many points. Being irreducible, it consists of a single point. This proves the first statement of the corollary. Composing with coordinate functions, we get the second one. □

Remark 6.3.31. For $\mathbb{K} = \mathbb{C}$, the corollary can also be deduced from the maximum modulus principle. Indeed, a regular function f on A is holomorphic. Since A is compact in the Euclidian topology, the modulus $|f|$ achieves its maximum on A. Hence, f is constant on every connected component of A (with respect to the Euclidean topology). The corollary follows since A is path connected by Theorem 6.7.13 in Section 6.6 below. □

Corollary 6.3.32. *Let $\pi : \mathbb{P}^n \setminus \{p\} \to \mathbb{P}^{n-1}$ be projection from the point $p = [1 : 0 : \cdots : 0]$. Let $A \subset \mathbb{P}^n$ be a projective algebraic subset such that $p \notin A$. Then $A' := \pi(A) \subset \mathbb{P}^{n-1}$ is an algebraic subset of $\mathbb{P}^{n-1}$. Moreover, the inclusion of homogeneous coordinate rings*

$$\mathbb{K}[A'] = \mathbb{K}[x_1, \dots, x_n]/\mathrm{I}(A') \longrightarrow \mathbb{K}[A] = \mathbb{K}[x_0, \dots, x_n]/\mathrm{I}(A)$$

is an integral ring extension, and $\dim A = \dim A'$.

Proof. The first statement is clear. For the second statement, we note that $\mathbb{K}[A] = \mathbb{K}[A'][\overline{x}_0]$ is finite over $\mathbb{K}[A']$. Indeed, since $p \notin A$, the vanishing ideal $\mathrm{I}(A)$ contains a form f of some degree $d \geq 1$ which is monic in x_0:

$$f = x_0^d + c_1(x_1, \dots, x_n)x_0^{d-1} + \ldots + c_d(x_1, \dots, x_n).$$

This shows that $\mathbb{K}[A'] \subset \mathbb{K}[A]$ is integral. For the last statement, write V_i and U_i for the coordinate charts on $\mathbb{P}^{n-1}$ and $\mathbb{P}^n$, respectively. Then, for $i = 1, \dots, n$, the inclusions of affine coordinate rings $\mathbb{K}[A' \cap V_i] \longrightarrow \mathbb{K}[A \cap U_i]$ are also finite: A polynomial in $\mathrm{I}(A \cap U_i)$ which is monic in x_0 is obtained by dehomogenizing f with respect to x_i. We conclude that $\dim A = \dim A'$. □

Corollary 6.3.33 (Projective Noether Normalization). *Let $A \subset \mathbb{P}^n$ be a projective algebraic set.*

1. *The dimension* $\dim A$ *is the least number* r *such that there is a linear subspace* $L \subset \mathbb{P}^n$ *of dimension* $n - r - 1$ *with* $A \cap L = \emptyset$.
2. *Let* $r = \dim A$*, and let* L *be any linear subspace as above. Then projection from* L *defines a morphism*

$$\pi : A \to \mathbb{P}^r$$

which is surjective and has finite fibers. Moreover, the map of homogeneous coordinate rings

$$\mathbb{K}[y_0, \ldots, y_r] \longrightarrow \mathbb{K}[A]$$

is a Noether normalization. In particular,

$$\dim A = \dim \mathbb{K}[A] - 1.$$

Proof. If $r = n$, then $A = \mathbb{P}^n$, and we are done. If $r < n$, there is a point $p \in \mathbb{P}^n \setminus A$. After a change of coordinates, we may suppose that $p = [1 : 0 : \cdots : 0]$. So the result follows from the preceeding Corollary by induction on $n - r$. □

Remark 6.3.34. A morphism $\varphi : A \to B$ of projective algebraic sets is called a **finite morphism** if for every point $q \in B$ there is an open affine neighborhood V of q in B such that $U := \varphi^{-1}(V)$ is affine, and the induced morphism $U \to V$ is finite in the sense of Chapter 3. We conclude from the proofs of the last two corollaries that the morphism $\pi : A \to \mathbb{P}^r$ above is finite. In the projective case, a morphism is finite iff it has finite fibers (see Harris (1992), Lemma 14.8). The example of the inclusion $\mathbb{A}^1 \setminus \{o\} \to \mathbb{A}^1$ shows that this is wrong in the affine case. □

Exercise 6.3.35. Show: The points corresponding to reducible polynomials $f = gh$ form an algebraic subset of $\mathbb{P}(\mathbb{K}[x_0, \ldots, x_n]_d)$. □

We finish this section by briefly treating Grassmanians. These are natural generalizations of projective spaces and provide important examples of projective varieties.

Definition 6.3.36. Given an n-dimensional vector space W over the field $\mathbb{K}$, the **Grassmannian** $\mathbb{G}(k, W)$ is the set

$$\mathbb{G}(k, W) = \big\{k\text{-dimensional linear subspaces of } W\big\}.$$

If $W = \mathbb{K}^n$, we write $\mathbb{G}(k, n)$ for $\mathbb{G}(k, W)$. □

Remark 6.3.37. Note that $\mathbb{G}(k, W)$ can also be thought of as the set of $(k-1)$-dimensional linear subspaces of the projective space $\mathbb{P}(W)$. □

To show that $\mathbb{G}(k, W)$ carries the structure of a projective variety, let $V \subset W$ be a k-dimensional linear subspace, and let $v_1, \ldots, v_k$ be a basis for V. Then $v_1 \wedge \cdots \wedge v_k$ is a nonzero vector of the exterior product $\bigwedge^k W$. This vector is determined by V up to scalar (choosing a different basis means to multiply the vector by the determinant of the change of basis matrix). We, thus, obtain a well-defined map

$$\mathbb{G}(k, W) \to \mathbb{P}(\bigwedge^k W) \tag{6.2}$$

whose image is the set of points corresponding to the totally decomposable vectors of $\bigwedge^k W$. This map is injective: if $v_1 \wedge \cdots \wedge v_k \in \bigwedge^k W$ represents a point p in the image, the kernel of the linear map

$$W \to \bigwedge^{k+1} W, \ w \mapsto w \wedge v_1 \wedge \ldots \wedge v_k,$$

is the unique linear subspace of W sent to p.

Definition 6.3.38. The map (6.2) is called the **Plücker embedding** of $\mathbb{G}(k, W)$ into $\mathbb{P}(\bigwedge^k W)$. The homogeneous coordinates on $\mathbb{P}(\bigwedge^k W)$ are called the **Plücker coordinates** on $\mathbb{P}(\bigwedge^k W)$. □

Note that if $p \in \mathbb{P}(\bigwedge^k W)$ corresponds to the linear subspace $V = \langle v_1, \ldots, v_k \rangle$ of W under the Plücker embedding, then the Plücker coordinates of p are the $k \times k$ minors of the $n \times k$ matrix with columns v_j.

Exercise* 6.3.39. With notation as above, show:

1. The Plücker embedding is a closed embedding.
2. Each coordinate chart of $\mathbb{P}(\bigwedge^k W)$ intersects $\mathbb{G}(k, W)$ in an affine space of dimension $k(n-k)$. □

We give $\mathbb{G}(k, W)$ the **structure of a projective variety** by identifying it with its image under the Plücker embedding.

6.4 Hilbert Functions and Hilbert Polynomials

Numerical invariants of a projective algebraic set such as the dimension are useful in that they allow us to partition a given classification problem into handy pieces. In this section, we will rediscover the dimension as the degree of the Hilbert polynomial, and we will use this polynomial to obtain other important invariants. Theorem 6.4.5, which shows the existence of the polynomial, is the fourth major result of Hilbert treated in this book. Hilbert's goal when proving the result was to encode the infinitely many values of what is nowadays called the Hilbert function in finite terms. The general context for the Hilbert function is that of graded modules.

Definition 6.4.1. Let $S = \bigoplus_{d \geq 0} S_d$ be a graded ring. A **graded module** over S is an S-module with a decomposition $M = \bigoplus_{d \in \mathbb{Z}} M_d$ as Abelian groups such that $S_d M_e \subset M_{d+e}$ for all d, e. An element of M_d is, then, called a **homogeneous element** of M of **degree** d. A **graded submodule** of M is a submodule generated by homogeneous elements. If $N = \bigoplus N_d$ is another graded S-module, a **graded homomorphism** from M to N is an S-module homomorphism $\phi : M \to N$ such that $\phi(M_d) \subset N_d$ for any d. □

If we consider S as a graded module over itself, its graded submodules are precisely its homogeneous ideals. The characterization of homogeneous ideals in Proposition 6.1.2 extends from the ideal to the submodule case:

Aushuehren

Furthermore and if $N = \bigoplus N_d$ is a graded submodule of $M = \bigoplus M_d$, then the quotient $M/N = \bigoplus M_d/N_d$ is graded as well. The direct sum of a collection of graded S-modules is naturally graded, and so are the kernel and the image of a graded homomorphism.

Example 6.4.2. Let S be a graded ring. Given a graded S-module $M = \bigoplus M_d$ and $\ell \in \mathbb{Z}$, the ***ℓ*th twist of *M***, written $M(\ell)$, is the graded S-module

$$M(\ell) = \bigoplus_{d\in\mathbb{Z}} M_{d+\ell}.$$

That is, $M(\ell)$ is isomorphic to M as an S-module, but its grading is shifted in degrees by ℓ. In particular, for each ℓ, we have the graded S-module $S(\ell)$ in which the free generator 1 of S has degree $-\ell$. Since each homomorphism of S is multiplication by an element of S, each graded homomorphism $S(k) \rightarrow S(\ell)$ is multiplication by a homogeneous element of S of degree $k - l$.

By specifying a basis together with a degree for each basis vector, a free S-module F becomes a **graded free *S*-module** (with a basis of homogeneous elements). That is, as a graded S-module, F is isomorphic to a direct sum of graded modules of type $S(\ell)$, for various ℓ. □

Each graded piece of a graded S-module M is an S_0-module and, thus, a $\Bbbk$-vector space if S is a graded $\Bbbk$-algebra. These vector spaces are *of finite dimension* if, in addition, S is *Noetherian*, and M is *finitely generated*. Indeed, in this case, M is Noetherian by Exercise 1.10.9. On the other hand, if M_e would not be of finite dimension for some e, the truncation $M_{\geq e} = \bigoplus_{d\geq e} M_d$ would be a submodule of M which is not finitely generated.

Definition 6.4.3. Let S be a Noetherian graded $\Bbbk$-algebra, and let $M = \bigoplus_{d\in\mathbb{Z}} M_d$ be a finitely generated graded S-module. The function

$$H(M, \_) : \mathbb{Z} \longrightarrow \mathbb{Z}, \quad d \longmapsto H(M, d) := \dim_{\Bbbk} M_d,$$

is called the **Hilbert function** of M. □

Example 6.4.4. Let S be the polynomial ring $\Bbbk[x_0, \ldots, x_n]$. Then

$$H(S, d) = \binom{d+n}{n}$$

for all $d \geq 0$. In fact, the formula holds for all $d \in \mathbb{Z}$ if we set $S_d = 0$ for $d < 0$. Thus, $H(S, d)$ agrees for $d \geq -n$ with the polynomial expression

$$\frac{(d+n)(d+n-1)\cdots(d+1)}{n!}.$$

We refer to this fact by saying that $H(S, \_)$ is of **polynomial nature**. □

More generally, we have:

Theorem 6.4.5 (Polynomial Nature of Hilbert Functions). *Let S be the polynomial ring $\Bbbk[x_0, \ldots, x_n]$, and let M be a finitely generated graded S-module. Then there is a unique polynomial $P_M(t) \in \mathbb{Q}[t]$ such that*

$$H(M, d) = P_M(d) \textit{ for all } d \gg 0.$$

Furthermore, $\deg P_M \leq n$. □

Definition 6.4.6. In the situation of the theorem, P_M is called the **Hilbert polynomial** of M. □

Following Hilbert, we will use *graded* free resolutions to reduce Theorem 6.4.5 to the special case considered in Example 6.4.4. Here is the relevant notation:

Definition 6.4.7. Let $S = \bigoplus_{d \geq 0} S_d$ be a graded ring. A **graded complex** of S-modules is a complex of S-modules where all modules and homomorphisms are graded. Similarly, we define the notions **graded free resolution** and **graded homomorphism of graded complexes**. □

In the context of graded free resolutions, we often write homomorphisms "from right to left" since this is consistent with how information on the resolutions is printed by computer algebra systems. Note that a graded homomorphism $F = \bigoplus_{i=1}^{s} S(\ell_i) \longleftarrow G = \bigoplus_{j=1}^{t} S(k_j)$ is given by an $s \times t$-matrix whose ij entry is a homogeneous element of S of degree $k_j - \ell_i$, for each pair i, j.

Example 6.4.8. If $S = \Bbbk[w, x, y, z]$, the matrix

$$\phi = \begin{pmatrix} x+y+z & w^2 - x^2 & z^3 \\ 1 & x & xy + z^2 \end{pmatrix}$$

defines a graded homomorphism

$$S \oplus S(-1) \xleftarrow{\phi} S(-1) \oplus S(-2) \oplus S(-3).$$

□

In the graded case, the recipe from Section 2.8 for constructing free resolutions yields a *graded* free resolution if we choose *homogeneous* generators at each stage. In the special case where S is the polynomial ring $\Bbbk[x_0, \ldots, x_n]$, we get a **graded version of the syzygy theorem**: Each finitely generated graded S-module M has a graded free resolution of length $\leq n+1$, with finitely generated graded free S-modules. Indeed, this follows from our constructive proof of the syzygy theorem in Chapter 2 and Remark 6.1.10 on the behaviour of Buchberger's algorithm in the graded case.

Example 6.4.9. Consider the ideal $I = \langle f_1, f_2, f_3 \rangle \subset S = \Bbbk[w, x, y, z]$, where $f_1 = x^2 - wy$, $f_2 = xy - wz$, and $f_3 = y^2 - xz$. Then I defines the twisted cubic curve in $\mathbb{P}^3$, and

$$\begin{tikzcd}
0 & S/I \arrow[l] & S \arrow[l] & S(-2)^3 \arrow[l, "(f_1,f_2,f_3)"'] & S(-3)^2 \arrow[l, "\begin{pmatrix} x & w \\ -y & -x \\ z & y \end{pmatrix}"'] & 0 \arrow[l]
\end{tikzcd}$$

is a graded free resolution of S/I. Note that f_1, f_2, f_3 are precisely the 2×2 minors of the 3×2 matrix in the resolution (with appropriate signs). This is no accident. It is, in fact, a consequence of the theorem of Hilbert-Burch, proved by Hilbert in his 1890 paper to give examples of free resolutions (see Eisenbud (1995), Theorem 20.15). □

Given a graded free resolution

$$\begin{tikzcd}
0 & M \arrow[l] & F_0 \arrow[l] & F_1 \arrow[l, "\phi_1"'] & \cdots \arrow[l] & F_{i-1} \arrow[l] & F_i \arrow[l, "\phi_i"'] & F_{i+1} \arrow[l, "\phi_{i+1}"'] & \cdots \arrow[l]
\end{tikzcd} ,$$

where all free modules are finitely generated, we usually collect all copies of S involving the same twist when writing F_i:

$$F_i = \bigoplus_j S(-j)^{\beta_{ij}} . \tag{6.3}$$

The β_{ij} are known as the **graded Betti numbers** of the resolution. A convenient way of visualizing these numbers is to write a **Betti diagram** as in the following example:

```
            0     1     2     3
------------------------------
    0:      1     -     -     -
    1:      -     2     1     -
    2:      -     2     3     1
------------------------------
total:      1     4     4     1
```

A number `i` in the top row of the diagram refers to the ith free module F_i of the resolution. More precisely, the column with first entry `i` lists the number of free generators of F_i in different degrees and, in the bottom row, the total number of free generators (that is, the rank of F_i). If `k:` is the first entry of a row containing a number β in the column corresponding to F_i, then F_i has β generators in degree $k + i$. That is, in (6.3), β is the number β_{ij} with $j = k+i$. The diagram above indicates, for instance, that F_2 has one generator in degree 3 and three generators in degree 4. In total, the diagram corresponds to a graded free resolution of type

$$S(-2)^2 \oplus S(-3)^2 \longleftarrow S(-3) \oplus S(-4)^3 \longleftarrow S(-5) \longleftarrow 0 .$$

Example 6.4.10. Resolving the homogeneous coordinate ring of the twisted cubic curve as in Example 6.4.9, we get the Betti diagram below:

```
               0     1     2
------------------------------
       0:      1     -     -
       1:      -     3     2
------------------------------
total:         1     3     2
```

□

In general, the β_{ij} cannot be called invariants of M since they depend on the choices made when constructing the resolution. Over a Noetherian graded $\Bbbk$-algebra S, the concept of minimal free resolutions takes care of this problem. To show the uniqueness of such a resolution, we need a graded version of Nakayama's lemma. In comparison with the local version, we replace the uniquely determined maximal ideal $\mathfrak{m}$ by the ideal S_+, which is the uniquely determined homogeneous maximal ideal if S is a graded $\Bbbk$-algebra.

Theorem 6.4.11 (Lemma of Nakayama, Graded Version). *Let S be any graded ring, let M be a finitely generated graded S-module, and let $N \subset M$ be a graded submodule. Then*

$$N + S_+M = M \;\; \textit{iff} \;\; N = M.$$

Proof. Reducing to the case $N = 0$ as in the proof of the local version, it suffices to show that $S_+M = M$ implies $M = 0$. Since M is finitely generated, $M_d = 0$ for $d \ll 0$. Suppose that $M \neq 0$, let d be the least d such that $M_d \neq 0$, and let $m \in M_d$ be a nonzero element. If $S_+M = M$, then m can be written as a sum $m = \sum_i s_i m_i$, with elements $s_i \in S_+$ and $m_i \in M$, and where we may assume that all s_i and m_i are nonzero and homogeneous. Then all $d_i = \deg s_i$ are strictly positive, so that $d - d_i < d$ for each i. This contradicts the fact that the M_{d-d_i} are zero by the choice of d. □

If S is a graded $\Bbbk$-algebra, and M is a graded S-module, then the quotient $\overline{M} = M/S_+M$ is a $\Bbbk$-vector space, and each graded homomorphism $\phi : M \to N$ induces a $\Bbbk$-vector space homomorphism $\overline{\phi} : \overline{M} \to \overline{N}$. As in the local case, Nakayama's lemma gives:

Corollary 6.4.12. *Let S be a graded $\Bbbk$-algebra, and let M be a finitely generated graded S-module. Then $m_1, \ldots, m_r \in M$ generate M as an S-module iff the residue classes $\overline{m}_i = m_i + S_+M$ generate $\overline{M} = M/S_+M$ as a $\Bbbk$-vector space. In particular, any minimal set of generators for M corresponds to a $\Bbbk$-basis for $\overline{M}$, and any two such sets have the same number of elements.* □

Let, now, S be a *Noetherian* graded $\Bbbk$-algebra, and let M be a finitely generated graded S-module. A **minimal free resolution** of M is obtained by choosing a minimal set of homogeneous generators at each stage of constructing a graded free resolution of M. Given any graded free resolution

$$\begin{tikzcd} 0 & M \arrow[l] & F_0 \arrow[l, "\phi_0"'] & F_1 \arrow[l, "\phi_1"'] & \ldots \arrow[l] & F_{i-1} \arrow[l] & F_i \arrow[l, "\phi_i"'] & F_{i+1} \arrow[l, "\phi_{i+1}"'] & \ldots \arrow[l] \end{tikzcd}$$

with finitely generated free modules, the images of the basis vectors of F_i under ϕ_i form a minimal set of generators for $\operatorname{im}\phi_i$ iff $\operatorname{im}\phi_{i+1} \subset S_+F_i$. That is, if we regard ϕ_{i+1} as a matrix, then ϕ_{i+1} does not have a nonzero scalar entry. In fact, the jth row of ϕ_{i+1} has an entry in $\Bbbk \setminus \{0\}$ iff the image of the jth basis vector of F_i under ϕ_i is an S-linear combination of the images of the other basis vectors.

Example 6.4.13. The resolution of the homogeneous coordinate ring of the twisted cubic curve in Example 6.4.9 is minimal. □

Minimal free resolutions are uniquely determined up to graded isomorphisms of complexes. This is a consequence of the following more general result:

Proposition 6.4.14. *Let S be a Noetherian graded $\Bbbk$-algebra, let M be a finitely generated graded S-module, and let*

$$\begin{tikzcd} 0 & M \arrow[l] & F_0 \arrow[l, "\phi_0"'] & F_1 \arrow[l, "\phi_1"'] & F_2 \arrow[l, "\phi_2"'] & \cdots \arrow[l] \end{tikzcd}$$

and

$$\begin{tikzcd} 0 & M \arrow[l] & G_0 \arrow[l, "\psi_0"'] & G_1 \arrow[l, "\psi_1"'] & G_2 \arrow[l, "\psi_2"'] & \cdots \arrow[l] \end{tikzcd}$$

be graded free resolutions with finitely generated graded S-modules. Suppose that the first resolution is minimal. Then there is a graded homomorphism of complexes

$$\begin{tikzcd} 0 & M \arrow[l] \arrow[d, "\mathrm{id}_M"'] & F_0 \arrow[l, "\phi_0"'] \arrow[d, "\alpha_0"'] & F_1 \arrow[l, "\phi_1"'] \arrow[d, "\alpha_1"'] & F_2 \arrow[l, "\phi_2"'] \arrow[d, "\alpha_2"'] & \cdots \arrow[l] \\ 0 & M \arrow[l] & G_0 \arrow[l, "\psi_0"'] & G_1 \arrow[l, "\psi_1"'] & G_2 \arrow[l, "\psi_2"'] & \cdots \arrow[l] \end{tikzcd}$$

such that each α_i is injective and identifies F_i with a direct summand of G_i:

$$G_i \cong F_i \oplus G_i', \quad \textit{for some graded free } S\textit{-module } G_i'.$$

Proof. Following the recipe from Exercise 2.8.17, starting from homogeneous free generators for F_0, we find a graded homomorphism α_0 such that the diagram

$$\begin{tikzcd} M \arrow[d, "\mathrm{id}_M"'] & F_0 \arrow[l, "\phi_0"'] \arrow[d, "\alpha_0"'] \\ M & G_0 \arrow[l, "\psi_0"'] \end{tikzcd}$$

commutes. If we regard α_0 as a matrix with entries in S, all entries and, in fact, all minors are homogenous. On the other hand, by Corollary 6.4.12, the induced maps on vector spaces $\overline{\phi_0}$ and, thus, also $\overline{\alpha_0}$ are injective. Hence, there is a $\operatorname{rank} F_0 \times \operatorname{rank} F_0$ minor of α_0 which is nonzero modulo S_+. Since the minor is homogeneous, it must be a nonzero scalar, so that the corresponding $\operatorname{rank} F_0 \times \operatorname{rank} F_0$ matrix is invertible over S. This shows that α_0 has the desired properties. The result follows by induction. □

Exercise 6.4.15. With S and M as in the proposition, design an algorithm which computes a minimal free resolution starting from any given graded free resolution (of finite length, with finitely generated free modules).
Hint. Use nonzero scalar entries of the given matrices as pivot elements as for Gaussian elimination. □

The proposition shows that the graded Betti numbers β_{ij} of a minimal free resolution depend on the finitely generated S-module M only. We, therefore, call these numbers the **graded Betti numbers of $\boldsymbol{M}$**, written $\beta_{ij}(M) = \beta_{ij}$.

Remark 6.4.16. Due to the local version of Nakayama's lemma, the concept of minimal free resolutions makes also sense over a *local* Noetherian ring R. If a finitely generated R-module M is given, its ***i*th Betti number** is the rank of the ith free module in the minimal free resolution of M. □

Proof of Theorem 6.4.5 (Hilbert). The uniqueness of P_M is clear. For the existence, consider any graded free resolution of M of length $\leq n+1$, with finitely generated free modules $F_i = \bigoplus_j S(-j)^{\beta_{ij}}$, where $S = \Bbbk[x_0, \dots, x_n]$:

$$\begin{tikzcd} 0 & M \arrow[l] & F_0 \arrow[l] & F_1 \arrow[l] & F_2 \arrow[l] & \cdots \arrow[l] & F_{n+1} \arrow[l] & 0 \arrow[l] \end{tikzcd}$$

Then, for each d, the graded pieces of degree d fit into an induced exact sequence of finite dimensional $\Bbbk$-vector spaces. Computing the alternating sum of the dimensions as in Exercise 2.8.4, we get

$$\begin{aligned} H(M,d) &= \sum_{i=0}^{n+1} (-1)^i \sum_j \beta_{ij} H(S(-j),d) \\ &= \sum_{i=0}^{n+1} (-1)^i \sum_j \beta_{ij} \binom{n-j+d}{n} \end{aligned}$$

(see Example 6.4.4). For each $d \geq j-n$, the value $H(S(-j),d)$ agrees with the polynomial expression

$$\frac{(d-j+n)(d-j+n-1)\cdots(d-j+1)}{n!}.$$

Hence, if P_M is the polynomial

$$P_M(t) = \sum_{i=0}^{n+1} (-1)^i \sum_j \beta_{ij} \binom{t-j+n}{n} \in \mathbb{Q}[t],$$

then $H(M,d) = P_M(d)$ for each $d \geq \max\{j-n \mid \beta_{ij} \neq 0 \text{ for some } i\}$. □

In algebraic geometry, the module M in Hilbert's theorem is the homogenous ccordinate ring of a projective algebraic set.

Definition 6.4.17. If $A \subset \mathbb{P}^n$ is an algebraic set, the **Hilbert polynomial** of A, written $P_A(t)$, is defined to be the Hilbert polynomial of the homogeneous coordinate ring $\mathbb{K}[A]$. □

Theorem 6.4.18. *If $A \subset \mathbb{P}^n$ is a projective algebraic set of dimension r, then its Hilbert polynomial is of type*

$$P_A(t) = d\,\frac{t^r}{r!} \;+\; \textit{terms of degree} < r,$$

where d is a strictly positive integer.

Proof. By Remark 3.3.2, there is a Noether normalization

$$\mathbb{K}[y_0, \dots, y_m] \subset \mathbb{K}[A] = \mathbb{K}[x_0, \dots, x_n]/\mathrm{I}(A)$$

such that the y_j are linear forms in the x_i. Then $A \cap \mathrm{V}(y_0, \dots, y_m) = \emptyset$ since, otherwise, the y_j would not be algebraically independent over $\mathbb{K}$. Hence, by Exercise **??**, projection from $\mathrm{V}(y_0, \dots, y_m)$ defines finite morphisms $A \cap \pi^{-1}(U_j) \to \pi(A) \cap U_j$, where the U_j are the coordinate charts of $\mathbb{P}^m$. In particular, $r = \dim A = m$. Furthermore, by the second defining condition of a Noether normalization, $\mathbb{K}[A]$ is a *finitely generated* graded $\mathbb{K}[y_0, \dots, y_r]$-module. Considering graded free resolutions over $\mathbb{K}[y_0, \dots, y_r]$, we see that the Hilbert function of $\mathbb{K}[A]$ is of type

$$H(\mathbb{K}[A], d) = \sum_{i=0}^{r+1} (-1)^i \sum_j \alpha_{ij} \binom{r - j + d}{r}.$$

It follows that $P_A(t) \in \mathbb{Q}[t]$ is a polynomial of degree $\leq r$, and that $r!\; P_A(t) \in \mathbb{Z}[t]$. On the other hand, since $\Bbbk[y_0, \dots, y_r]$ is a graded subring of $\mathbb{K}[A]$, we have

$$H(\mathbb{K}[A], d) \geq \binom{r+d}{r} \quad \text{for all } \; d.$$

We conclude that P_A has exactly degree r, and that its leading coefficient is strictly positive. □

Here are two corollaries of the proof:

Corollary 6.4.19. *Let $A \subset \mathbb{P}^n$ be a projective algebraic set, and let $C(A) \subset \mathbb{A}^{n+1}$ be the affine cone over A. Then*

$$\dim \mathbb{K}[A] = \dim C(A) = \dim A + 1.$$

Proof. We have $\dim \mathbb{K}[A] = \dim C(A)$, and this number is obtained via a Noether normalization as in the proof of the theorem. □

Corollary 6.4.20. *Let $A \subset \mathbb{P}^n$ be a projective algebraic set. Then* $\dim A$ *is the least number r such that there is a linear subspace $L \subset \mathbb{P}^n$ of dimension $n - r - 1$ with $A \cap L = \emptyset$.*

Proof. The projection from a linear subspace $\mathbb{P}^{n-r-1} \subset \mathbb{P}^n$ with $X \cap \mathbb{P}^{n-r-1} = \emptyset$ induces a morphism $X \to \mathbb{P}^r$, which is finite onto its image. If r is minimal, then the map is onto $\mathbb{P}^r$ and corresponds to a Noether normalization of the coordinate ring. □

Definition 6.4.21. *In the situation of Theorem 6.4.18, we write* $\deg A = d$, *and call this number the* **degree** *of* A. □

Though our definition is of purely algebraic nature, the degree has a geometric meaning: We will show in Proposition 6.6.11 that $\deg A$ is the number of points in which a *general* linear subspace of $\mathbb{P}^n$ of complementary dimension $n - \dim A$ intersects A.

Example 6.4.22. Let $A \subset \mathbb{P}^n$ be a hypersurface, let $f \in S = \mathbb{K}[x_0, \dots, x_n]$ be a square-free form defining A, and let $d = \deg f$. Then

$$\begin{tikzcd} 0 & \mathbb{K}[A] \arrow[l] & S \arrow[l] & S(-d) \arrow[l, "f"'] & 0 \arrow[l] \end{tikzcd}$$

is a graded free resolution of $\mathbb{K}[A]$, so that

$$\begin{aligned} P_A(t) &= \binom{n+t}{n} - \binom{n+t-d}{n} \\ &= d\,\frac{t^{n-1}}{(n-1)!} \;+\; \text{terms of degree} < n-1. \end{aligned}$$

Hence, $\deg A = d = \deg f$, and we conclude that our general definition of degree is consistent with that for hypersurfaces given earlier. □

Definition 6.4.23. Let $A \subset \mathbb{P}^n$ be a projective algebraic set of dimension r, with Hilbert polynomial P_A. The **arithmetic genus** of A is defined to be

$$p_a(A) = (-1)^r (P_A(0) - 1).$$

Let $C \subset \mathbb{P}^n$ be a curve. Then the Hilbert function can be written in the form

$$p_C(t) = dt + 1 - p_a.$$

p_a is called the **arithmetic genus** of C. □

Example 6.4.24. A plane curve $C \subset \mathbb{P}^2$ of degree d has Hilbert polynomial

$$p_C(t) = \binom{t+2}{2} - \binom{t-d+2}{2} = dt + 1 - \binom{d-1}{2}.$$

So C has arithmetic genus $p_a = \binom{d-1}{2}$. □

Remark 6.4.25. The funny way to write the constant term of the Hilbert polynomial comes from the Riemann-Roch Theorem 8.3.2.

We already mentioned that the degree d has a geometric interpretation. The arithmetic genus p_a has an even more fundamental interpretation. For a smooth irreducible curve over the complex numbers the arithmetic genus p_a determines the Euclidean topology of the underlying 2-dimensional manifold. By Corollary 8.4.7 and 8.2.6, the Euler number of the underlying 2-dimensional real manifold is $2 - 2p_a$. We will return to the arithmetic genus in Chapter 7, where we prove Riemann's inequality 7.4.12, and in Chapter 8. □

Exercise 6.4.26. Let $A \subset \mathbb{P}^4$ be the projective closure of the surface considered in Example 4.7.20. Compute equations for A as well as $\deg A$. □

Computing the Hilbert polynomial of a homogeneous coordinate ring S/I via syzygies may be costly since this means to compute Gröbner bases for I as well as for every kernel needed to construct a graded free resolution. The ideas of Macaulay only require the computation of a Gröbner basis for I:

Theorem 6.4.27 (Macaulay). *Let $S = \Bbbk[x_0, \dots, x_n]$, let F be a graded free S-module, and let $M \subset F$ be a graded submodule. For any global monomial order $>$ on F, we have*

$$H(F/M, \_) = H(F/\mathbf{L}_{>}M, \_).$$

Proof. By Macaulay's Theorem 2.3.5, the standard monomials of degree d represent $\Bbbk$-vector space bases for both $(F/M)_d$ and $(F/\mathbf{L}_{>}(M))_d$. □

Computing the intitial ideal of a homogeneous ideal J and then the Hilbert polynomial of $p_{S/\mathbf{L}(J)}$ is one of the fastest ways to obtain information about $\mathrm{V}(J)$.

Exercise 6.4.28. Let $A \subset \mathbb{P}^n$ be a projective algebraic set, and let B be its image under the d-uple embedding of $\mathbb{P}^n$ into $\mathbb{P}^N$, with $N = \binom{n+e}{d}$. Show that P_A and P_B have the same constant term:

$$P_A(0) = P_B(0).$$

□

Exercise 6.4.29. Let $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ be a hypersurface of bidegree (a, b). Show that C has degree $\deg C = a + b$ and arithmetic genus $p_a(C) = (a-1)(b-1)$. □

Exercise 6.4.30. Describe an algorithm which computes the minimal resolution from an arbitrary finite free resolution. Give a simplified algorithm which computes only the graded Betti numbers of the minimal resolution. □

Let $X \subset \mathbb{P}^n$ be a projective variety of dimension r, and let $H = \mathrm{V}(h) \subset \mathbb{P}^n$ be a hypersurface which does not contain X. By the Principal Ideal Theorem **??**, every component Z of $X \cap H$ has dimension $r - 1$.

Definition 6.4.31. *The* **intersection multiplicity** *of X and H along Z is the length of the Artinian ring $\mathcal{O}_{Z,\mathbb{P}^n}/(I_X^a + \langle h^a \rangle)\mathcal{O}_{Z,\mathbb{P}^n}$, where I_X^a and $\langle h^a \rangle$ denote the corresponding ideals in an affine chart $U_i \cong \mathbb{A}^n$ intersecting Z:*

$$i(X, H; Z) = \operatorname{length} \mathcal{O}_{Z,\mathbb{P}^n}/(I_X^a + \langle h^a \rangle)\mathcal{O}_{Z,\mathbb{P}^n}.$$

Example 6.4.32. The two hypersurfaces of Exercise 4.1.16 intersect along their common intersection curve with multiplicity two.

Theorem 6.4.33 (Bézout's Theorem, second version). *Let $X \subset \mathbb{P}^n$ be a projective variety and let $H \subset \mathbb{P}^n$ be a hypersurface which does not contain X. Let $Z_1, \dots, Z_s$ be the irreducible components of $X \cap H$. Then*

$$\deg X \cdot \deg H = \sum_{j=1}^{s} i(X, H; Z_j) \deg Z_j.$$

For the proof we need some preparations.

Definition 6.4.34. *Let M be a module over a ring R and $m \in M$. Then*

$$\operatorname{Ann}(m) = \{r \in R \mid rm = 0\}$$

is called the **annihilator** *of m.*

$$\operatorname{Ann}(M) = \{r \in R \mid rm = 0 \, \forall m \in M\}$$

is the annihilator of M. An **associated prime** $\mathfrak{p}$ *of M is a prime ideal which occurs as annihilator of an element.*

$$\mathfrak{p} = \operatorname{Ann}(m)$$

for some $m \in M \setminus 0$.

$$\operatorname{Ass} M = \{\mathfrak{p} \textit{ prime } \mid \mathfrak{p} = \operatorname{Ann}(m) \textit{ for some } m \in M\}$$

denotes the set of associated primes.

Thus with this notation the associated primes of an ideal I in the sense of Chapter 1 are the associated primes of the quotient R/I as R-module, and not the associated primes of the R-module I. This inconsistency in notation is unfortunate, but has a long tradition. In practise it rarely leads to confusion. The associated primes of R/I are of much more interest than the associated primes of the module I.

Remark 6.4.35. Note, that in case of an module over an affine coordinate ring

$$\mathrm{V}(\operatorname{Ann}(M)) = \bigcup_{\mathfrak{p} \in \operatorname{Ass} M} \mathrm{V}(\mathfrak{p})$$

is the **support** of M. The **minimal primes** in $\operatorname{Ass}(M)$ correspond to the irreducible components of the support of M. Non-minimal primes are called **embedded primes**, because their zero loci is strictly contained in a component.

Exercise 6.4.36. Let M be an R-module, and let $\mathfrak{q}$ be a prime ideal of R. Prove $M_{\mathfrak{q}} = 0$ iff $\mathfrak{q} \supset \mathfrak{p}$ for an associated prime $\mathfrak{p} \in \mathrm{Ass}(M)$ □

The set of associated primes is never empty for a module $M \neq 0$ over a Noetherian ring.

Lemma 6.4.37. *Every maximal element in the set* $\{\mathrm{Ann}(m) \mid m \in M \setminus 0\}$ *is a prime ideal.*

Proof. Let $\mathfrak{p} = \mathrm{Ann}(m)$ be maximal among the annihilators. Suppose $x, y \in \mathfrak{p}$ and $x \notin \mathfrak{p}$. Then $xm \neq 0$ and $\mathrm{Ann}(m) \subset \mathrm{Ann}(xm)$ and $y \in \mathrm{Ann}(xm) = \mathrm{Ann}(m) = \mathfrak{p}$ by the maximality. So $\mathfrak{p}$ is prime. □

Thus, by the Noetherian property there exist an associated prime.

Exercise 6.4.38. Every associated prime of a graded module is homogeneous. □

Exercise 6.4.39. Let

$$0 \to M' \to M \to M'' \to 0$$

be a short exact sequence. Prove:

$$\mathrm{Ass}(M') \subset \mathrm{Ass}(M) \subset \mathrm{Ass}(M') \cup \mathrm{Ass}(M'')$$

□

Proposition 6.4.40. *Let M be a finitely generated graded S-module. M has a filtration*

$$0 = M^0 \subset M^1 \subset \ldots \subset M^r = M$$

by graded submodules such that the quotients $M^i/M^{i-1} \cong (S/\mathfrak{p}_i)(a_i)$ for a homogeneous prime ideal $\mathfrak{p}_i$ and a twist a_i.

Proof. If $m \in M$ is a homogeneous element of degree a with $\mathfrak{p} = \mathrm{Ann}(m)$ an associated prime, then

$$(S/\mathfrak{p})(-a) \hookrightarrow M\,,\ r + \mathfrak{p} \mapsto rm$$

is an embedding.

Consider the set of graded submodules $N \subset M$, which have a filtration as in the proposition. This set is nonempty, because M has an associated prime. Let $M' \subset M$ be maximal in this set. We have to show $M' = M$. Suppose $M' \subsetneq M$. Consider an associated prime $\mathfrak{p}$ of M/M' and the inclusion $(S/\mathfrak{p})(a) \hookrightarrow M/M'$. Let M^+ be the preimage of $(S/\mathfrak{p})(a)$ in M. Then $M' \subsetneq M^+$ has a one step longer filtration. This contradicts the maximality of M'. □

Exercise 6.4.41. Let $\mathfrak{p}$ be a minimal prime of M. Then $\mathfrak{p}$ occurs precisely length $M_{(\mathfrak{p})}$-times in any filtration of M. □

Proof of Bézout's Theorem 6.4.33. We compute the Hilbert polynomial of $M = S/(I_X + I_H) = R_X/hR_X$ in two ways. The short exact sequence

$$0 \longrightarrow R_X(-\deg h) \xrightarrow{h} R_X \longrightarrow M \longrightarrow 0$$

gives us

$$\begin{aligned} p_M(t) &= \deg X(t^r/r! - (t - \deg H)^r/r!) + \text{ lower terms} \\ &= \deg X \cdot \deg H \; t^{r-1}/(r-1)! + \text{ lower terms} \; . \end{aligned}$$

On the other hand, the filtration of M gives

$$\begin{aligned} p_M(t) &= \sum_j p_{S/\mathfrak{p}_j}(t + a_j) \\ &= \sum_{\dim \mathrm{V}(\mathfrak{p}_j) = r-1} \deg \mathrm{V}(\mathfrak{p}_j)\, t^{r-1}/(r-1)! + \text{ lower terms} \\ &= \big(\sum_{j=1}^{s} i(X, H; Z_j) \deg Z_j\big)\, t^{r-1}/(r-1)! + \text{ lower terms} \; , \end{aligned}$$

because the number, in which $\mathrm{I}(Z_j)$ occurs in the filtration, coincides with $i(X, H; Z_j)$. Comparing the leading coefficients gives

$$\deg X \cdot \deg H = \sum_{j=1}^{s} i(X, H; Z_j) \deg Z_j$$

as desired. □

Exercise 6.4.42. Let M be a graded S-module. Prove that the Hilbert polynomial of M has degree

$$\deg p_M(t) = \dim \operatorname{supp} M = \max\{\dim \mathrm{V}(\mathfrak{p}) | \mathfrak{p} \in \operatorname{Ass}(M)\},$$

and that the leading coefficient is

$$\sum_{\dim \mathrm{V}(\mathfrak{p}) = \dim \operatorname{supp} M} \text{length}\, M_{(\mathfrak{p})} \frac{\deg \mathrm{V}(\mathfrak{p})}{r!}.$$

□

Apriori knowledge of the Hilbert function or Hilbert polynomial can ease Gröbner basis computation tremendously. We illustrate this in an example.

Example 6.4.43. Consider the morphism

$$\mathbb{P}^1 \to \mathbb{P}^3, \quad [x_0 : x_1] \mapsto [x_0^4, x_0^3x_1, x_0x_1^3, x_1^4].$$

We want to compute the equations of the image curve. One way to do this is to compute a Gröbner basis of

$$\langle y_0 - x_0^4, y_1 - x_0^3x_1, y_2 - x_0x_1^3, y_3 - x_1^4\rangle$$

with respect to a product order. Another way is to guess the equations and then argue. Clearly,

$$J = \langle y_1y_2 - y_0y_3, y_1^3 - y_0^2y_2, y_1^2y_3 - y_0y_2^2, y_1y_3^2 - y_2^3\rangle$$

is contained in the kernel $I = \ker \varphi$ of

$$\varphi : S = \mathbb{k}[y_0, \dots, y_3] \to R = \mathbb{k}[x_0, x_1].$$

To prove, that this is the Gröbner basis of the kernel, we compare various Hilbert functions. The image is $\operatorname{im}\varphi = \mathbb{k}[x_0^4, x_0^3x_1, x_0x_1^3, x_1^4] \subset R$. Hence, $S/I \cong \operatorname{im}\varphi$ has Hilbert function

$$h_{S/I}(t) = h_R(4t) = 4t + 1 \text{ for } t \geq 2.$$

On the other hand the lead terms of the generators of J with respect to the reversed lexicographic order and variables sorted $y_1 > y_2 > y_3 > y_0$ generate the ideal $J' = \langle y_1y_2, y_1^3, y_1^2y_3, y_2^3\rangle$. A $\mathbb{k}$-basis of S/J' is represented by the monomials in

$$\mathbb{k}[y_3, y_0] \oplus \mathbb{k}[y_3, y_0]y_1 \oplus \mathbb{k}[y_3, y_1]y_2 \oplus \mathbb{k}[y_3, y_0]y_2^2 \oplus \mathbb{k}[y_0]y_1^2.$$

Thus, $h_{S/J'}(t) = t + 1 + 2t + t - 1 + 1 = 4t + 1$ for $t \geq 2$. On the other hand,

$$h_{S/J'}(t) \geq h_{S/J}(t) \geq h_{S/I}(t).$$

Thus, equality holds, $I = J$, $\mathbf{L}(I) = J'$ and our 4 generators form a Gröbner basis. This completes our goal.

Finally, we can now easily compute the shape of the free resolution of S/I.

$$\begin{aligned}
M_2 &= \langle y_1y_2\rangle &&: y_1^3 &&= \langle y_2\rangle, \\
M_3 &= \langle y_1y_2, y_1^3\rangle &&: y_1^2y_3 &&= \langle y_1, y_2\rangle, \\
M_4 &= \langle y_1y_2, y_1^3, y_1^2y_3\rangle &&: y_2^3 &&= \langle y_1\rangle
\end{aligned}$$

Thus, the shape of the free resolution computed as in 2.8.11 is

$$0 \leftarrow S/I \leftarrow S \leftarrow S(-2) \oplus S(-3)^3 \leftarrow S(-4)^4 \leftarrow S(-5) \leftarrow 0,$$

and this gives the minimal free resolution, since no constant term is contained in syzygy matrices for degree reasons. Computing the Hilbert polynomial from the free resolution, we get

$$p_{S/I}(t) = \binom{t+3}{3} - \binom{t+1}{3} - 3\binom{t}{3} + 4\binom{t-1}{3} - \binom{t-2}{3} = 4t + 1,$$

which agrees already for $t \geq 2$ with the Hilbert function.

For practical computations, the idea of using the Hilbert function in Gröbner basis computations leads to the following spead up of the elimination algorithm. The key point is, that a Gröbner basis with respect to a weighted reversed lexicographic order is much sheaper to compute than a Gröbner basis with respect to an elimination order.

Algorithm 6.4.44 (Hilbert function driven elimination). Input: *A homogeneous ideal $I \subset \Bbbk[x_0, \ldots, x_n, y_0, \ldots, y_m]$ with weighted variables of possibly different degrees.* **Output**: *$I \cap \Bbbk[y_0, \ldots, y_m]$*
1. Compute a Gröbner basis with respect to the weighted reverse lexicographic order.
2. Compute the Hilbert function of $\Bbbk[x, y]/I$.
3. Compute a Gröbner basis with respect to an elimination order, but skip all Buchberger tests in a given degree, when there are already enough leading terms to account for the Hilbert function.

Example 6.4.45. Here is an example, where the Hilbert function driven Buchberger allows to compute the elimination ideal, while without this the computation takes much too long. Camera positioning.

Another way to present the Hilbert function is as follows:

Definition 6.4.46. *Let M be a graded module with $\dim M_d < \infty$ for all d. Then*

$$H_M(s) = \sum_{d \in \mathbb{Z}} \dim M_d s^d \in \mathbb{Z}[[s, s^{-1}]]$$

is called the **Hilbert series** *of M.*

Lemma 6.4.47. *Let M be a finitely generated graded module over the polynomial ring $\Bbbk[x_0, \ldots, x_n]$. Then H_M is the rational function*

$$H_M(s) = \frac{\sum_{ij} (-1)^i \beta_{ij} s^j}{(1-s)^{n+1}},$$

where the β_{ij} are the graded Betti numbers of M. If $r = \dim \operatorname{supp} M$ then the rational function $H_M(s)$ has a pole of order precisely $r + 1$ at $s = 1$.

Proof. The Hilbert series of $\Bbbk[x_1, \ldots, x_n]$ is

$$\frac{1}{(1-s)^{n+1}} = \sum_{t=0}^{\infty} \binom{n+t}{n} s^t.$$

Thus, expanding the rational function

$$\frac{\sum_{ij} (-1)^i \beta_{ij} s^j}{(1-s)^{n+1}}$$

at $s = 0$ yields a series whose coefficients satisfy the same formula as the Hilbert function $h_M(t) = \dim M_t$:

$$h_M(t) = \sum_{i=0}^{n+1} (-1)^i \sum_j \beta_{ij} \binom{n-j+t}{n}.$$

If we consider syzygies over a linear Noether normalization

$$\operatorname{supp} M \to \mathbb{P}^r$$

of $\operatorname{supp} M$, then we see that $H_M(s)$ has a pole of order at most $r+1$ at $s = 1$. It cannot have a pole of smaller order, because otherwise the coefficients of $H_M(s)$ would not grow fast enough. □

Remark 6.4.48. A similar formula holds for graded modules over a polynomial ring with generators x_i of different degrees. If $\deg x_i = d_i$ then the denominator takes the form $(1 - s^{d_0}) \cdot \ldots \cdot (1 - s^{d_n})$.

Example 6.4.49. Let $d_0, \ldots, d_n \in \mathbb{Z}_>0$ be a set of integers with no common divisor. We consider the group action of $\Bbbk^*$ on $\Bbbk^{n+1} \setminus 0$ defined by

$$\Bbbk^* \times \Bbbk^{n+1} \to \Bbbk^{n+1},\ (\lambda, (a_0, \ldots, a_n)) \mapsto (\lambda^{d_0} a_0, \ldots, \lambda^{d_n} a_n).$$

The **weighted projective space**

$$\mathbb{P}(d_0, \ldots, d_n) = (\Bbbk^{n+1} \setminus 0)/\Bbbk^*$$

is defined as the orbit space under this action. In case $d_0 = d_1 = \ldots = d_n = 1$ this is the ordinary projective space $\mathbb{P}^n$. We give $\mathbb{P}(d_0, \ldots, d_n)$ the structure of a projective variety as follows. Consider the polynomial ring $S = \Bbbk[x_0, \ldots, x_n]$ with grading induced by $\deg x_i = d_i$. Let $\ell = \operatorname{lcm}(d_0, \ldots, d_n)$ and let $m_0, \ldots, m_N \in S_\ell$ be a basis formed by monomials. Then

$$\mathbb{P}(d_0, \ldots, d_n) \to \mathbb{P}^N$$

induced by

$$a = (a_0, \ldots, a_n) \mapsto [m_0(a) : \ldots : m_N(a)]$$

is a well-defined embedding. However $\mathbb{P}(d_0, \ldots, d_n)$ is in general not smooth. The standard charts might carry some quotient singularities:

$$U_i = \{[a] \mid a_i = 1\} \cong \Bbbk^n/\mu_{d_i},$$

where μ_d denotes the group of d-th roots of unity.

Exercise 6.4.50. Prove that

$$\mathbb{P}(1,1,2) \cong V(x_0 x_2 - x_1^2) \subset \mathbb{P}^3.$$

□

Exercise 6.4.51. Consider $S = \Bbbk[x_0, \ldots, x_n]$ the polynomial ring with the grading induced by $\deg x_i = d_i$ and the corresponding weighted projective space. Let $I \subset S$ be a homogenous ideal with respect to this grading.

1. Prove that

$$V(I) = \{[a] \in \mathbb{P}(d_0, \ldots, d_n) \mid f(a) = 0 \text{ for all homogeneous } f \in I\}$$

is an algebraic subset of $\mathbb{P}(d_0, \ldots, d_n)$, and that every algebraic subset arises in this way.

2. Let

$$H_{S/I}(s) = \frac{\sum_{ij}(-1)^i \beta_{ij} s^j}{(1 - s^{d_0}) \cdot \ldots \cdot (1 - s^{d_n})}$$

be the Hilbert series of S/I according to Remark 6.4.48. Prove that $V(I) \subset \mathbb{P}(d_0, \ldots, d_n)$ has dimension r iff $H_{S/I}(s)$ has a pole of order $r + 1$ at $s = 1$.

□

Exercise 6.4.52. Complete the proof of Theorem 3.3.8.

Hint: Consider the projective closure in a suitable weighted projective space $\mathbb{P}(1, d_1, \ldots, d_n)$, where $w = (d_1, \ldots, d_n) \in \mathbb{Z}^n_{>0}$ is a weight vector, such that the Gröbner basis for the given monomial order $>$ and the weight order $>_w$ coincides. □

Exercise 6.4.53. Let $I \subset \Bbbk[x_0, \ldots, x_n] = \Bbbk[x]$ be a homogeneous ideal, and let

$$\varphi : \Bbbk[y_0, \ldots, y_m] \to \Bbbk[x]/I, y_i \mapsto f_i + I$$

be the substitution homomorphism induced by homogeneous forms $f_i \in \Bbbk[x_0, \ldots, x_n]$ of degree $\deg f_i = d_i$. Let

$$J = I\Bbbk[x, y] + \langle y_0 - f_0, \ldots, y_m - f_m \rangle$$

be the ideal of the graph in $\mathbb{P}(\deg x_0, \ldots, \deg x_n, d_0, \ldots d_m)$ of the corresponding rational map

$$\mathrm{V}(I) \dashrightarrow \mathbb{P}(d_0, \ldots, d_m).$$

Prove

$$H_{\Bbbk[x]/I}(s) = H_{\Bbbk[x,y]/J}(s).$$

□

6.5 Dimension Formulas

Theorem 6.5.1 (on the dimension of intersections). *Let $X, Y \subset \mathbb{P}^n$ be two subvarieties. Then every component Z of $X \cap Y$ has dimension*

$$\dim Z \geq \dim X + \dim Y - n.$$

If the right hand side is non-negative then the intersection $X \cap Y$ is non-empty.

Proof. Consider the **join** $J(X,Y) \subset \mathbb{P}^{2n+1}$ defined by the ideal $I(X)+I(Y) \subset \Bbbk[x_0,\dots,x_n,y_0,\dots,y_n]$. $J(X,Y)$ is the union of all lines joining a point of $X \subset \mathbb{P}^n \subset \mathbb{P}^{2n+1}$ with a point of $Y \subset \mathbb{P}^n \subset \mathbb{P}^{2n+1}$ contained in two complementary linear subspaces $\mathbb{P}^n \subset \mathbb{P}^{2n+1}$. With $\mathbb{P}^n \cong \Delta = \mathrm{V}(x_0-y_0,\dots,x_n-y_n) \subset \mathbb{P}^{2n+1}$ the "diagonal" we have

$$X \cap Y = \Delta \cap J(X,Y).$$

A Gröbner basis of $J(X,Y)$ is the union of the Gröbner basis for X and for Y. So $\dim J(X,Y) = \dim X + \dim Y + 1$. On the other hand Δ is defined by $n+1$ equations. Thus the generalized Principle Ideal Theorem 4.6.19 gives the desired inequality for the dimension of each component of $X \cap Y$.

For the second statement we consider the affine cones $C(X), C(Y) \subset \mathbb{A}^{n+1}$. The origin $0 \in \mathbb{A}^{n+1}$ lies in the intersection of the cones. Since every component of the intersection $C(X) \cap C(Y)$ has dimension at least $\dim X + 1 + \dim Y + 1 - n - 1 \geq 1$, there is at least one component containing the origin properly. This component is a cone again. Hence, we obtain $X \cap Y \neq \emptyset$. □

Remark 6.5.2. The reader might ask, why we did not prove Bézout's Theorem in a more general version for intersections $X \cap Y \subset \mathbb{P}^n$, say in case all components Z of $X \cap Y$ have expected $\dim Z = \dim X + \dim Y - n$. The reason is that length $\mathcal{O}_{Z,\mathbb{P}^n}/(I_X + I_Y)\mathcal{O}_{Z,\mathbb{P}^n}$ no longer gives the correct intersection multiplicity for the Theorem.

Exercise 6.5.3. Consider the surface $X \subset \mathbb{P}^4$ from Example 4.7.20 and Exercise 6.4.26 and let $Y = \mathrm{V}(x_1 - x_3, x_3 - x_4) \subset \mathbb{P}^4$ be a plane passing through the improper node $p = [0:0:0:0:1]$. Prove that

$$\operatorname{length} \mathcal{O}_p/(I_X + I_Y)\mathcal{O}_p = 3,$$

although there are 3 intersection points away from the node. Thus, adding the various length gives at least 3+1+1+1=6, which is larger than $\deg X \deg Y = 5 \cdot 1$

The reason why the numbers do not match is, that the module $S/(I_X+I_H)$ for the intersection of X with a hyperplane H containing Y has already $\mathfrak{m}_p$ as an associated prime. Thus, the full intersection ring gets too large. □

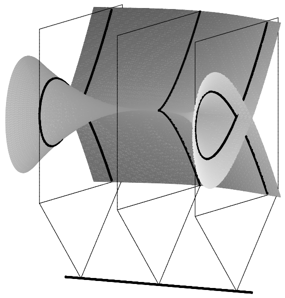
The general correct definition of the intersection multiplicity was a topic of Gröbner's research [1951]. In case of an intersection of two varieties of "expected" dimension $\dim X + \dim Y - \dim \mathbb{P}^n$, the correct definition was finally given by Serre [1957]:

$$i(X,Y;Z) = \sum_{i \geq 0} (-1)^i \operatorname{length} \operatorname{Tor}^i_{\mathcal{O}_{\mathbb{P}^n,Z}}(\mathcal{O}_{\mathbb{P}^n,Z}/I_X\mathcal{O}_{\mathbb{P}^n,Z}, \mathcal{O}_{\mathbb{P}^n,Z}/I_Y\mathcal{O}_{\mathbb{P}^n,Z}).$$

A disadvantage of this formula is that $i(X,Y;Z) > 0$ is no longer obvious. Fortunately, this is still true . In case we have a component Z of excess dimension, that is of dimension $\dim Z > \dim X + \dim Y - \dim \mathbb{P}^n$, one can apply the intersection theory of Fulton [1998] and/or Flenner, O'Carrel and Vogel [1999].

Let $\varphi : X \to Y$ be a morphism. For $q \in Y$ we call $X_q = \varphi^{-1}(q)$ the **fiber** of φ over q. On the right is the illustration of an affine piece of the surface

$$\begin{tikzcd} X = \mathrm{V}(y^2z - x^2(t^2z - x)) \subset \mathbb{P}^2 \times \mathbb{A}^1 \arrow[d, "\varphi"'] \\ Y = \mathbb{A}^1 \end{tikzcd}$$



and three fibers of the projection to the t-axis.

Theorem 6.5.4 (on the fiber dimension). *Let $\varphi : X \to Y$ be a projective morphism.*

1. *The function*
$$q \mapsto \dim X_q$$
is upper semi-continous on Y.
2. *If φ is a surjective map between varieties then there exists a non-empty open subset of $U \subset Y$, such that*
$$\dim X_q = \dim X - \dim Y$$
for all $q \in U$.

In particular, for a surjective projective morphism,

$$\dim X_q \geq \dim X - \dim Y$$

holds for every $q \in Y$.

Proof. 1.) We may assume that $X \subset Y \times \mathbb{P}^n$ is a closed subset. Let $q \in Y$ be a point and $\dim X_q = r$. Choose a linear subspace $\mathbb{P}^{n-r-1} \subset \mathbb{P}^n$ which does not intersect $X_q \subset \mathbb{P}^n$. Then $A = X \cap (Y \times \mathbb{P}^{n-r-1}) \subset Y \times \mathbb{P}^n$ is an algebraic set, whose image $pr_1(A) \subset Y$ contains all points q', where the fiber $X_{q'}$ has dimension $> r$ by Theorem 6.5.1 and perhaps some other points. Since the image is algebraic by 6.3.26 and $q \notin pr_1(A)$ the open set $V = Y \setminus pr_1(A)$ is an open neighborhood of q with $\dim X_{q'} \leq r$ for all $q' \in V$.

2.) We may assume that Y is affine and that $X \subset Y \times \mathbb{P}^n$. Consider the function fields $\Bbbk(Y) \subset \Bbbk(X)$.

$$\operatorname{trdeg}_{\Bbbk(Y)} \Bbbk(X) = \operatorname{trdeg}_{\Bbbk} \Bbbk(X) - \operatorname{trdeg}_{\Bbbk} \Bbbk(Y).$$

Let $I \subset \Bbbk[Y][x_0, \ldots, x_n]$ be the ideal of $X \subset Y \times \mathbb{P}^n$. We compute a Gröbner basis for $I \subset \Bbbk(Y)[x_0, \ldots, x_n]$ over the function field $\Bbbk(Y)$. The resulting Gröbner basis corresponds to a variety of dimension $\operatorname{trdeg}_{\Bbbk(Y)} \Bbbk(X)$ defined over $\Bbbk(Y)$. In such a computation of a Gröbner basis we have to invert finitely many leading coefficients in $\Bbbk[Y]$. Let f be the product of all these leading coefficients. Then for a point $q \in U = Y \setminus \mathrm{V}(f)$ the Gröbner basis of the ideal $I_q = \langle f(x, q) \mid f \in I \rangle$ defining X_q is obtained by substituting q into the coefficients of the Gröbner basis for $I \subset \Bbbk(Y)[x_0, \ldots, x_n]$. Thus, $\dim X_q = \operatorname{trdeg}_{\Bbbk(Y)} \Bbbk(X) = \dim X - \dim Y$ for all $q \in U$. We have proved more: the Hilbert function of $\Bbbk[x_0, \ldots, x_n]/I_q$ is the same for all $q \in U$.

The last statement follows from combining 1.) and 2.). □

Remark 6.5.5. 1. Assertion 6.5.4.1 does not hold without the hypothesis of projectivity. An example where the assertion does not hold is Example **??**.2.
2. An example of a projective morphism between varieties, where the fiber dimension is not constant, is the blow-up 7.2.1 below.

The following result has a very similar proof.

Theorem 6.5.6 (Reduction mod p). *Let $I = \langle f_1, \ldots, f_r \rangle \subset \mathbb{Q}[x_0, \ldots, x_n]$ be a homogeneous ideal defined by polynomials f_i with integer coefficients. For a prime number p we denote by $I_p \subset \mathbb{F}_p[x_0, \ldots, x_n]$ the ideal generated by the reduction of the f_i mod p. For all but finitely many primes the Hilbert function of $\mathbb{Q}[x_0, \ldots, x_n]/I$ and $\mathbb{F}_p[x_0, \ldots, x_n]/I_p$ coincide.*

Proof. We compute a normalized Gröbner basis of $I \subset \mathbb{Q}[x_0, \ldots, x_n]$. In this process we divide by finitely many leading terms. Let B be the set of primes, which devides a numerator of some of these leading terms. For p a prime outside B the computation of the Gröbner basis of I_p has exactly the same steps. In particular, $\mathbf{L}(I)$ and $\mathbf{L}(I_p)$ are generated by the same monomials. The result follows with Corollary 6.4.27. □

Remark 6.5.7. 1. Within Grothendieck's theory of schemes (eg. Hartshorne [1977], Chapter II and III), Theorem 6.5.6 and Theorem 6.5.4 have indeed a common generalization.
2. For practical purposes, Theorem 6.5.6 on the reduction mod p is of great importants. As long as we are only interested in the qualitative behavior of a system of equations, say in the dimension or degree, we can use a Gröbner basis computation mod p, which is much faster than the computations over $\mathbb{Q}$, because the bit length of the coefficients do not grow over $\mathbb{F}_p$. In doing so, we have to choose p outside B, which we usually do not know in advance. However, when choosing moderate size p, the chances for $p \in B$ are really low. The authors never had the bad luck to choose $p \in B$.

Exercise 6.5.8. Let $X \subset \mathbb{P}^n$ be a variety defined over $\mathbb{Q}$, and let $\mathrm{I}(X) = \langle f_1, \ldots, f_r \rangle$ be generators with integral coefficients. Let $I_p = \langle f_1, \ldots, f_r \rangle \subset$

$\mathbb{F}_p[x_0, \ldots, x_n]$ be generated by their reductions mod p. Prove: If X is non-singular, then $X_p = \mathrm{V}(I_p)$ is non-singular for all but finitely many primes p. □

Exercise 6.5.9. Let $f_1, \ldots, f_r \in \mathbb{Z}[x_0, \ldots, x_n]$ be homogeneous polynomials, and let $I \subset \mathbb{Q}[x_0, \ldots, x_n]$ and $I_p \subset \mathbb{F}_p[x_0, \ldots, x_n]$ denote the ideals generated by them over $\mathbb{Q}$ and $\mathbb{F}_p$, respectively. Prove

$$h_{\mathbb{F}_p[x_0,\ldots,x_n]/I_p}(t) \geq h_{\mathbb{Q}[x_0,\ldots,x_n]/I}(t) \text{ for all } t \in \mathbb{Z}.$$

□

6.6 Bertini's Theorem and other Applications

The dimension formulas have many applications. One of the most important is Bertini's theorem.

For a given projective space $\mathbb{P}^n = \mathbb{P}(V)$ the space of hyperplanes is natural the projective space of the dual vector space

$$\check{\mathbb{P}}^n = \mathbb{P}^n(V^*).$$

Theorem 6.6.1 (Bertini). *Let $X \subset \mathbb{P}^n$ be a smooth projective variety of dimension r. There exists a non-empty open subset $U \subset \check{\mathbb{P}}^n$, such that $X \cap H$ is smooth of dimension $r-1$ for every $H \in U$.*

Remark 6.6.2. It is true that for $\dim X \geq 2$ and $H \in U$ the intersection $X \cap H$ is also connected, hence irreducible. Frequently, this is considered to be part of Bertini's Theorem. The connectedness statement follows easily from cohomology theory of coherent sheaves, in particular Serre duality, which we do not treat in this book. See Hartshorne [1977] III.7.9. We will sketch a proof for fields $\mathbb{k}$ of characteristic zero in the appendix to this section.

Proof. We may assume that X is **non-degenerate**, i.e. that X spans $\mathbb{P}^n$. Then $X \cap H$ is singular at p iff $T_pX \subset H$. Since $T_pX \cong \mathbb{P}^r$ there exists an $\mathbb{P}^{n-r-1} \subset \check{\mathbb{P}}^n$ of hyperplanes H with $H \supset T_pX$. Consider the diagram

$$\begin{tikzcd} N = \{(p,H) \in X \times \check{\mathbb{P}}^n \mid T_pX \subset H\} \arrow[r] \arrow[d] & \check{\mathbb{P}}^n \\ X & \end{tikzcd}$$

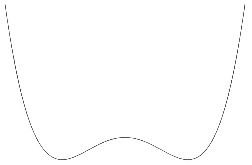
The fibers of $N \to X$ are $(n-r-1)$-dimensional. Hence, $\dim N = n-1$ and the image $\check{X}$ of N in $\check{\mathbb{P}}^n$ is at most a hypersurface. The open set $U = \check{\mathbb{P}}^n \setminus \check{X}$ has the desired property. □

Definition 6.6.3. *$\check{X} \subset \check{\mathbb{P}}^n$ is called the dual variety of $X \subset \mathbb{P}^n$. More generally, for possibly singular varieties $X \subset \mathbb{P}^n$ the dual variety is defined as the closure of the image of*

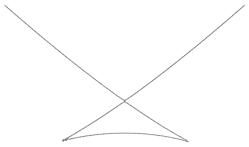
$$N^0 = \{(p, H) \in X^0 \times \check{\mathbb{P}}^n \mid T_pX \subset H\} \to \check{\mathbb{P}}^n$$

where $X^0 = X \setminus X_{sing}$ denotes the set of smooth points of X.

Example 6.6.4. For a plane curve $C \subset \mathbb{P}^2$ the dual variety is again a plane curve $\check{C} \subset \check{\mathbb{P}}^2$.



The curve defined by $y = x^4 - x^2$.



The dual curve in the chart $b = 1$.

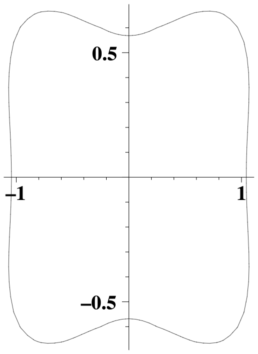
In this example the dual curve has equation

$$27a^4 - 4a^2b^2 + 144a^2bc - 16b^3c + 128b^2c^2 - 256bc^3 = 0$$

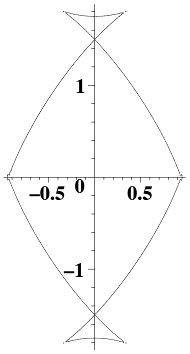
in coordinates a, b, c dual to x, y, z.

Exercise 6.6.5. Prove: An ordinary double point of $\check{C}$ corresponds to a bitangent of C. A cusp of $\check{C}$ corresponds to a flex of C. □

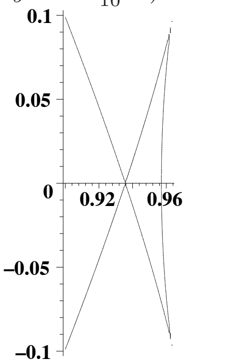
Exercise 6.6.6. Consider the curve $\mathrm{V}(x^4 + 4y^4 - x^2z^2 - y^2z^2 - \frac{1}{10}z^4) \subset \mathbb{P}^2$.



The curve with equation $x^4 + 4y^4 - x^2 - y^2 - \frac{1}{10}$



The dual curve.



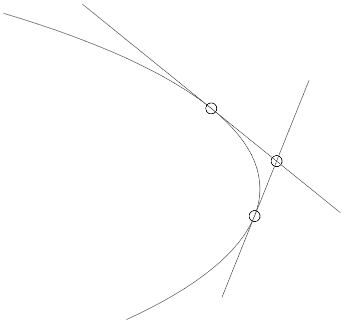
A detail of the dual curve.

Verify that the dual curve is defined by the equation

$$\begin{aligned} &a^{12} + \frac{20}{13}a^{10}b^2 + \frac{1297}{676}a^8b^4 + \frac{205}{169}a^6b^6 + \frac{239}{338}a^4b^8 + \frac{35}{169}a^2b^{10} + \frac{49}{676}b^{12} \\ &+\frac{290}{13}a^{10}c^2 - \frac{1210}{169}a^8b^2c^2 - \frac{3335}{338}a^6b^4c^2 + \frac{385}{338}a^4b^6c^2 - \frac{1355}{338}a^2b^8c^2 + \frac{385}{338}b^{10}c^2 \\ &+\frac{23430}{169}a^8c^4 - \frac{34000}{169}a^6b^2c^4 + \frac{34595}{338}a^4b^4c^4 - \frac{9250}{169}a^2b^6c^4 + \frac{30}{169}b^8c^4 \\ &+\frac{9600}{169}a^6c^6 + \frac{61800}{169}a^4b^2c^6 + \frac{37800}{169}a^2b^4c^6 - \frac{5400}{169}b^6c^6 - \frac{164800}{169}a^4c^8 \\ &-\frac{80000}{169}a^2b^2c^8 + \frac{800}{169}b^4c^8 + \frac{192000}{169}a^2c^{10} + \frac{48000}{169}b^2c^{10} - \frac{64000}{169}c^{12} = 0 \end{aligned}$$

Here a, b, c are dual coordinates to x, y, z. □

Exercise 6.6.7. Suppose $\operatorname{char} \Bbbk = 0$. Prove for an irreducible plane projective curve, that the double dual curce is the original curve, i.e. $\check{\check{C}} = C$.



□

Remark 6.6.8. In case of $\operatorname{char} \Bbbk = p > 0$, the double dual curve is not necessarily the original curve. For example, each tangent of the curve $\mathrm{V}(x^p + yz^{p-1})$ passes through the point $[1 : 0 : 0]$, so the dual curve is the line $L \subset \check{\mathbb{P}}^2$ dual to this point. A curve different from a line with the property that every tangent line passes through a fixed point is called **strange**. Strange curves exist only in $\operatorname{char} \Bbbk = p > 0$, by the exercise above. One can prove that strange curves are not smooth.

Exercise 6.6.9. ($\operatorname{char} \Bbbk = 0$). Prove $\check{\check{X}} = X$ for arbitrary varieties. □

Corollary 6.6.10. *Let $X \subset \mathbb{P}^n$ be a variety. There exists a open set $U \subset \check{\mathbb{P}}^n$ such that $(X \cap H)_{sing} = X_{sing} \cap H$ for all $H \in U$.*

Proof. $U = \check{\mathbb{P}}^n \setminus \check{X}$ has this property. □

Corollary 6.6.11. *Let $X \subset \mathbb{P}^n$ be a variety of dimension r and degree d. A general linear subspace $\mathbb{P}^{n-r} \subset \mathbb{P}^n$ intersects X in d distinct points.*

Proof. Combine Bertini's Theorem with Bézout's Theorem 6.4.33. □

Exercise 6.6.12. Let $X \subset \mathbb{P}^n$ an absolutely irreducible non-degenerate variety of dimension r. Prove

$$\deg X \geq n - r + 1.$$

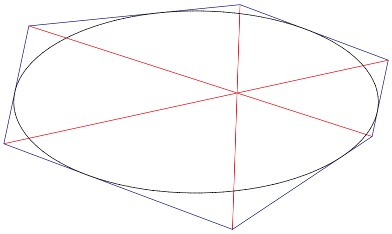
□

Exercise 6.6.13. Consider the d-th Veronese embedding

$$\mathbb{P}^n \hookrightarrow \mathbb{P}^N$$

with $N = \binom{n+d}{d}$. The dual variety $\check{X}$ of the image X can be identified with the set of singular hypersurfaces of degree d in $\mathbb{P}^n$. What is the degree of X? See Ge'lfand, Kapranov and Zelevinsky [1994] for a beautiful treatise on dual varieties. □

Exercise 6.6.14. Deduce Brianchon's Theorem from Pascal's Theorem 5.5.4 and projective duality: A hexagon in $\mathbb{P}^2$ is circumscribed to a smooth conic, if and only if the lines joining opposite vertices intersect in a point.



□

The dimension formulas and Bertini's theorem give another proof that every variety is birational to a hypersurface:

Theorem 6.6.15. *Let the ground field $\Bbbk$ be infinite. A variety $X \subset \mathbb{P}^n$ of dimension r can be birationally projected onto a hypersurface $X' \subset \mathbb{P}^{r+1}$*

Proof. We will project X from a center $\mathbb{P}^{n-r-2} \subset \mathbb{P}^n$ to $\mathbb{P}^{r+1}$. The induced map $X \to X' \subset \mathbb{P}^{r+1}$ is everywhere defined and finite if the center does not intersect X, which is the case for a general choice of the projection center. The problem is to prove that $X \to X'$ is birational. The preimage of a line $L \subset \mathbb{P}^{r+1}$ is a $\mathbb{P}^{n-r}$ containig the center of projection. For general choices this linear space will intersect X in $d = \deg X$ many distinct points by Bertini's Theorem. $X \to X'$ is birational in a neighborhood of one of these points iff none of the $d-1$ secant lines of X through the point intersect the center $\mathbb{P}^{n-r-2}$. We can acchieve this if we choose the primage $\mathbb{P}^{n-r}$ of the line first and then the center of projection $\mathbb{P}^{n-r-2} \subset \mathbb{P}^{n-r}$ such that it intersects none of the $\binom{d}{2}$ secant lines. □

Exercise 6.6.16. With the notation as in the proof of 6.6.15 and the additional assumption that $X \subset \mathbb{P}^n$ is smooth, prove that X' is either smooth and $X \to X'$ an isomorphismen, or X'_{sing} is of pure dimension $r-1$. □

The proof of the following theorem is of a simular flavour.

Theorem 6.6.17. *Every smooth projective curve can be embedded into $\mathbb{P}^3$.*

Proof. Suppose $C \subset \mathbb{P}^n$. If $n \leq 3$ there is nothing to prove. If $n \geq 4$ then we consider the secant variety. Consider the variety $\{(p_1, p_2, q) \in C \times C \times \mathbb{P}^r \mid p_1 \neq p_2 \text{ and } q \in \overline{p_1 p_2}\}$. The secant variety $Sec(C) \subset \mathbb{P}^n$ is the image of the closure of this set. Note that all tangent lines of C are contained in $Sec(C)$. By the dimension formula $\dim Sec(C) \leq 3$. Thus for $n \geq 4$ we can find a point $q \in \mathbb{P}^n \setminus Sec(C)$. The projection from p induces an isomorphism from C onto its image in $\mathbb{P}^{n-1}$. □

Exercise 6.6.18. Why is Theorem 6.6.17 not true for singular curves? □

Exercise 6.6.19. Prove that every smooth projective variety of dimension d can be embedded into $\mathbb{P}^{2d+1}$. □

Remark 6.6.20. Surfaces which can be embedded into $\mathbb{P}^4$ satisfy an identity between their numerical invariants, (see eg. Hartshorne [1977], Appendix A. Example 4.1.3). A famous result of Severi says that a non-degenerate smooth surface $X \subset \mathbb{P}^5$ can be isomorphically projected into $\mathbb{P}^4$ iff X is projectively equivalent to the Veronese surface, i.e. the image $\mathbb{P}^2 \hookrightarrow \mathbb{P}^5$ under the 2-uple embedding.

Exercise 6.6.21. a) Prove with Computer algebra the easy part of Severi's theorem: $\mathbb{P}^2 \hookrightarrow \mathbb{P}^5$ can be projected isomorphically. b) Describe the set of points in $\mathbb{P}^5$ from which one can project the Veronese surface isomorphically, and give a proof of the easy part without Computer algebra. (c) Prove the hard part of Severi's Theorem, i.e. no other surface in $\mathbb{P}^5$ can be projected isomorphically. □

6.7 Appendix: Monodromy Arguments

In this appendix we will prove the irreducibility of a general hyperplane $X \cap H$ section of a variety $X \subset \mathbb{P}^n$ of dimension $\dim X \geq 2$. We start by investigating general hyperplane sections of curves. Our first step is to establish the path connectedness of irreducible curves.

Theorem 6.7.1. *Let $f \in \mathbb{C}[x,y]$ be an irreducible polynomial and $C = \mathrm{V}(f) \subset \mathbb{A}^2(\mathbb{C})$ the corresponding plane algebraic curve. Then C equipped with the Euclidean topology is path connected.*

The proof of this result is interesting in its own. However, it requires some basic knowledge in Galois theory and analytic continuation of algebraic functions of one complex variable.

Proof. Let

$$f(x,y) = g_d(x)y^d + \ldots + g_0(x)$$

with coefficients $g_j(x) \in \mathbb{C}[x]$. If our coordinates are choosen general, then $d = \deg f$, a_d is a non-zero constant and $\deg g_j \leq d-j$. In that case we have counted with multiplicities precisely d solutions $(a,b) \in C$ for any given value $a \in \mathbb{C}$, and these solutions are distinct, iff the resultant $R(x) = Res(f, f_y)$ does not vanish at a. Moreover, the solutions depend continously on a. In particular, f has no isolated zeroes. In what follows we will not assume general coordinates. Then $g_d(x)$ might be a non-constant polynomial and some of the roots of $f(a,y)$ might approach infinity, if a approaches a zero of $g_d(x)$ in $\mathbb{A}^1(\mathbb{C}) = \mathbb{C}$. Let $B = \mathrm{V}(g_d(x)R(x)) \subset \mathbb{C}$. The projection onto the x-coordinate induces an unramified d sheeted covering

$$pr_1 : C \setminus pr_1^{-1}(B) \to \mathbb{C} \setminus B.$$

Since C has no isolated points, it suffices to prove that $C\setminus B$ is path connected. We will prove this with monodromy and Galois theory.

Let $p \in \mathbb{C} \setminus B$ be a base point. For each closed path

$$\gamma : [0,1] \to \mathbb{C} \setminus B \text{ with } \gamma(0) = \gamma(1) = p$$

and each preimage point $p_i \in \Gamma = pr_1^{-1}(p)$ path lifting defines a path $\gamma_i : [0,1] \to C \setminus pr_1^{-1}(B)$ which starts in $\gamma_i(0) = p_i$ and ends in a possibly different point $p_j = \gamma(1) \in \Gamma$. Thus, path lifting of γ induces a permutation

$$\mu(\gamma) : \Gamma \to \Gamma, \quad p_i \mapsto \gamma_i(1)$$

of Γ. We call the subgroup G, generated all permutations $\mu(\gamma)$ the **monodromy group** of the covering $pr_1 : C \setminus pr_1^{-1}(B) \to \mathbb{C} \setminus B$.

Path connectedness follows, if we can prove that G acts transitively on Γ. The key point is to identify G with a Galois group.

Consider the field extension $\mathbb{C}(x) \subset \mathbb{C}(x)[y]/\langle f \rangle$. Since f is irreducible, $\mathbb{C}[x,y]/\langle f \rangle$ is a domain, and $\mathbb{C}(x)[y]/\langle f \rangle$ is simply its quotient field. Let $K \supset \mathbb{C}(x)[y]/\langle f \rangle$ a splitting field of $f \in \mathbb{C}(x)[y]$. The splitting field K can be constructed explicitely as follows. Suppose that $p = 0 \in \mathbb{C}$ for notational convenience. We denote by $\mathbb{C}\{x\}$ the ring of convergent power series and by $\mathbb{C}\{x\}[x^{-1}] = Q(\mathbb{C}\{x\})$ the quotient field of meromorphic power series. We construct the splitting field of f over $\mathbb{C}(x)$ as a subfield of $\mathbb{C}\{x\}[x^{-1}]$. Let $p_i = (0, b_i) \in \Gamma$ be a point. By the Theorem on implicit functions , there exists an holomorphic power series $y_i(x) \in \mathbb{C}\{x\}$ with constant term $y_i(0) = b_i$, such that C near b_i equals the graph of y_i.

More precisely, there are $\epsilon, \delta > 0$, such that for

$$U_\epsilon(b_i) = \{y \in \mathbb{C} \mid |y - b_i| < \epsilon\} \text{ and } U_\delta(0) = \{x \in \mathbb{C} \mid |x| < \delta\},$$

we have

$$C \cap (U_\epsilon(b_i) \times U_\delta(0)) = \{(x, y_i(x)) \mid x \in U_\delta(0)\}.$$

Then

$$K \cong \mathbb{C}(x)[y_1(x), \ldots, y_d(x)] \subset \mathbb{C}\{x\}[x^{-1}],$$

indeed

$$f(x,y) = g_d(x)(y - y_1(x)) \cdot \ldots \cdot (y - y_d(x)).$$

We now consider the analytic continuation of our functions $y_i(x)$ along one of the closed path $\Gamma : [0,1] \to \mathbb{C} \setminus B$. This is possible, since for each point $\gamma(t)$ the implicit function theorem guarantees the existence of power serieses $y_{i,t}(x - \gamma(t)) \in \mathbb{C}\{x - \gamma(t)\}$, whose graphs parametrize C locally above $\gamma(t)$. For any t' in domain of convergence of the powerseries $y_{i,t}$, the function $y_{i,t}(x - \gamma(t))$ coincides with some $y_{i,t'}(x - \gamma(t'))$ in their common domain of definition, because both parametrize the same piece of C. At the end of the path the analytic continuation ends up with the same set of power series $y_1(x), \dots, y_d(x)$, however, possibly permuted. The permutation coincides with $\mu(\gamma)$.

We now claim, that each of these permutation induces an automorphism of the field K over $\mathbb{C}(x)$. Consider

$$\varphi : \mathbb{C}(x)[Y_1, \dots, Y_d] \to K \subset \mathbb{C}\{x\}[x^{-1}], \quad Y_i \mapsto y_i(x).$$

To prove that $\sigma = \mu(\gamma)$ gives an automorhism of K, we have to show that for any $F \in \ker \varphi$ the function $F(y_{\sigma(1)}(x), \dots, y_{\sigma(d)}(x)) = 0 \in K$. This follows from analytic continuation. The function $F(y_{1,t}(x-\gamma(t)), \dots, y_{d,t}(x-\gamma(t)))$ stays identically zero by the identity theorem for functions in one complex variable . Thus G is a subset of the Galois group $\mathrm{Gal}(f)$ of f.

The Theorem follows, if we can prove that $G = \mathrm{Gal}(f)$, because the Galois group of an irreducible polynomial acts transitively on the roots. So the following theorem completes the proof. □

Theorem 6.7.2. *With notation as above, the monodromy group G coincides with the Galois group of f over $\mathbb{C}(x)$.*

Proof. Let $h \in K^G$ an invariant function. Then by the definition of G, the invarinat function h has a well-defined meromorphic continuation to $\mathbb{C} \setminus B$. Moreover, also in B and infinity, the continuation of h cannot have an essential singularity, because it is a polynomial function in the local roots $y_{i,t}(x-\gamma(t))$ with coefficients in $\mathbb{C}(x)$. Thus, h extend to a meromorphic function on $\mathbb{P}^1(\mathbb{C})$. So h is rational. This proves $K^G = \mathbb{C}(x)$, and hence $G = \mathrm{Gal}(K/\mathbb{C}(x)) = \mathrm{Gal}(f)$ by the main theorem of Galois theory.

□

Remark 6.7.3. The image of a closed path γ in G depends only on the homotopy class of γ. What we really have is an group homomorphism

$$\pi_1(\mathbb{C} \setminus B, p) \to \mathrm{Aut}(\mathit{\Gamma}).$$

Here we use the notation $\pi_1(X, p)$ for Poincaré's **fundamental group** of homotopy classes of closed loops in a topological space X with base point p, and $\mathrm{Aut}(\mathit{\Gamma})$ denotes group of permutation group of the set $\mathit{\Gamma}$.

Thus to determine the image G, it suffices to apply path lifting to generators of $\pi_1(\mathbb{C} \setminus B, p)$. As it is well known , generators of $\pi_1(\mathbb{C} \setminus B)$ are small loops around each point of $b \in B$ connected via a path forwards and backwards to p.

This gives a numerical method to detect irreducibility of plane curves.

Corollary 6.7.4. *Any irreducible quasi-projective curve C over $\mathbb{C}$ is path connected with respect to the Euclidean topology.*

Proof. Consider a birational projection of C onto a plane curve C'. By the proof of the theorem, any non-empty Zariski open part of C' is path connected. Since we have an isomorphism of Zariski open parts of C' and C, and since C has no isolated points, C is path connected as well. □

Our next goal is to establishing the uniform position of a general hyperplane section of an irrreducible curve. This may be considered as an appropriate version of our desired irreduciblity result in case of curves.

Definition 6.7.5. *Let $\Gamma = \{p_1, \ldots, p_d\} \subset \mathbb{P}^n$ be a collection of d distinct points. Γ is in* **linearly uniform position***, if any subset of n points of Γ spans a $\mathbb{P}^{n-1}$. Γ is in (arithmetically)* **uniform position***, if the homogeneous ideals of any two subsets of Γ with the same number of elements have the same Hilbert function. The arithmetically uniform position is the stronger statement.*

Our goal is to prove that the general hyperplane section of an irreducible curve $C \subset \mathbb{P}^{n+1}$ over a field $\Bbbk$ of characteristic 0 is in uniform position. The assertion is not true in positive characteristic.

Exercise 6.7.6. Consider the curve

$$\mathrm{V}(x_0^2 - x_1x_4, x_1^2 - x_2x_4, x_2^2 - x_3x_4) \subset \mathbb{P}^4$$

over a field of characteristic 2. Prove that the points of a general hyperplane section form the vertices of a cube. □

To prove uniform position, we treat the case $\Bbbk = \mathbb{C}$ first. Let $C \subset \mathbb{P}^n(\mathbb{C})$ be an irreducible curve of degree d. Consider the Zariski open set $U = \check{\mathbb{P}}^n \setminus \check{C}$ of transversal hyperplanes. U is path connected in the Euclidean topology. Pick a base point $H_0 \in U$ and consider the monodromy action of the fundamental group $\pi_1(U, H_0)$ on $\Gamma = C \cap H_0 = \{p_1, \ldots, p_d\}$ defined by path lifting: Let

$$\gamma : [0,1] \to U,\ t \mapsto H_t$$

be a continuous path with $\gamma(0) = H_0$. Then by the continuity of roots of algebraic systems of equations there exist d continues paths

$$\gamma_i : [0,1] \to C \text{ with } \gamma_i(0) = p_i,$$

such that $C \cap H_t = \{\gamma_1(t), \ldots, \gamma_d(t)\}$ for all t. Since all H_t intersect transversally, a loop in U starting and ending in H_0 induces a permutation of Γ:

$$\Gamma \to \Gamma,\ p_i = \gamma_i(0) \mapsto \gamma_i(1),$$

which in fact depends only on the homotopy class of the closed loop. Thus, if $\pi_1(U, H_0)$ denotes Poincaré fundamental group consisting of homotopy classes of closed loops starting and ending at H_0, we obtain a homomorphism

$$\mu : \pi_1(U, H_0) \to \mathrm{Aut}(\Gamma)$$

to the symmetric group of permutations of Γ.

Theorem 6.7.7 (Harris' Monodromy Theorem). *Let $C \subset \mathbb{P}^n(\mathbb{C})$ be an irreducible curve of degree d. The monodromy action of $\pi_1(U, H_0)$ on $\Gamma = C \cap H_0$ gives the full symmetric group.*

Proof. We assume that C is not a line. We have to prove that ρ is surjective. For this, it is enough to prove that $\pi_1(U, H_0)$ acts double transitive and that the image contains a simple transposition. Applying if necessary a birational projection we may assume $n = 2$. Since the double dual $\check{\check{C}} \cong C$ by 6.6.7, there are only finitely many tangents lines passing through any point $q \in \mathbb{P}^2$, and all but finitely many tangent lines are simple tangents, i.e. tangent in precisely one smooth point of C, which is not a flex.

Consider $C' = C \setminus C_{sing}$ and the fibers X_p of the incidence variety

$$X = \{(p, H) \in C' \times U | p \in C \cap H\} \to C'.$$

C' is path connected by Corollary 6.7.4 and all X_p are path connected, since they are Zariski open subset of a $\mathbb{P}^1$. So X is path connected, which implies that $\pi_1(U, H_0)$ acts transitively. To see double transitivity, we choose a smooth point $p \in C'$ and choose the base point H_0 in the fiber X_p. The image of $\pi(X_p, H_0)$ lies in the stabilizer of p. Since

$$C'' = \bigcup_{H \in X_p} (C \cap H \setminus \{p\})$$

is still path connected by Corollary 6.7.4, we obtaion double transitivity. To exhibit a simple transposition, we look at a general point $H_1 \in \check{C}$. Then $H_1 \cap C$ is tangent at precisely one point with multiplicity 2. A small loop in U near H_1 around $\check{C}$ will interchanges the two nearby intersection points and leaves the other $d-2$ points unchanged. □

We denote with C^k the product $C \times C \times \ldots \times C$ and with $\Delta = \bigcup \Delta_{i,j}$ the union of the various diagonals.

Corollary 6.7.8. *The closure of $X_k = \{((p_1, \ldots, p_k), H) \in (C^k \setminus \Delta) \times U \mid \{p_1, \ldots, p_k\} \subset H \cap C\}$ in $C^k \times \check{\mathbb{P}}^n$ is irreducible for every k.*

Proof. X_k is non empty only for $k \leq d = \deg C$. It is path connected and irreducible, since we can connect any two points in the fiber of X_k over H_0 by a closed path in the smooth part of X_k according to Harris' Monodromy Theorem 6.7.7. □

Corollary 6.7.9. *The general hyperplane section $\Gamma \cap H$ of an irreducble curve lies in uniform position.*

Proof. Suppose that two subsets of Γ of the same cardinality k have different Hilbert functions. Since the values of the Hilbert function of a collection of points varies semicontineously with the points and H is general, this would give a decomposition of X_k into at least two components, a contradiction to Corollary 6.7.8. □

Remark 6.7.10. Much more general statements than Corollary 6.7.9 can be deduced. For example, the graded Betti numbers of the the image of any subset $\mathit{\Gamma}_1$ under the projection from the span of $\mathit{\Gamma}_2$ for disjoint subsets $\mathit{\Gamma}_1 \cup \mathit{\Gamma}_2 \subset \mathit{\Gamma}$ depend only on $\deg \mathit{\Gamma}_1$ and $\deg \mathit{\Gamma}_2$.

We now turn to arbitrary fields $\Bbbk$ of characteristic zero. First, if $X \subset \mathbb{P}^n$ is a quasi-projective algebraic set, then only finitely many coefficients occur in any finite set of defining equations of $\overline{X}$ and the complement $\overline{X} \setminus X$. The subfield $\Bbbk_0 \subset \Bbbk$ generated by these coefficients is a field of definition of X. Since $\Bbbk_0$ is a finitely generated field extension of $\mathbb{Q}$ and because $\mathbb{C}$ is algebraically closed with uncountable transcendence degree over $\mathbb{Q}$, there exists an embedding $\Bbbk \hookrightarrow \mathbb{C}$. Pick one and consider $X(\mathbb{C}) \subset \mathbb{P}^n(\mathbb{C})$. Then we apply

Lemma 6.7.11 (Lefschetz principle). *Let $\mathcal{P}$ be a property of algebraic sets which can be formulated by the solvability of a system of algebraic equations and inequalities with coefficients in the field of definition of X. If $X(\mathbb{C})$ satisfies $\mathcal{P}$ then $X(\overline{\Bbbk})$ satisfies $\mathcal{P}$, where $\overline{\Bbbk}$ denotes an algebraic closure of a field of definition of X.*

Proof. Clear, since we can embed $\Bbbk \hookrightarrow \mathbb{C}$. □

Let $C \subset \mathbb{P}^n$ be an absolutely irreducible curve over a field of characteristic zero. Let $U = \check{\mathbb{P}}^n \setminus \check{C}$ be the quasi-projective variety of transversal hyperplanes. For each k, the algebraic set $X_k = \{((p_1, \dots, p_k), H) \in (C^k \setminus \mathit{\Delta}) \times U \mid \{p_1, \dots, p_k\} \subset H \cap C\}$ in $C^k \times \check{\mathbb{P}}^n$ is absolutely irreducible, because it is irreducible over $\mathbb{C}$ by Corollary 6.7.8.

Corollary 6.7.12. *There exists a hyperplane $H \in U$ defined over the field of definition of C such that $\mathit{\Gamma} = C \cap H$ lies in uniform position in H.*

Proof. For each fixed t, the space of hyperplane H such that there exist two subsets $\mathit{\Gamma}_1$, $\mathit{\Gamma}_2$ of $C \cap H$ with the same number of points, but different values $h_{\mathit{\Gamma}_1}(t) \neq h_{\mathit{\Gamma}_2}(t)$, is a proper algebraic subset $B_t \subset U$ by Corollary 6.7.8 and the Lefschetz principle.

Since the Hilbert function $h_{\mathit{\Gamma}_1}(t)$ of a finite set of points takes value $\deg \mathit{\Gamma}_1$ for $t \geq \deg \mathit{\Gamma}_1$, there are only finitely many values t which we have to consider. Hence $B = \bigcup_{t \leq \deg \mathit{\Gamma}} B_t \subset U$ is an proper algebraic subset as well. (Without the bound for t, we would just conclude, that B is a countable union of proper algebraic subsets.) Therefore and because the field of definition $\Bbbk_0$ is infinite, the set of $\Bbbk_0$-rational points in $U \setminus B \subset \check{\mathbb{P}}^n$ is Zariski dense. □

To prove the irreducibilty of a general hyperplane section of a variety $X \subset \mathbb{P}^n$ of dimension $r \geq 2$, we consider the ground field $\mathbb{C}$ first, and start by extending the Mondromy Theorem 6.7.7 to this case.

Theorem 6.7.13. *Let $X \subset \mathbb{P}^n$ be a quasi-projective variety defined over $\mathbb{C}$. Then $X(\mathbb{C})$ is path connected.*

Proof. Adapt the proof of Theorem 6.7.1 and Corollary 6.7.4. □

Consider the Grassmannian

$$\mathbb{G} = \mathbb{G}(n-r+1, \mathbb{C}^{n+1}) = \{\mathbb{P}^{n-r} \subset \mathbb{P}^n\}$$

of complementary dimensional linear subspaces, see Exercise 6.3.39 for a definition of the Grassmannian as a projective variety. Let U be the open subset of transversal subspaces to X:

$$U = \{P \in \mathbb{G} \mid P \text{ intersects in } X \text{ in } d \text{ distinct points}\},$$

where $d = \deg X$. Pick a base point $P_0 \in U$ and consider the monodromy action of $\pi_1(U, P_0)$ on $\Gamma = X \cap P_0$.

Theorem 6.7.14. *Let $X \subset \mathbb{P}^n(\mathbb{C})$ be an irreducible variety of dimension r and degree d. The monodromy action of $\pi_1(U, P_0)$ on $\Gamma = X \cap P_0$ gives the full symmetric group.*

Proof. With minor modifications as before. □

Corollary 6.7.15. *Suppose* $\operatorname{char} \Bbbk = 0$. *A general hyperplane section $X \cap H$ of an irreducible variety $X \subset \mathbb{P}^n$ of dimension $r \geq 2$ is irreducible.*

Proof. We first consider the case $X \subset \mathbb{P}^n(\mathbb{C})$. Consider a flag $P_0 \subset H_0$ of a general complementary linear subspace P_0 and a general hyperplane H_0. By Berini's Theorem 6.6.1 H_0 intersects $X \setminus X_{sing}$ transversally. Suppose $X \cap H_0$ is reducible. Then every general hyperplane section is reducible. Since a loop

$$\gamma : [0,1] \to U, t \mapsto P_t$$

can be lifted to a loop of flags $t \mapsto (P_t, H_t)$ with P_t, H_t transversal to X, the monodromy action would distinguish between pairs of points in $X \subset P_0$, which do, respectively, which do not lie on the same irreducible component of $X \cap H_0$. This contradicts the Monodromy Theorem. Thus, $X \cap H_0$ is irreducible. For arbitrary fields of chararcteristic 0, the statement follows by applying the Lefschetz principle. □

Remark 6.7.16. As we see from the above, path lifting allows to establish an algorithmic test for absolute irreducibility of an algebraic set. Path lifting itself can be computed by numerical methods. An implementation numerical primary deomposition based on these ideas, has been developed by Sommese, Verschelde and Wampler, see Sommese, Wampler [2005].

Remark 6.7.17. A projective algebraic set $A \subset \mathbb{P}^n$ is called **non-degenerate**, if $I(A)$ contains no linear form, equivalently, if A spans $\mathbb{P}^n$. The study of **degenerate** algebraic sets can be reduced to non-degenerate ones by passing to a projective space of smaller dimension. The homogeneous coordinate ring of an non-degenerate algebraic set is generated by $n+1$ linear forms.

A **quasi-projective** variety W is a open subset of a projective variety, open with respect to the subspace topology. So $W = V \setminus A$, where V is a projective variety and $A \subset V$ an algebraic set. Quasi-projective varieties include both affine and projective varieties.

Graded modules over the polynomial are even better behaved then modules over a local ring.

Definition 6.7.18. *A* **graded ring** *R is a ring together with a decomposition*

$$R = \bigoplus_{d \geq 0} R_d,$$

such that the multplication respects the grading $R_d \times R_e \to R_{d+e}$. A **graded module** *$M$ over R is a module together with a decomposition*

$$M = \bigoplus_{d \in \mathbb{Z}} M_d,$$

such that $R_d \times M_e \to M_{d+e}$. We require, that homomorphisms of graded modules preseverve the degree.

Example 6.7.19. $S = \Bbbk[x_0, \ldots, x_n]$ is a graded ring. A homogenous ideal I is a graded module, the quotient ring $R = S/I$ is another example of a graded ring. In particular, for $X \subset \mathbb{P}^n$ we have the homogeneous ideal $I = I_X = \mathrm{I}(X)$ of X and the homogeneous coordinate ring $R_X = S/I_X$ of X.

Lemma 6.7.20 (Lemma of Nakayama in the graded case). *Let R be a graded ring, and let $R_{>0} = \bigoplus_{d>0} R_d$ be the ideal of elements of positive degree. Let $N \subset M$ be finitely generated graded R-modules. If $N + R_{>0}M = M$ then $N = M$.*

Proof. Since N and M are finitely generated, $N_d = M_d = 0$ for $d \ll 0$. Suppose $N \subsetneq M$. Consider the smallest d such that $N_d \subsetneq M_d$. Suppose $m \in M_d \setminus N_d$. By assumption $m = n + \sum_i r_i m_i$ for $n \in N$, $r_i \in R_{>0}$ and $m_i \in M$. Since we have graded modules, we may assume that this equation is homogeneous, i.e. $n \in N_d$, $r_i \in R_{d_i}$ and $m_i \in M_{d-d_i}$. Since $d_i > 0$ we have $d - d_i < d$. Hence by induction hypothesis $m_i \in N_{d-d_i}$. Hence $m \in N$, a contradiction. □

Corollary 6.7.21. *If R_0 is a field and M a finitely generated graded R module. Then $\dim_{R_0} M/R_{>0}M$ is the minimal number of generators of M.* □

With respect to $S = \bigoplus S_d$, we are mainly concerned with the case where the graded piece S_0 is a field $\Bbbk$. Then S is a $\Bbbk$-algebra, which we call a **graded $\Bbbk$-algebra**. We usually assume that S is finitely generated as a $\Bbbk$-algebra. Then every finitely generated S-module M is Noetherian by Exercise 1.10.9, and the graded pieces M_d are *finite dimensional* $\Bbbk$-vector spaces. Their dimensions are important numerical invariants of M.

Our next topic is that of a minimal free resolution which makes equally sense over local rings and in the graded case. In fact, in both cases, we can apply Nakayama's lemma whose graded version is as follows (note that the *homogeneous* maximal ideal plays the role of the maximal ideal considered earlier):

If $S = \bigoplus S_d$ is a graded ring such that $S_0 = \Bbbk$ is a field, then S is a $\Bbbk = S/S_+$-algebra, and the graded pieces S_d are $\Bbbk$-vector spaces. We, then, say that S is a **graded $\Bbbk$-algebra**.

Chapter 7

Rational Maps

In this Chapter we study rational maps and morphisms of curves to projective spaces.

In Section 7.2 we prove that a plane curve has an embedded resolution of singularities via blow-ups, and deduce that any curve is birational to a smooth projective curve. Our proof is based on the Cremona resolution, which transforms an arbitrary irreducible plane curve to a plane curve with only ordinary singularities. The detailed study of the blow-up and the quadratic transformation gives a first glimpse on the beauty of birational geometry in higher dimensions.

In Section 7.3 we introduce the concept of a divisor D on a smooth projective curve and its Riemann-Roch space $L(D)$. Divisors are the basic tool for the study of morphisms to projective spaces. In Section 7.4 we prove Riemann's inequality for the dimension of Riemann-Roch spaces, as a byproduct of an algorithm, which computes a bases of $L(D)$. Key point is here the completeness of hypersurface systems of large degree. As Corollary we will see the preliminary result that the arithmetic genus of a smooth projective curve, does not depend on the embedding, and that divisors of sufficiently large degree are very ample.

The final section 7.5 is devoted to the δ-invariant of a curve singularity both from a geometric and an arithmetic point of view. As an application we give an valuation formula for the intersection multiplicity of plane curves, and prepare the proof of the equality of the arithmetic and geometric δ-invariant: The proof will be complete, once we have established the equality of arithmetic and geometric genus with the help of the Riemann-Roch Theorem in Section 8.3. We begin recalling some basic facts about rational maps.

To easy our notation we make the following convention through out this Chapter: We work over an algebraically closed field $\Bbbk = \overline{\Bbbk}$. Only some times we comment on the field of definition of an object. For example we only mention in an exercise, that the resolution of singularities of a plane curve is defined over its the field of definition. A curve refers to an one dimensional variety. Its is important for computations, that the Riemann-Roch spaces $L(D)$ on a smooth curve C are defined over the common field of definition of the curve C and the divisor D.

7.1 Basic Facts

Basic facts about rational maps and rational functions have been established in Theorem 2.6.22, Proposition 4.2.26 and Corollary 6.3.30. We summarize what we know in case of varieties.

Definition 7.1.1. A **rational map** $f\colon X \dashrightarrow Y$ between two varieties X and Y is a morphism $f\colon U \to Y$ defined on a non-empty dense open subset $U \subset X$.

Example 7.1.2. If Y is affine, say $Y \subset \mathbb{A}^m$, then f is given by a tuple $f_1, \dots, f_m \in \Bbbk[U] \subset \Bbbk(X)$ of rational functions, such that $g(f_1, \dots, f_m) = 0$ for every equation $g \in I(Y) \subset \Bbbk[y_1, \dots, y_m]$ of Y.

Consersely, given a tuple of $f_1, \dots, f_m \in \Bbbk(X)$ which satisfies $g(f_1, \dots, f_m) = 0 \in \Bbbk(X)$ for every $g \in I(Y)$, we obtain a rational map $f\colon X \dashrightarrow Y \subset \mathbb{A}^m$ defined on the non-empty open set $U = \{p \in X \mid f_j \text{ is regular at } p \text{ for all } j\}$.

Example 7.1.3. Consider

$$\varphi\colon \mathbb{P}^2 \dashrightarrow \mathbb{P}^2, \quad [x:y:z] \mapsto [yz:xz:xy].$$

φ is defined outside $B = \mathrm{V}(xy, xz, yz) = \{[1:0:0], [0:1:0], [0:0:1]\}$. $\varphi^2 = \varphi \circ \varphi\colon \mathbb{P}^2 \dashrightarrow \mathbb{P}^2$ is given by

$$[x:y:z] \mapsto [x^2yz : xy^2z : xyz^2] = xyz[x:y:z]$$

Thus, outside the 3 lines $\mathrm{V}(xyz)$, we have $\varphi^2 = id_{\mathbb{P}^2}$.

Definition 7.1.4. Two rational maps $f, g\colon X \dashrightarrow Y$ are called equal (more precisely equivalent) if there exists a non-empty open set $U \subset X$ on which both maps are defined and equal.

Remark 7.1.5. If we want to be logically completely correct, we call $f|_U : U \to Y$ a representative of a rational map and defined a rational map $f\colon X \dashrightarrow Y$ to be an equivalence class of representatives.

With this notion we have $\varphi^2 = id_{\mathbb{P}^2}$.

Remark 7.1.6. In general rational maps $f\colon X \dashrightarrow Y$ and $g\colon Y \dashrightarrow Z$ cannot be composed. This reason is, that the image of f could be completely contained in the set, where $g\colon Y \dashrightarrow Z$ is not defined.

Definition 7.1.7. A rational map $f\colon X \dashrightarrow Y$ is called **dominant**, if $f(U) \subset Y$ is dense, i.e. $f(U) \cap V \neq \emptyset$ for all non empty open subset $V \subset Y$.

Thus, if f is dominant then $g \circ f\colon X \dashrightarrow Z$ is defined for every rational map $g\colon Y \dashrightarrow Z$. In Theorem 2.6.22 we proved

Theorem 7.1.8. *Let X and Y be varieties. There is a bijection*

$$\left\{ \begin{array}{c} \textit{dominant rational} \\ \textit{maps } \varphi\colon X \dashrightarrow Y \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \textit{inclusions} \\ \Bbbk(Y) \overset{\varphi^*}{\hookrightarrow} \Bbbk(X) \end{array} \right\}.$$

Definition 7.1.9. A dominant rational map $f\colon X \dashrightarrow Y$, which has a rational inverse, is called **birational**. In that case, X and Y are called **birationally equivalent**.

Corollary 7.1.10. *The category of birational equivalence classes of varieties with dominant rational maps as morphism is equivalent to the category of finitely generated fields over $\Bbbk$ with field inclusions as morphisms.*

Definition 7.1.11. For a given function field $\Bbbk(X)$, any variety Y birational to X is called a (geometric) **model** of the function field.

Rational maps from curves are considerably simpler than the general case. In Section 7.2 we will prove that any curve is birational to a smooth projective curve. Moreover rational maps from smooth curve can be extended to morphisms:

Theorem 7.1.12. *Let C be a smooth curve and $\varphi\colon C \dashrightarrow \mathbb{P}^r$ a rational map. φ has an extension to a morphism $\overline{\varphi}\colon C \to \mathbb{P}^r$.*

Proof. We may assume that C is irreducible and not contained in the hyperplane $V(x_0)$. The induced rational map $C \dashrightarrow U_0 = D(x_0) \cong \mathbb{A}^r$ is given by a tuple of rational functions $f_j = \varphi^*(x_i/x_0)$. Projectively, we have the map $C \dashrightarrow \mathbb{P}^r$ with $p \mapsto [1 : f_1(p) : \ldots : f_r(p)]$.

More general, for any tuple of rational functions, the map

$$C \dashrightarrow \mathbb{P}^r,\ p \mapsto [f_0(p) : f_1(p) : \ldots : f_r(p)]$$

is defined, where none of the rational functions has a pole and not all have simultaneously a zero. To extend such rational map over a point p, where it is not yet defined, we consider $n = \min\{v_p(f_j) | j = 0, \ldots, r\}$. Let $t \in \mathfrak{m}_p \subset \mathcal{O}_p \subset \Bbbk(C)$ a generator of $\mathfrak{m}_p$. Then $[t^{-n}f_0 : t^{-n}f_1 : \ldots : t^{-n}f_r]$ is defined in p and coincides with φ, where both tuples have a defined value, because the common factor t^{-n} does not matter projectively. □

In Section 7.2 we will proof that any irreducible curve is birational to a smooth projective curve. Using this we obtain from Theorem 7.1.8 and Theorem **??** the following:

Corollary 7.1.13. *Let $\Bbbk = \overline{\Bbbk}$ be algebraically closed. There is an equivalence between the categories*

$$\left\{\begin{matrix} \textit{smooth irreducible projective} \\ \textit{curves } C \textit{ over } \Bbbk \textit{ and} \\ \textit{dominant morphism } C \to E \end{matrix}\right\} \leftrightarrow \left\{\begin{matrix} \textit{finitely generated fields} \\ \Bbbk(C) \textit{ with } \operatorname{trdeg}_{\Bbbk} \Bbbk(C) = 1 \textit{ and} \\ \textit{algebraic field extensions } \Bbbk(E) \subset \Bbbk(C) \end{matrix}\right\}$$

Proof. Any non-constant rational map $C \dashrightarrow E$ between smooth irreducible projective curves, extends to a morphism $C \to E$, which is onto hence dominant. Indeed the image is an algebraic subset by Theorem 6.3.28, which is different from a point. Conversely any dominant morphism induces a field extension by Theorem 7.1.8. □

Hence the smooth projective model of a finitely generated function field of transcendence degree 1 is unique up to isomorphism.

Exercise 7.1.14. By the corollary a smooth irreducible projective curve C is completely determined by its function field $\Bbbk(C)$. The point of C can be recovered from the function field via valuation: Prove that the map $p \mapsto v_p$ defines a bijection

$$\{ \text{ points } p \text{ of } C\} \leftrightarrow \left\{\begin{matrix} \text{Discrete valuations} \\ v : \Bbbk(C) \setminus \{0\} \to \mathbb{Z} \\ \text{with } v(\Bbbk \setminus \{0\}) = 0 \end{matrix}\right\}$$

Hint: Use that an everywhere regular function on a projective curve is constant (Corollary 6.3.30).

Exercise 7.1.15. Let $X \subset \mathbb{P}^n$ be an irreducible variety, and let $f_0, \ldots, f_m \in \Bbbk[x_0, \ldots, x_n]$ be homogeneous polynomials of the same degree d, which not all vanish on X. Consider the rational map

$$X \dashrightarrow \mathbb{P}^m, \quad a \mapsto [f_0(a) : \ldots : f_m(a)].$$

Let $R = \Bbbk[x_1, \ldots, x_n]/\mathrm{I}(X)$ be the homogeneous coordinate ring of X, and let

$$R^{m+1} \to R(d)$$

be the homomorphism defined by the image of $f_0, \ldots, f_m$ in R. Consider the syzygies among $f_0, \ldots, f_m$ in R, and let $\overline{\varphi} = (\overline{\varphi}_{ij}) : \oplus_j R(-b_j) \to R^{m+1}$ be their syzygy matrix. Let $\varphi = (\varphi_{ij})$ be a matrix of representative in $\Bbbk[x_0, \ldots, x_n]$. Prove

1. The ideal

$$J = \langle I \rangle + \langle (y_0, \ldots, y_m) \cdot \varphi \rangle \subset \Bbbk[x_0, \ldots, x_n, y_0, \ldots y_m]$$

has the graph of the the rational map $X \dashrightarrow \mathbb{P}^m$ in $\mathbb{P}^n \times \mathbb{P}^m$ as a component.

2. The image of X in $\mathbb{P}^m$ is defined by the ideal

$$(J : \langle f_0, \dots, f_m\rangle^\infty) \cap \Bbbk[y_0, \dots, y_m].$$

□

Exercise 7.1.16. Choose a 4×5 matrix φ of linear foms in $\Bbbk[x_0, x_1, x_2]$ randomly and let $f_0, \dots, f_4$ denote the maximal minors of φ. Compute the image $\mathbb{P}^2 \dashrightarrow \mathbb{P}^4$ in two ways:

1. by computing algebra relations with Proposition 2.5.12 ,
2. by using Exercise 7.1.15.

□

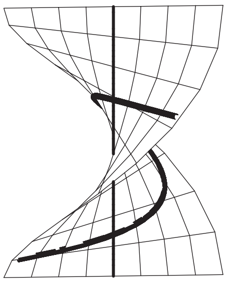
Exercise 7.1.17. Replace in Exercise 7.1.16 the matrix by a $n\times(n+1)$ matrix and the ground field $\Bbbk$ by a finite field $\mathbb{F}_q$. Compute the image of the rational map $\mathbb{P}^2 \dashrightarrow \mathbb{P}^n$ in the two ways. Which method is faster for large n in your experiment? □

7.2 Quadratic Transformation and Desingularization

Consider the origin $o \in \mathbb{A}^2$ and the variety

$$X = \mathrm{V}(xz_1 - yz_0) \subset \mathbb{A}^2 \times \mathbb{P}^1$$

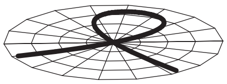
with coordinates $((x, y), [z_0 : z_1])$. Let $\sigma \colon X \to \mathbb{A}^2$ denote the morphism induced by the projection onto the first component. The only preimage of a point $(x, y) \neq 0$ is the point $((x, y), [x : y])$. On the other hand $E = \sigma^{-1}(o) = \{o\} \times \mathbb{P}^1$.



Definition 7.2.1. The morphism

$$\sigma \colon X \to \mathbb{A}^2$$

is called the **blow-up** of $o \in \mathbb{A}^2$. $E = \sigma^{-1}(o)$ is called the **exceptional** curve.



The intersection $X_0 = X \cap (\mathbb{A}^2 \times U_0) \cong \mathbb{A}^2$. Indeed, with affine coordinate $z = z_1/z_0$ on U_0, the equation restricts to $y - xz$ and hence $\Bbbk[X_0] \cong \Bbbk[x, z]$. $\sigma\,|_{X_0} \colon X_0 \to \mathbb{A}^2$ is given by $(x, z) \mapsto (x, xz)$. Similarly, for $X_1 = X \cap (\mathbb{A}^2 \times U_1) \cong \mathbb{A}^2$ we have

$$\sigma\,|_{X_1} \colon X_1 \to \mathbb{A}^2, \quad (y, w) \mapsto (yw, y).$$

Hence X is covered by two charts isomorphic to $\mathbb{A}^2$. The equations $w = 1/z$ and $y = xz$ hold in the intersection $X_0 \cap X_1$. In particular we see that X is a smooth irreducible surface.

If $f = f_r + f_{r+1} + \ldots \in \Bbbk[x, y]$ defines a curve C with multiplicity r at o then the preimage in X_0 is defined by

$$f(x, xz) = x^r (f_r(1, z) + x f_{r+1}(1, z) + \ldots).$$

Definition 7.2.2. The closure $C' = \overline{\sigma^{-1}(C \setminus \{o\})}$ of $\sigma^{-1}(C \setminus \{o\})$ is called the **strict transform** of C in X. The **total transform** is $\sigma^{-1}(C)$.

Thus, C' is defined on X_0 by the equation $f_r(1, z) + x f_{r+1}(1, z) + \ldots \in \Bbbk[x, z]$.

Lemma 7.2.3. *$C = \mathrm{V}(f)$ has an ordinary singular point of multiplicity r at o iff the strict transform C' is smooth in a neighborhood of E and intersects the exceptional curve E transversal in precisely r points. The intersetion points of C' with E correspond to the different tangents of C at o.*

Proof. After a linear change of coordinates we may assume that $\mathrm{V}(x)$ is not a tangent line of C at o. Then $f_r(1, z) + x f_{r+1} + \ldots = \lambda \prod_{i=1}^{r} (z - a_i) + x g(x, z)$ with $g = f_{r+1}(1, z) + x f_{r+2}(1, z) + \ldots$. Thus, $E \cap C' \cap X_0$ consists of the points $\{o\} \times [1, a_i]$. These are r distinct points, iff C has the r distinct tangents $\mathrm{V}(y - a_i x)$. In that case, C' is non-singular in these points and intersects $E = \mathrm{V}(x) \subset X_0$ transversally, because the equation has multiplicity 1 at these points with tangent $\mathrm{V}(z - a_i)$. □

Remark 7.2.4. More general we see that C' intersects E in precisely $r = \mathrm{mult}(C, o)$ points counted with multiplicities. The multiplicity of an intersection point is equal to the muliplicity of the corresponding tangent line.

Thus, blowing-up can improve the singularities. To repeat the blowing-up we generalize it to arbitrary varieties.

Definition 7.2.5. Let $p \in V \subset \mathbb{P}^n$ be a point on a quasi-projective algebraic set. Let $I(p) = \langle x_1, \ldots, x_n \rangle$ be the homogeneous ideal of $p \in \mathbb{P}^n$. Then the blow-up $\sigma_p : Bl_p V \to V$ of V in p is the closure of the graph $\mathit{\Gamma} \subset V \times \mathbb{P}^{n-1}$ of the projection

$$\pi_p : V \setminus \{p\} \to \mathbb{P}^{n-1}, q \mapsto [x_1(q) : \ldots : x_n(q)].$$

Exercise 7.2.6. Prove

1. The exceptional loci E of σ_p is the projectivized tangent cone of V at p

$$E = \sigma_p^{-1}(p) \cong \mathbb{P}(TC_p(V)) \subset \mathbb{P}(T_p\mathbb{P}^n) \cong \mathbb{P}^{n-1}.$$

The restriction $Bl_p V \setminus E \to V \setminus \{p\}$ is an isomorphism.

2. $\sigma_p : Bl_p V \to V$ does not depend on the embedding $V \subset \mathbb{P}^n$.

3. For two different points $p, q \in V$ we have
$$Bl_p(Bl_qV) \cong Bl_q(Bl_pV).$$

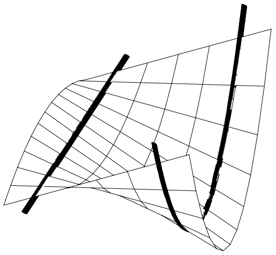
□

Remark 7.2.7. Quite frequently, $Bl_p(V) \cong \overline{\pi_p(V \setminus p)} \subset \mathbb{P}^{n-1}$. For example, if $V \subset \mathbb{P}^n$ is projective and no trisecant line of V passes through p. The last condition can be achieved, for example, by considering a suitable high Veronese image $V_d = \varphi_d(V)$ of V. Such image is cut out by quadrics, hence any trisecant would be contained in V_d, but V_d contains no line.

Example 7.2.8 (Cubic Scroll). Consider the Veronese surface $V_2 = \varphi_2(\mathbb{P}^2) \subset \mathbb{P}^5$. With coordinates $[x^2 : y^2 : z^2 : xy : xz : yz] = [w_0 : \ldots : w_5]$, the ideal of V_2 is defined by the 2×2 minors of
$$\begin{pmatrix} w_0 & w_3 & w_5 \\ w_3 & w_1 & w_4 \\ w_5 & w_4 & w_2 \end{pmatrix}.$$

To find the equation of the image $Bl_p(\mathbb{P}^2) \cong F \subset \mathbb{P}^4$ of V_2 under the projection from $p = [1 : 0 : 0 : 0 : 0 : 0] = \varphi_2([1 : 0 : 0])$, we eliminate w_0 from the equations, which leaves us with the 2×2 minors of
$$\begin{pmatrix} w_3 & w_1 & w_4 \\ w_5 & w_4 & w_2 \end{pmatrix}.$$

(Check this!). The resulting surface has degree 3 and is ruled by the lines $L_{[\lambda:\mu]} = \mathrm{V}(\lambda w_3 + \mu w_5, \lambda w_1 + \mu w_4, \lambda w_4 + \mu w_1)$. Thus F is a **ruled surface**, i.e. a surface ruled by lines. We call F the **cubic scroll**. Note, that $E = \mathrm{V}(w_1, w_4, w_2)$ is a further line on F. There are no more lines on F. However, there is a 2-dimensional family of conics on F. One of them is $C = \mathrm{V}(w_3, w_5, w_4^2 - w_1w_2)$. The pencil of lines joins points of E and C.



It corresponds to the strict transform of the pencil of lines $\mathrm{V}(\lambda x + \mu z)$ through $p = [1 : 0 : 0] \in \mathbb{P}^2$. Each conic is the image of a line $L \subset \mathbb{P}^2$ which does not pass through $p = [1 : 0 : 0]$. The exceptional curve of the blow-up is E.

Exercise 7.2.9. Let $F \subset \mathbb{P}^4$ be a non-degenerate smooth surface of degree 3. Prove that F is projectively equivalent to the cubic scroll. *Hint:* Project from a point on F. □

Examples of further ruled surfaces are the **rational normal scrolls**: For $0 \leq b \leq a \leq n-1$ with $a+b = n-1$ consider the surface $S(a,b) \subset \mathbb{P}^n$ defined

$$\operatorname{rank} \begin{pmatrix} x_0 \; x_1 \; \dots \; x_{a-1} \; y_0 \; \dots \; y_{b-1} \\ x_1 \; x_2 \; \dots \; \; x_a \; \; y_1 \; \dots \; \; y_b \end{pmatrix} < 2$$

in suitable coordinates $[x_0 : x_1 : \dots : x_a : y_0 : \dots : y_b]$ on $\mathbb{P}^{a+b+1} = \mathbb{P}^n$, i.e. we consider the ideal I generated by all 2×2 minors of the matrix above. It is easy to see that $\Bbbk[x_0, x_a - y_0, y_b] \hookrightarrow \Bbbk[x_0, \dots, y_b]/I$ is a projective Noether normalization, and hence $S(a,b)$ a surface of minimal degree $n-1$, compare Exercise 7.2.10. Moreover $I = \mathrm{I}(S(a,b))$ follows from the minimality of the degree by a counting argument with the initial ideal $\mathbf{L}(I)$ in appropriate coordinates. In particular we see that these surfaces are arithmetically Cohen-Macaulay. For $b > 0$ the surface $S(a,b)$ is smooth, and has the following description: Choose disjoint linear subspace $\mathbb{P}^a, \mathbb{P}^b \subset \mathbb{P}^n$ and rational normal curves $\varphi_a : \mathbb{P}^1 \hookrightarrow \mathbb{P}^a$ and $\varphi_b : \mathbb{P}^1 \hookrightarrow \mathbb{P}^b$ of degree a and b respectively. Then $S(a,b)$ is ruled by the family of lines joining $\varphi_a(p)$ and $\varphi_b(p)$, i.e.

$$S(a,b) = \bigcup_{p \in \mathbb{P}^1} \overline{\varphi_a(p)\varphi_b(p)}.$$

The surface $S(n-1,0)$ is the projective cone over the rational normal curve in $\mathbb{P}^{n-1}$.

With this notation, the cubic scroll is projectively equivalent to $S(2,1)$. The surface $S(1,1)$ is a quadric in $\mathbb{P}^3$.

****** fix ******Move the next exercise to section 6 after Bertini and Bezout

Exercise 7.2.10. Prove the following:

1. A non-degenerate variety $X \subset \mathbb{P}^n$ of dimension r has degree $\deg X \geq n - r + 1$.
2. A non-degenerate curve $C \subset \mathbb{P}^n$ of degree n in $\mathbb{P}^n$ is projectively equivalent to the rational normal curve . *Hint:* Project from a point on $C \subset \mathbb{P}^n$.

The complete classification of varieties of minimal degree is due to Bertini **?**, see also Harris **?**. A generalization to algebraic sets is the topic of a recent paper by Hulek, Eisenbud and Popescu **?**. For surfaces, we return to this topic in Exercise 7.2.11. □

Exercise 7.2.11. [Surfaces of minimal degree] Let $X \subset \mathbb{P}^n$ be a non-degenerate irreducible surface of minimal degree $d = n-1$. Prove that X is projectively equivalent to either a rational normal scroll $S(a,b)$ or to the Veronese surface $\mathbb{P}^2 \hookrightarrow \mathbb{P}^5$.

Exercise 7.2.12. Let $\{p_1, \dots, p_r\} \subset \mathbb{P}^2$ be a collection of r points. Prove that the image of $\mathbb{P}^2$ under the rational map defined by $L(d, p_1, \dots, p_r)$ for $d \geq r+2$ is isomorphic to the blow-up of $\mathbb{P}^2$ in the points $p_1, \dots, p_r$ (in any order). □

Returning to the question, how to improve singularities, we can formulate now:

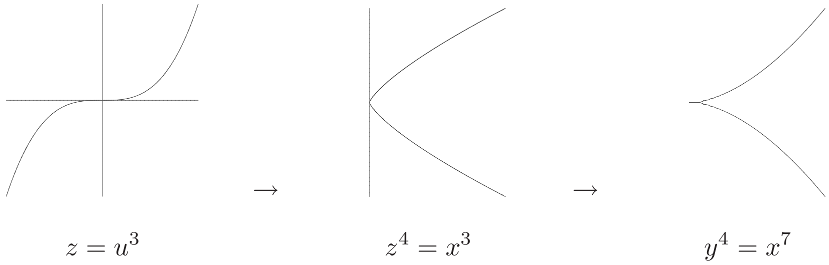
Theorem 7.2.13 (Embedded Resolution of Singularities). *Let $C \subset \mathbb{A}^2$ be a plane curve. There exists a sequence of blow-ups*

$$X^{(k)} \xrightarrow{\sigma_k} X^{(k-1)} \xrightarrow{\sigma_{k-1}} \dots \xrightarrow{\sigma_2} X^{(1)} \xrightarrow{\sigma_{(1)}} \mathbb{A}^2,$$

such that the strict transform $C^{(k)}$ of C in $X^{(k)}$ is smooth.

We call $C^{(k)}$ a **desingularization** of C.

Example 7.2.14. Consider $C \subset \mathbb{A}^2$ defined by $f(x,y) = y^4 - x^7$. The first strict transform is entirely contained in X_0 and defined by $f'(x,z) = z^4 - x^3$, because $f(x,zx) = x^4(z^4 - x^3)$. Since $f'(uz,z) = z^3(z-u^3)$ the second strict transform C'' is defined by $z - u^3$. Hence C'' is smooth.



C'' is parametrized by $u \mapsto (u,z) = (u,u^3)$. This gives for C' the parametrization $u \mapsto (x,z) = (uz,z) = (u^4,u^3)$. Finally, for C we obtain $u \mapsto (x,y) = (x,xz) = (u^4,u^7)$.

Remark 7.2.15. In general the resolution of singularities via blowing-up gives as in the example above, a method to compute a power series parametrization of the various branches of a singularity. For each point $p \in C^{(k)}$ we can solve the final equation $f^{(k)}$ locally with with the implicit function theorem to obtain a (formal) parametrization $t \mapsto (a_k(t), b_k(t)) \in C^{(k)}$ with $a_k, b_k \in \mathbb{k}[[t]]$, because p is a smooth point of $C^{(k)}$. Then successive substitutions give a parametrization

$$t \mapsto (a(t), b(t)) \in C$$

with $a, b \in \mathbb{k}[[t]]$ power series of higher order. If $\operatorname{char}(\mathbb{k}) = 0$ and $\operatorname{mult}(a(t)) = n$, we can achieve that $a(t) = t^n$ with a change of coordinates in $\mathbb{k}[[t]]$. The normalized parametrization

$$t \mapsto (t^n, b(t))$$

is uniquely determined up to an n-th root of unity by the branch. It is called the **Puiseux expansion** of the branch of the singularity. The set of branches of the singularity, is the set of such parametrizations. One can show that there is a bijection between the branches and the irreducible factors of $f \in \Bbbk[[x,y]]$ in the powers series ring. For the classical algorithm to compute Puiseux expansions based on the Newton polytope of f we refer to Brieskorn-Knörrer **?** or Walker **?**. An implemantation can be found in `SINGULAR`.

Exercise 7.2.16. Compute a Puiseux expansion of $\mathrm{V}(y^4 - x^6 + x^7)$. □

Although, the statement of Theorem 7.2.13 is local, we will use some global, i.e. projective arguments to deduce it. The problem is to make precise what has improved after one blow-up, and to establish that we reach a smooth curve after finitely many steps.

Key to our proof is a detailed study of the quadratic transformation.

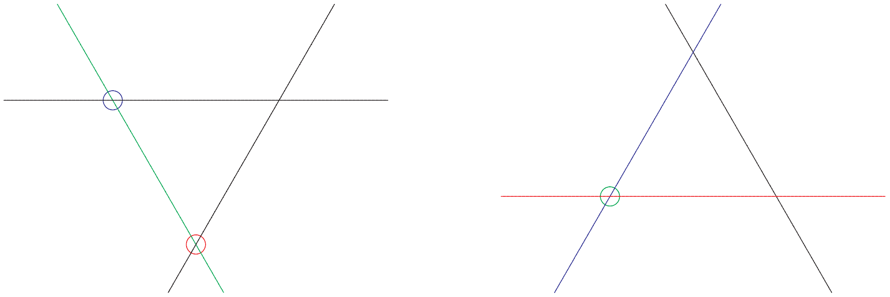
Definition 7.2.17. The birational map

$$q\colon \mathbb{P}^2 \dashrightarrow \mathbb{P}^2, \quad [x:y:z] \mapsto [\frac{1}{x} : \frac{1}{y} : \frac{1}{z}] = [yz : xz : xy]$$

is called **quadratic (Cremona) transformation**. $p_0 = [1:0:0], p_1 = [0:1:0]$ and $p_2 = [0:0:1]$ are called the fundamental points of q. The rational map q is not defined in these points. For a curve $C \subset \mathbb{P}^2$, we call the closure of the image $C' = \overline{q(C \setminus \{p_0,p_1,p_2\}}$ the strict transform of C.

Remark 7.2.18. More generally, a **Cremona transformation** is a birational transformation $\mathbb{P}^2 \dashrightarrow \mathbb{P}^2$. One can show, that the group of Cremona transformations is generated by linear automorphisms and the quadratic transformation. Max Noether **?** gave an incomplete proof of this fact. The discovery of the gap and the completion of the proof is due to Castelnuovo **?**.

To get an idea about the geometry of this transformation, it is convenient to work with an affine chart, which contains all three fundamental points. In other words we can define a quadratic tranformation for any set three non-colinear points, by composing q with a suitable linear automorphisms.



****** fix ******Move to Chapter 5 make an exercise out of this

In terms of affine coordinates $p_j = (x_{j,1}, \dots, x_{j,2}) \in \mathbb{A}^2$ of the fundamental points a point with homogeneous coordinates $[\lambda_0 : \lambda_1 : \lambda_2]$ say normalized such that $\sum_j \lambda_j = 1$ corresponds to the affine point

$$p = \sum_j \lambda_j p_j$$

and the proper real convex combination in case of real fundamental points correspond to the inner part of the triangle. Thus homogeneous coordinates generalize convex coordinates from convex geometry.

Proposition 7.2.19. *The graph of the quadratic transformation is isomorphic to the blow-up of $\mathbb{P}^2$ in the 3 fundamental points. The projection onto the second factor is again a blow up of three points, the strict transforms of the 3 fundamental lines $L_{ij} = \overline{p_i p_j}$ are the exceptional curves of the second projection.*

Proof. The graph G is defined as the closure of the graph of the morphism $U \to \mathbb{P}^2$, which represents Q. We use coordinates [x:y:z] and [u:v:w] on $\mathbb{P}^2 \times \mathbb{P}^2$. The graph is defined by

$$\operatorname{rank} \begin{pmatrix} yz & xz & xy \\ u & v & w \end{pmatrix} < 2$$

outside the fundamental points. However, over the fundamental points p_0, p_1, p_2 we need additional equations. If J denotes the ideal of minors of the matrix above then $I = J : \langle xy, xz, yz \rangle^\infty$ is the defining ideal of the graph. We claim $I = \langle vy - ux, ux - wz \rangle$, which we can easily verify with computer algebra. The inclusion $I \supset \langle vy - ux, ux - wz \rangle$, which is all we need in the following, follows from the identiy

$$(vy - ux) \begin{pmatrix} yz \\ xz \\ xy \end{pmatrix} = (vyz - uxz, wyz - uxy, wxz - vxy) \begin{pmatrix} y & x & 0 \\ 0 & 0 & x \\ 0 & 0 & -y \end{pmatrix}$$

and a similar calculation for the second generator.

Now restistricted to the open set $G_{21} = G \cap \{z = 1, v = 1\} = G \cap (U_2 \times U_1)$, we obtain $\Bbbk[G_{21}] \cong \Bbbk[x, y, u, w]/(y - ux, w - ux) \cong \Bbbk[x, u]$ and the projection onto the first factor is

$$\mathbb{A}^2 \cong G_{21} \to U_2 \cong \mathbb{A}^2,\ (x, u) \mapsto (x, ux).$$

This is the chart of a blow-up with exceptional curve defined by $x = 0$. The projection onto the second factor

$$\mathbb{A}^2 \cong G_{21} \to U_1 \cong \mathbb{A}^2,\ (x, u) \mapsto (u, ux).$$

is similarly a chart of a blow-up with exceptional curve this time defined by $u = 0$, which is the strict transform under the first projection of the line $V(y) \subset U_2$ to G_{21}, because $y = ux$ on G_{21}. Since the sets $G_{ij} = G \cap (U_i \times U_j)$ for $i \neq j$ cover G the proposition follows. □

Corollary 7.2.20. *The graph of the quadratic transformation is the intersection*

$$G = (\mathbb{P}^2 \times \mathbb{P}^2) \cap \mathbb{P}^6 \subset \mathbb{P}^8$$

of a codimension 2 linear subspace with the Segre product.

Proof. Indeed, G is defined by two linear equations in the Segre embedding of $\mathbb{P}^2 \times \mathbb{P}^2 \subset \mathbb{P}^8$. In suitable coordinates $z_0, \ldots, z_6$ of $\mathbb{P}^6$ the homogeneous ideal is given by the 2×2 minors of the matrix

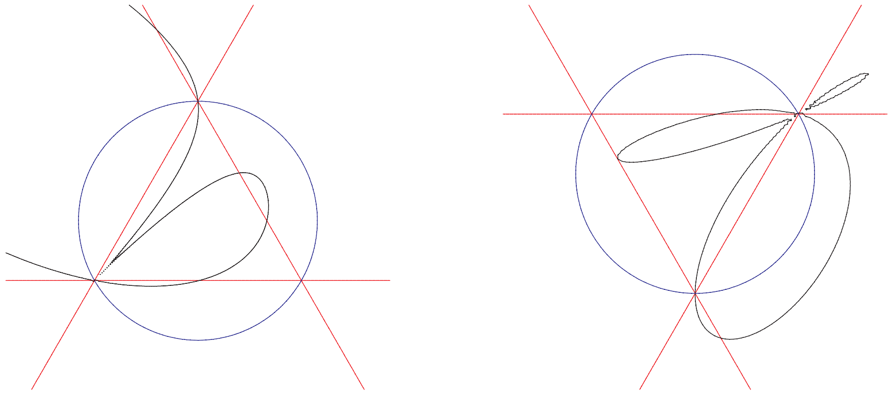
$$\begin{pmatrix} z_0 & z_1 & z_2 \\ z_3 & z_0 & z_4 \\ z_5 & z_6 & z_0 \end{pmatrix}.$$

□

Exercise 7.2.21. Let denote $p_0 = [1:0:0], p_1 = [0:1:0] \in \mathbb{P}^2$. The linear system $L(2; p_0, p_1) = \langle xy, xz, yz, z^2 \rangle$ defines a birational map $\varphi : \mathbb{P}^2 \dashrightarrow \mathbb{P}^1 \times \mathbb{P}^1 \subset \mathbb{P}^3$. Describe the graph of φ and the exceptional curves of both projections. ****** fix ******add picture of the projection with all the lines! □

If we want to study the strict transform of an affine plane curve in a blow-up of the origin we can equivalently study the quadratic transformation with one of the fundamental points in the origin of the corresponding projective curve. If the fundamental lines are not tangent to C, then the intersection points with the exceptional curve correspond to the intersection points of the transform $q(C)$ with the corresponding new line in the image.

How does C change under the transformation?



****** fix ******add equations

Note, that the circle in one side is the strict transforms of the line at infinity of the visible chart on the other side. The original curve has an non-ordinary triple point, the strict transform has an ordinary triple point.

Proposition 7.2.22. *Let C be a plane curve of degree d which has multiplicity r_0, r_1, r_2 in the fundamental points p_0, p_1, p_2 of a quadratic transformation q. Then $q(C)$ has degree $2d - r_0 - r_1 - r_2$ and three new singular points with multiplicity $d - r_1 - r_2$, $d - r_0 - r_2$ and $d - r_0 - r_1$.*

Proof. The preimage $q^{-1}(L)$ of a general line is a conic through the 3 fundamental points with no common tangent with C in these points. So $\deg C' = 2d - r_0 - r_1 - r_2$. The singular points in the fundamental points of q^{-1} have multiplicity $d - r_i - r_j$, because this is the intersection number of the strict transform of $\overline{p_i p_j}$ with C in G. □

Theorem 7.2.23 (Cremona resolution). $\mathbb{k} = \overline{\mathbb{k}}$. *Every irreducible plane curve can be transformed by a sequence of quadratic transformations into a plane curve with only ordinary singularities.*

Proof. Let C be an irreducible plane curve of degree d. Then by the bound on the number of singular points 5.4.12, the difference

$$\binom{d-1}{2} - \sum_{p \in C} \binom{r_p}{2}$$

is non-negative. We compute how this expression changes under a suitable quadratic transformation. Let $p_0 \in C$ be a singular point of multiplicity r. We like to choose $p_1, p_2 \notin C$ such that the fundamental lines are not tangent to C and that the fundamental lines intersect C in smooth points outside p_0. This is possible over a field of char($\mathbb{k}$) $= 0$, because not every line through p_0 is tangent to C by Exercise 6.6.7. If char($\mathbb{k}$) $\neq 0$ then possibly every line through p_0 is a tangent line to C, see Remark 6.6.8, and we might need an additional quadratic tranformation, as we will explain at the end of the proof. Assume now, that we can find p_1 and p_2 satisfying the desired requirement for the fundamental triangle p_0, p_1, p_2. Then the strict transform $C' = q(C)$ has degree $d' = 2d - r$ and three new ordinary singular points of multiplicities d, $d - r$ and $d - r$. Since

$$\binom{2d-r-1}{2} - \binom{d}{2} - 2\binom{d-r}{2} = \binom{d-1}{2} - \binom{r}{2},$$

we have

$$\binom{d'-1}{2} - \sum_{p' \in C'} \binom{r'_p}{2} \leq \binom{d-1}{2} - \sum_{p \in C} \binom{r_p}{2},$$

because the contribution of the singular points outside the fundamental triangles is unchanged. The inequality is strict, if the strict transform of C in the blow-up of $\mathbb{P}^2$ at p_0 is not smooth at some points of the exceptional curve E. Thus such quadratic transformations based at some non-ordinary point either decreases the difference above (in which case the number of non-ordinary points might increase), or decreases the number of non-ordinary points. Since both numbers are non-negative, we obtain a curve with only ordinary singularities after a finite number of transformations. This completes the proof over a field of characteristic zero.

If $\mathrm{char}(\Bbbk) \neq 0$, we might need additional steps. Suppose that everline through p_0 is tangent to C somewhere. Then the dual curve of C is the line dual to p_0. We claim, that after a general quadratic transformation q based at points q_0, q_1, q_2 outside C, the strict transform C' of C, which now has degree $2d$ and three more ordinary singular points of multiplicity d, has $q(p_0) \in C'$ no longer as a "strange point". A line through $q(p_0)$ corresponds to a conic D through p_0 and q_0, q_1, q_2. We choose such a nowhere tangent conic D through p_0 first and then $q_0, q_1, q_2 \in D$. Consider the blow-up of $\mathbb{P}^2$ at p_0. That is the ruled surface F of Example 7.2.8. A smooth conic through p_0, which is nowhere tangent to C, corresponds to a transversal hyperplane section of the strict transform $C'' \subset F \subset \mathbb{P}^4$ of C in F. Such hyperplane exists by Bertini's Theorem 6.6.1. Pick such D. Furthermore, pick a general line L transversal to $C \cup D$. The two intersection points q_1, q_2 of $D \cap L$ are not strange points of C, because C has at most one strange point. So only finitely many lines through q_1 respectively q_2 are tangent to C, and we can choose one of the infinitely many points $q_0 \in D$, such that $\overline{q_0 q_1}$ and $\overline{q_0 q_2}$ are neither tangent nor pass through a singular point of C. The quadratic transformation based on q_0, q_1, q_2 satisfies the assertion that $q(p_0)$ is no longer a strange point, because D is not tangent. The number of non-ordinary singular points and the difference

$$\binom{d-1}{2} - \sum_{p \in C} \binom{r_p}{2}$$

remains unchanged under q. Indeed

$$\binom{d-1}{2} = \binom{2d-1}{2} - 3\binom{d}{2}.$$

Thus, the induction goes through also for fields of characteristic $\neq 0$. □

Proof of the resolution of singularities 7.2.13. We first treat the case of an irreducible plane curve. The key is to compare the blow-up of $\mathbb{P}^2$ at a point p_0 with a quadrtic tranformation q based on a triangle p_0, p_1, p_2, which includes p_0. By Proposition 7.2.19, points on the exceptional curve E of the blow-up of $\mathbb{P}^2$ at p_0 correspond to the points on one of the fundamental lines in the image of q with the exceptions of the two points on E, which correspond to the direction of two fundamental lines $\overline{p_0 p_1}$ and $\overline{p_0 p_2}$. If the two lines $\overline{p_0 p_1}$ and $\overline{p_0 p_2}$ are not tangent to C at p_0 then every singularity of the strict transform C' of C in the blow-up is still visible under the quadratic tranformation. This is satisfied in the Cremona resolution process. So we can keep track about what happens at the succesive exceptional curves in a resolution process by considering the corresponding fundamental lines in a Cremona resolution. The additional quadratic transformation possibly needed, if $\mathrm{char}(\Bbbk) \neq 0$, do not matter to us, because they only introduce ordinary singularities away from our original singularities. By the theorem, we reach a situation where the transformed curve has only ordinary singular points, and one more blow-up for each remaining singular point will resolve these singularities. This proves the resolution of singularities for irreducible plane curves.

If we have a reduced curve consisting of several components, then by the preceeding result we may assume that all components are smooth after some blow-ups. What remains, is to prove that two smooth curves C_1, C_2, which intersect in a point p, can be separated by a sequence of blow-ups. We leave it as an easy exercise to prove that two smooth curves which intersect with multiplicity $i(C_1, C_2, p) = m$ get separated after precisely m blow-ups. □

Exercise 7.2.24. Let C, D be two plane curve in $\mathbb{A}^2$, which are smooth in o. Then $\imath(C, D; o) = m$ iff the strict transfoms $C^{(k)}$ and $D^{(k)}$ in $X^{(k)}$ intersect exceptional curve $E^{(k)}$ of σ_k in the same point for $k < m$ and in different points for $k = m$, where

$$X^{(k)} \xrightarrow{\sigma_k} X^{(k-1)} \qquad \text{for } 1 \leq k \leq m$$

denotes the blow up of the $X^{(k-1)}$ in the common point of $C^{(k-1)}$ and $D^{(k-1)}$ on $E^{(k-1)}$ for $k \geq 2$, and σ_1 denotes the blow up of the point $o \in X^{(0)} = \mathbb{A}^2$.

Corollary 7.2.25. $\Bbbk = \overline{\Bbbk}$. *The assertion of Theorem 7.2.23 holds for reduced plane curves.* □

Theorem 7.2.26 (Desingularization). *Let C be a projective curve. There exists a birational morphism*

$$\eta\colon \widetilde{C} \to C$$

from a smooth projective curve $\widetilde{C}$. The smooth curve $\widetilde{C}$ is unique up to an unique isomorphism.

We call $\eta : \widetilde{C} \to C$ the **desingularization** of C and $\widetilde{C}$ a **smooth model** of the function field $\Bbbk(C)$.

Proof. We first assume that C is irreducible. By Theorem 3.5.2 C has birational map $\pi\colon C \dashrightarrow C' \subset \mathbb{P}^2$ onto a plane curve C'. Let $\sigma\colon \widetilde{C} \to C'$ be an embedded resolution of singularities 7.2.13. By Theorem 7.1.12, the birational map $\pi^{-1} \circ \sigma\colon \widetilde{C} \dashrightarrow C$ extends to a morphism $\eta\colon \widetilde{C} \to C$ since C is projective. For uniqueness, given two resolution of $\eta_i\colon \widetilde{C}_i \to C$ the morphism $\varphi\colon C_1 \to C_2$ is the extension of the rational map $\eta_2^{-1} \circ \eta_1 : C_1 \dashrightarrow C_2$ according to Theorem 7.1.12. Hence this is the unique isomorphism which makes the diagram

$$\begin{tikzcd}
\widetilde{C}_1 \arrow[rr, "\varphi"] \arrow[dr, "\eta_1"'] & & \widetilde{C}_2 \arrow[dl, "\eta_2"] \\
& C &
\end{tikzcd}$$

commutative. Finally, if C is reducible, say $C = C_1 \cup \ldots \cup C_r$ with irreducible components C_i, then the disjoint union of their desingularizations $\widetilde{C} = \widetilde{C}_1 \cup \ldots \cup \widetilde{C}_r$ with $\eta\colon \widetilde{C} \to C$ defined by $\eta \mid_{\widetilde{C}_i} : \widetilde{C}_i \xrightarrow{\eta_i} C_i \hookrightarrow C$ is the unique resolution of singularities. □

Corollary 7.2.27. *Every curve is birational to a smooth projective curve.* □

Exercise 7.2.28. Prove that the smooth model and the birational morphism can be defined over the field of definition of the curve. □

Definition 7.2.29. Let C be a smooth projective curve and $C' \subset \mathbb{P}^2$ a plane curve birational to C of degree d with only ordinary multiple points. Then

$$g := \binom{d-1}{2} - \sum_{p \in C'} \binom{r_p}{2}$$

is called the **geometric genus** of C.

We will see later in Proposition 8.2.1, that the geometric genus of C is well defined, i.e. independent of the choice of C'. To make the definition precise at this point, we should speak of g as the geometric genus of the plane model C'.

By Proposition 5.4.13 a curve with geometric genus $g = 0$ can be parametrized by $\mathbb{P}^1$ (in case of an algebraically closed field).

Remark 7.2.30. Desingularization of higher dimensional varieties is much more involved. The existence of a desingularization for varieties of any dimension was proved by Hironaka **?** over fields of characteristic $\operatorname{char}(\Bbbk) = 0$. His precise result is that for any variety X there exists a birational morphism $\widetilde{X} \to X$ from a smooth variety $\widetilde{X}$ such that the restriction $\eta : \widetilde{X} \setminus \eta^{-1}(\operatorname{sing} X) \to X \setminus \operatorname{sing} X$ is an isomorphism. Resolution of singularities in positive characteristic is known for dimension $\dim \leq 3$ and but open otherwise.

The first Computer Algebra implementation of Hironaka's algorithm is due to Bodnar and Schicho **?**. A `SINGULAR` implementation was given by Frühbis-Krüger and Pfister **?** . These implementation are based on Villarmajor's **?** variant of Hironaka's algorithm.

For some application of the resolution of singularities the theory of alterations due to Johan de Jong **??** is sufficient: There always exists a dominant surjective morphism $\eta : \widetilde{X} \to X$ from a smooth variety $\widetilde{X}$ of the same dimension as X. In other words, for a function field $\Bbbk(X)$ of any characteristic there exists a finite field extension $\Bbbk(\widetilde{X}) \supset \Bbbk(X)$, which has a smooth projective model.

Another feature that is different in higher dimension, is that there exist no unique smooth model of a function field. We have seen that $\mathbb{P}^2$ and the cubic scroll F from Example 7.2.8 are both smooth projective models of $\Bbbk(x_1, x_2)$.

Indeed in dimension dim ≥ 2 we can always blow up a smooth point (or a smooth subvariety of codim ≥ 2) to obtain a larger smooth model than a given one. Castelnuovo **?** proved that a surface has a smooth minimal model, i.e. a smooth model, which does not arise as the blow up of a smooth surface, and that this model is unique unless the the surface is birational to $\mathbb{P}^2$ or to $C \times \mathbb{P}^1$, see Beauville **?** or Barth-Hulek-Peters-van de Ven **?** .

For higher dimensional varieties starting with the work of Mori **?** , the search for suitable minimal models, known under the heading minimal model program, is an important area of current research in birational geometry. After an intense period of work on this program Birkar, Cascini, Hacon and McKernan **?** finally established the existence of minimal models. The minimal model program is main tool to extend the so called Enriques classification of surfaces to higher dimensions. All these topics are far beyond the scope of this book.

7.3 Divisors and Rational Maps on Curves

Let $C \subset \mathbb{P}^n$ be a smooth projective curve. We want to describe morphisms $C \to \mathbb{P}^r$. Key ingredient in our description is the concept of divisors.

Definition 7.3.1. Let C be a smooth projective curve. A **divisor** on C is a formal finite sum

$$D = \sum_{j=1}^{s} n_i p_i \,,$$

where the $n_i \in \mathbb{Z}$ are integers and the p_i are points of C. The set of divisors

$$\mathrm{Div}(C) = \{D = \sum_{p \in C} n_p p \mid \text{ all but finitely many } n_p = 0\}$$

is an abelian group, the free abelian group generated by the points of C. The **degree of a divisor** $D = \sum n_p p$ is $\deg D = \sum n_p$. This make sense, because only finitely many n_p are non-zero for a given divisor.

$$\deg\colon \mathrm{Div}(C) \to \mathbb{Z}$$

is a group homorphism. The **support** of a divisor $D = \sum n_p p$ is the set $\mathrm{supp}(D) = \{p \in C \mid n_p \neq 0\}$

Example 7.3.2. Let $f \in \Bbbk(C) \setminus \{0\}$ be a rational function. Then the **divisor of zeroes and poles** of f is

$$(f) = \sum_{p \in C} v_p(f) p \,.$$

Divisors of rational functions are called **principal divisors**.

For a smooth curve $C \subset \mathbb{P}^n$ and a homogeneous polynomial $h \in \Bbbk[x_0, \ldots, x_n]$ which does not vanish on any component of C, we define the intersection **intersection divisor** is

$$C.h = \sum_{p \in C \cap H} \imath(C, h; p)p,$$

where $\imath(C, h, p) = \operatorname{length} \mathcal{O}_{C,p}/h_p = \dim_{\Bbbk} \mathcal{O}_{\mathbb{P}^n,p}/(I(C) + h)\mathcal{O}_{\mathbb{P}^n,p}$ denotes the intersection multiplicity. Here $(I(C) + h)\mathcal{O}_{\mathbb{P}^n,p}$ denotes the ideal generate by dehomogenization of $I(C)$ and h in any affine chart of $\mathbb{P}^n$ which contains p. If h is the defining equation of an hypersurface H then we write also

$$C.H = C.h$$

Remark 7.3.3. Note, that

$$\deg(C.h) = \deg C \cdot \deg h$$

holds by Bézout's Theorem 6.4.33.

Since every rational function f on C is quotient of homogeneous polynomials of the same degree $f = g/h$, every principal divisor is difference of two intersection divisors:

$$(f) = C.g - C.h.$$

Hence principal divisors have degree 0. In particular, a rational function on a smooth projective curve has as many poles as zeroes counted with multiplicities.

Remark 7.3.4. The concept of divisors plays also a crucial role in higher dimension. The group of **Weil divisors** on a variety X of higher dimension is the free abelian group generated by codimension 1 subvarieties. We will make little use of this concept. However it is convienent to associate to an homogeneous polynomial $h \in \Bbbk[x_0, \ldots, x_n]$ the divisor

$$H = \mu_1 H_1 + \ldots + \mu_s H_s \in \operatorname{Div}(\mathbb{P}^n)$$

in case that $h = u h_1^{\mu_1} \cdot \ldots \cdot h_s^{\mu_s}$ is factorization in irreducibles and $H_k = \mathrm{V}(h_k)$ the corresponding subvarieties.

Since for a smooth curve all local rings $\mathcal{O}_{C,p}$ are DVRs we have the identiy

$$C.h = \mu_1 C.H_1 + \ldots + \mu_s C.H_s.$$

It is natural to define $C.H$ for a divisor $H \in \operatorname{Div}(\mathbb{P}^n)$ as the right hand side of this expression.

Example 7.3.5. Generalizing the intersection divisor, we define the **pull back divisor** $\varphi^* H$ as follows: For a morphism $\varphi \colon C \to \mathbb{P}^r$ and a divisor $H = \mu_1 H_1 + \ldots + \mu_s H_s \in \operatorname{Div}(\mathbb{P}^r)$ defined by a homogeneous polynomial h, with no component of H containing the image of a component of C, the **pull back divisor** $\varphi^* H$ as

$$\varphi^* H := \sum v_p(h)p = \sum_{k=1}^{s} \mu_k \varphi^* H_k,$$

where we regard h as a function in a neighborhood of $p \in C$ by dehomogenizing $h \mapsto h_{\varphi(p)}$ in an affine neighborhood of $\varphi(p) \in \mathbb{P}^n$ and pullback $h_{\varphi(p)} \mapsto \varphi^*(h_{\varphi(p)}) = h_{\varphi(p)} \circ \varphi$. Since dehomogenization is well-defined up to a unit in $\mathcal{O}_{\varphi(p)}$, the vanishing order $v_p(\varphi^* h_p)$ is independent of the choice of $h_{\varphi(p)}$.

Thus, $\varphi^* H$ generalizes the intersection divisor $C.H$ for $C, H \subset \mathbb{P}^n$.

Definition 7.3.6. Two divisors $D, D' \in \operatorname{Div}(C)$ are **linearly equivalent**, $D \equiv D'$, if there is a function $f \in \Bbbk(C)$, such that $D - D' = (f)$.

Example 7.3.7. If $C \subset \mathbb{P}^n$ and g, h are homogeneous forms on $\mathbb{P}^n$ of the same degree not containing C then

$$C.g \equiv C.h\,.$$

Indeed, we have $(g/h) = C.g - C.h\,.$

Definition 7.3.8. A divisor $D = \sum n_p p$ is **effective**, $D \geq 0$, if all $n_p \geq 0$. If $D \in \operatorname{Div}(C)$ is a divisor then the **complete linear system of divisors** linear equivalent to D is

$$|D| = \{D' \in \operatorname{Div}(C) \mid D' \geq 0 \text{ and } D \equiv D'\}.$$

We will see, that either $|D|$ is empty or that $|D|$ carries the structure of a projective space.

Definition 7.3.9. Let $D \in \operatorname{Div}(C)$. We call the vector space

$$L(D) = \{f \in \Bbbk(C)^* \mid (f) + D \geq 0\} \cup \{0\}$$

the space of rational functions with **poles up to order** D, or the **Riemann-Roch space** of D.

$$\ell(D) = \dim L(D)$$

denotes its dimension.

The name Riemann-Roch space is motivated by the prominent role of the Theorem of Riemann-Roch 8.3.2 in the theory of these function spaces. If $D = \sum_{p \in C} n_p p$ is not effective, say $n_q < 0$ for some point $q \in C$ then functions $f \in L(D)$ have a zero of order at least n_q at q. So the name function with poles up to order D is a bit misleading. Since by Corollary **??** we may assume for most purposes that D is effective, we stick to this abuse of notation.

Lemma 7.3.10. *Let C be a smooth irreducible projective curve and let $D \in \operatorname{Div}(C)$ be a divisor on C. i) If $p \in C$ is a point, then $L(D-p) \subset L(D)$ is a subspace of codimension at most 1.*
ii) $L(D) = 0$ if $\deg D < 0$. In case $\deg D = 0$ the Riemann-Roch space $L(D)$ is one-dimensional iff D is a principal divisor.
iii) $\ell(D) \leq \deg D + 1$

Proof. i) If $p \notin \operatorname{supp}(D)$ then $f \in L(D)$ lies in $L(D-p)$ iff f vanishes at p. Since $f(p) = 0$ gives one linear equation on $L(D)$ the result follows. In case $p \in \operatorname{supp}(D)$, the argument is similar. One more coefficient in the Laurent expansion of f at p has to vanish.
ii) Since an effective divisor have degree ≥ 0, an $f \in L(D) \setminus \{0\}$ gives $\deg D = \deg(f) + \deg D = \deg((f)+D) \geq 0$. If equality holds then $(f) + D = 0$, hence $D = (1/f)$.
iii) Suppose $L(D) \ni f \neq 0$. If we pick a point $p \notin \operatorname{supp}(D)$, where f does not vanish then $f \notin L(D-p)$. Hence, $\ell(D-p) = \ell(D) - 1$. We can continue substracting points until we reach $\ell(D - p_1 - \ldots - p_r) = 1$ after $r = \ell(D) - 1$ steps. Then by ii) $\deg D - r \geq 0$, i.e, $\ell(D) - 1 \leq \deg D$. □

Corollary 7.3.11. *$\mathbb{P}(L(D)) \cong |D|$ via $f \mapsto (f) + D$.*

Proof. We apply *ii*). □

Corollary 7.3.12. *If $D \equiv D'$ then $L(D) \cong L(D')$ via $f \mapsto fg$ where $g \in K(C)$ is the rational function with $(g) = D' - D$. The function g and the isomorphism is unique up to a scalar factor $\lambda \in \Bbbk^*$.*

Proof. We apply *ii*). □

Definition 7.3.13. A **linear system of divisors** on a smooth projective curve is a linear subspace $P \subset |D|$ of a complete linear system. If a point p is common to all divisors of a linear system P then p is called a **base point** of P. A linear system without base points is called **base point free**.

Example 7.3.14. A complete linear system $|D|$ is base point free iff $L(D-p) \subset L(D)$ has codimension 1 for every point $p \in C$.

Theorem 7.3.15. *Let C be a smooth projective curve. There is a bijection*

$$\left\{ \begin{matrix} \textit{base point free} \\ \textit{linear systems} \\ \textit{of dimension } r \end{matrix} \right\} \longleftrightarrow \left\{ \begin{matrix} \textit{non-degenerate} \\ \textit{morphisms} \\ \varphi\colon C \to \mathbb{P}^r \end{matrix} \right\} / \operatorname{PGL}(r+1)$$

Proof. Given φ, then $\varphi^* \mathrm{V}(x)$, where x is a linear form on $\mathbb{P}^r$, gives an r-dimensional system without base points, because $\mathrm{V}(x_0, \ldots, x_r) = \emptyset$.

Conversely, given a linear system $P \subset |D|$, say with $D \in P$, we can consider the subspace $F \subset L(D)$ with

$$F = \{f \in L(D) \mid (f) + D \in P\}$$

Then $1 \in F$ and if $1, f_1, \dots, f_r$ is a basis, then the rational map given by this tuple gives a morphism defined at all points outside the support of D. If we extend across these points, we see that $D - \varphi^* \mathrm{V}(x_0)$ is precisely the divisor B of base points of the system P. The base point free system of this morphism is $P - B = \{D - B \mid D \in P\}$. The result follows. □

Corollary 7.3.16. *Given a non-degenerate projective curve $C \subset \mathbb{P}^r$, there exists an n_0 such that for $n \geq n_0$ every curve $C' \subset \mathbb{P}^n$, which can be projected birationally onto C from linear subspace of dimension $n - r - 1$ disjoint from C', is degenerate.*

Proof. Consider a desingularization $\eta \colon \tilde{C} \to C$ and $D = \eta^* H$ with $H = \mathrm{V}(x_0)$ a hyperplane. Then $n_0 = \ell(D)$ has this property. □

Thus to describe morphisms $C \to \mathbb{P}^r$, we first consider complete linear systems $|D|$ of divisors and their function spaces $L(D)$. The second step amounts to study projections from subspaces.

We denote by

$$\varphi_D \colon C \to \mathbb{P}^r$$

the morphism corresponding to a complete linear system $|D|$.

Note, that $\varphi_D = \varphi_{D-B}$ where B denotes the divisor of base points of $|D|$.

Theorem 7.3.17. *Let $|D|$ be a base point free complete linear system of divisors of dimension r. $\varphi_D \colon C \to \mathbb{P}^r$ is an embedding iff $L(D - p - q) \subset L(D)$ has codimension 2 for every pairs of points $p, q \in C$.*

Definition 7.3.18. A divisor D (or better the divisor class of D) is called **very ample** if D satisfies the equivalent conditions of the theorem.

Proof. Since $|D|$ is base point free, every $L(D - p) \subset L(D)$ has codimension 1. If $p \neq q$ and $L(D - p - q) \subset L(D)$ has not codimension 2 iff $L(D - p) = L(D - p - q)$, which means that p and q have the same image. Thus codimension 2 for all $p, q \in C$ with $p \neq q$ is equivalent to $\varphi_D \colon C \to \mathbb{P}^r$ being injective. However, injectivity does not imply, that C is isomorphic to the image.

Example 7.3.19. The map

$$\mathbb{P}^1 \to \mathbb{P}^2, \quad [1 : t] \mapsto [1 : t^2 : t^3]$$

is an injection but the image $\mathrm{V}(xz^2 - y^3)$ is not isomorphic to $\mathbb{P}^1$, due to the cusp.

Our assumption is $L(D - p - q) \subset L(D)$ has codimension 2 for every $p, q \in C$ including the case $p = q$.

To conclude, that for an injective morphism $C \to C' = \varphi(C) \subset \mathbb{P}^r$, the inverse map $\varphi^{-1}\colon C' \to C$ is a morphism, we need that $\varphi^*\colon \mathcal{O}_{C',\varphi(p)} \to \mathcal{O}_p$ is surjective. (Injective is clear by definition of the image.)

For this we consider a rational functions $f_0 \in L(D) \setminus L(D-p)$ and $f_1 \in L(D-p) \setminus L(D-2p)$ and suitable homogeneous coordinates $[x_0 : x_1 : \ldots : x_r]$ on $\mathbb{P}^n$. Since $\varphi\colon C \to C'$ is dominant and injective we have $\Bbbk(C') \cong \Bbbk(C)$ and we may regard f_0, f_1 both as rational function both on C and C'. The quotient $t = f_1/f_0 = \varphi^*(x_1/x_0)$ is an element of $\mathfrak{m}_{C',\varphi(p)} \subset \mathcal{O}_{C',\varphi(p)} \subset \mathcal{O}_{C,p}$ because in our coordinates $x_1/x_0 \in \mathfrak{m}_{\mathbb{P}^n,\varphi(p)} \subset \mathcal{O}_{\mathbb{P}^n,\varphi(p)}$. On the other hand, t is a generator of the maximal ideal $\mathfrak{m}_p \subset \mathcal{O}_{C,p}$.

Since $C \to C'$ is a finite morphism, $\mathcal{O}_{C,p}$ is a finite $\mathcal{O}_{C',\varphi(p)}$-module by Theorem **??**. We compare generators of these modules. We have $\mathfrak{m}_{C',\varphi(p)}\mathcal{O}_{C,p} = \mathfrak{m}_p$, because $t \in \mathfrak{m}_{C',\varphi(p)} \subset \langle t \rangle = \mathfrak{m}_p$ Hence

$$\Bbbk \cong \mathcal{O}_{C',\varphi(p)}/\mathfrak{m}_{C',\varphi(p)} \to \mathcal{O}_{C,p}/\mathfrak{m}_{C',\varphi(p)}\mathcal{O}_{C,p} \cong \mathcal{O}_{C,p}/\mathfrak{m}_p \cong \Bbbk$$

is an isomorphism. By Nakayama's Lemma 4.2.19, the inclusion $\mathcal{O}_{C',\varphi(p)} \hookrightarrow \mathcal{O}_{C,p}$ is an isomorphic of $\mathcal{O}_{C',\varphi(p)}$-modules as well. Hence we have equality.

Conversely, if φ is an embedding, then the image is smooth at p and we can find a hyperplane which passes through p but is not tangent. The pullback of its equation gives a divisor $D' \in |D|$ with $n_p = 1$ and $L(D-2p) \subsetneq L(D-p)$ follows. □

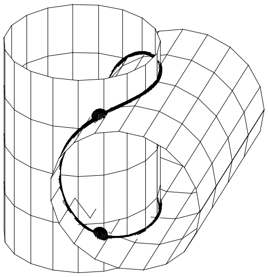
7.4 Riemann's Inequality

Let $C \subset \mathbb{P}^n$ be a smooth projective curve and let $D = \sum n_i p_i \in \operatorname{Div}(C)$ an effective divisor of degree d. We want to compute the complete linear series $|D|$, or equivalently, the space $L(D)$. The idea is simple.

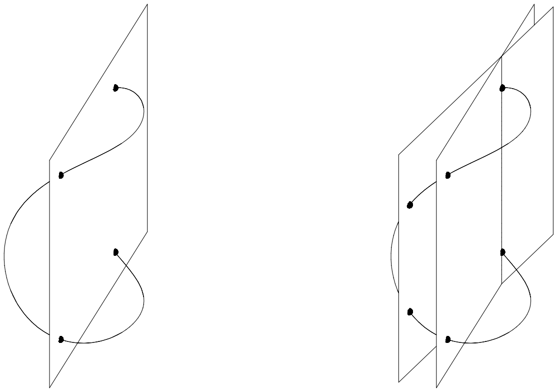
1. Let H be a hypersurface, which does not contain C, but passes through the points p_i with intersection multiplicity $\imath(C, H; p_i) \geq n_i$.
2. Condider the residual part $E = H.C - D$. E is an effective divisor on C of degree $\deg C \cdot \deg H - d$.
3. Then any hypersurface H', which does not contains C but intersects C in E, more precisely with $H'.C \geq E$, gives a divisor $D' = H'.C - E$, which is linearly equivalent to D, because the quotiont of the equations $h'/h \in L(D) \subset \Bbbk(C)$.

Example 7.4.1. Consider the smooth projective space curve C, defined by the affine equations

$$x^2 + y^2 - 1 = (x-1)^1 + z^2 - 1 = 0.$$



We compute divisors linearly equivalent to $D = p_1 + p_2$, where the points have affine coordinates $p_{1,2} = (\frac{1}{2}, -\frac{1}{2}\sqrt{(3)}, \pm\frac{1}{2}\sqrt{(3)})$. C has degree 4. The hyperplane with the affine equations $H = x - \frac{1}{2}$ passes D. The residual part of the intersection $C \cap H$ is $E = p_3 + p_4$ with $p_{3,4} = (\frac{1}{2}, \frac{1}{2}\sqrt{(3)}, \pm\frac{1}{2}\sqrt{(3)})$



Thus, 1 and $\frac{y+\frac{1}{2}\sqrt{(3)}}{x-\frac{1}{2}}$ are elements of $L(D)$. The proof of the Theorem 7.4.2 below will give, that this is a basis of $L(D)$ in particular $\ell(D) = 2$.

In general, it is not true that this approach will give the complete linear series. However, if we choose the hypersurface H of degree larger than a constant m_0, which depend only on C, the methods works. A key fact is the following theorem. For a linear form x not vanishing on (any component of) C, we denote (abusing of notation) $H = C.x \in Div(C)$ the intersection divisor, and speak of the hyperplane class of $C \subset \mathbb{P}^r$.

Theorem 7.4.2 (Completeness of hypersurface systems). *Let $C \subset \mathbb{P}^r$ be a smooth irreducible projective curve, and let $H \in Div(C)$ be a divisor representing the hyperplane class. Then, there exists a constant m_0, such that for $m \geq m_0$ the complete linear system $|mH|$ is cut out by hypersurfaces of degree m.*

Proof. We may assume, that $C \subset \mathbb{P}^r$ spans $\mathbb{P}^r$. Then $H_i = C.x_i$ is a divisor in $|\ H\ |$ for each coordinate function x_i. Let $D \in |mH|$. Then $D \equiv mH_i$, which means that there exist a rational function $f_i \in \Bbbk(C)$, such that

$$(f_i) = D - mH_i.$$

The fractions f_i/f_j satisfy

$$(f_i/f_j) = D - mH_i - (D - mH_j) = m(H_j - H_i) = (x_j^m/x_i^m).$$

Changing the f_i by suitable constants, this gives

$$f_i/f_j = x_j^m/x_i^m \quad \text{for all } i, j.$$

i.e. $f = x_i^m f_i$ is a well defined homogeneous element in the field of fraction of the homogeneous coordinate ring $R = \Bbbk[x_0, \dots, x_n]/I_C$ of C. Since D is effective, f_i is regular on $C \setminus \mathrm{V}(x_i)$. Thus, by Proposition 2.6.15, f_i lies in the affine coordinate ring of $C \cap U_i$ and $f_i = F_i/x_i^N$ for some homogeneous element $F_i \in \Bbbk[x_0, \dots, x_n]/I_C$ of degree N. We may assume, that N is the same for each i. Thus,

$$x_i^N f = x_i^{N+m} f_i = x_i^m F_i \in \Bbbk[x_0, \dots, x_n]/I_C.$$

The following proposition finishes the proof.

Proposition 7.4.3. *Let $R = \Bbbk[x_0, \dots, x_n]/\mathfrak{p}$ be the homogeneous coordinate ring of a variety. There exists an m_0 such that for $m \geq m_0$*

$$\left\{ \begin{matrix} f \in Q(R) \\ x_0^N f, \dots, x_n^N f \in R_{N+m} \end{matrix} \right\} \implies f \in R_m$$

holds.

move * the following defn**

Definition 7.4.4. An element $m \in M$ in a module over some ring A is a torsion element if $am = 0$ for some non zero divisor $a \in A$. M is torsion free, if $0 \in M$ is the only torsion element.

We actually proof a more general result. Consider a Noether normalization $\Bbbk[y_0, \dots, y_r] \hookrightarrow R = \Bbbk[x_0, \dots, x_n]/\mathfrak{p}$ induced by linear forms y_j. We may assume that the y_i are among the coordinates $x_0, \dots, x_n$ This makes R into a finitely generated graded torsion free $\Bbbk[y_0, \dots, y_r]$-module.

Our proof works for torsion free $S = \Bbbk[y_0, \dots, y_r]$-modules. Let M be a graded finitely generated torsion free module over $S = \Bbbk[y_0, \dots, y_r]$, and let

$$M_{Q(S)} = M[U^{-1}]$$

denote the localization of M in $U = S \setminus \{0\}$. Then $M_{Q(S)}$ is a vector space over the quotient field $Q(S)$, and

$$M \hookrightarrow M_{Q(S)},\, g \mapsto \frac{g}{1}$$

is injective, because M is torsion free.

Our element f is a homogeneous element of $R_{Q(S)}$, because $f = f_i/y_i^N = F_i/y_i^{N-m}$. Thus, Proposition 7.4.3 follows from the following more general result.

Proposition 7.4.5. *Let M be a finitely generated torsion free graded $S = \Bbbk[y_0, \ldots, y_r]$-module. There exists an m_0, such that for $m \geq m_0$*

$$\left\{\begin{matrix} f \in M_{Q(S)} \text{ and} \\ \exists\, N \text{ with } y_0^N f, \ldots, y_r^N f \in M_{N+m} \end{matrix}\right\} \implies f \in M_m$$

holds.

Proof. We first show, that M is a submodule of a finitely generated free S-module. Let $g_1, \ldots, g_s$ be homogeneous generators of M. Then the elements $g_j = g_j/1 \in M_{Q(S)}$ generate $M_{Q(S)}$. Thus, we may assume that $g_1, \ldots, g_t$ for some $t \leq s$ form a basis of $M_{Q(S)}$. For the remaining ones, we have expressions $b_i g_i = \sum_{j \leq t} a_{ij} g_j$ with homogenous elements $b_i, a_{ij} \in S$.

Set $b = \prod_{i=t+1}^{s} b_i$. The elements $g_j' = g_j/b \in M_{Q(S)}$ generate a graded free submodule M' with

$$M \subset M' = \oplus_{j=1}^{t} S g_j' \subset M_{Q(R)}.$$

As an intermediate step we claim

Claim. $f \in M_{Q(S)}$ and $y_i^N f \in M \,\forall i \implies f \in M'$

Indeed, if $f = \sum_{j=1}^{t} a_j g_j'$ with $a_j \in Q(S)$ and $b_{ij} = y_i^N a_j \in S$ then $y_k^N b_{ij} = y_i^N b_{kj} \in S$. Since $S = \Bbbk[y_0, \ldots, y_r]$ is factorial, this implies y_i^N divides b_{ij}, i.e. $a_j \in S$ and $f \in M'$, as claimed.

Consider now

$$\tilde{M} = \{f \in M' \mid \exists N \text{ such that } y_i^N f \in M \,\forall i\},$$

Then $\tilde{M}$ is finitely generated as a submodule of M'. Let $\tilde{g}_j$ be homogeneous generators, and consider an integer $\tilde{N}$, such that $y_i^{\tilde{N}} \tilde{g}_j \in M$ for all i,j. Then $(y_0, \ldots, y_r)^{(\tilde{N}-1)(r+1)+1} \tilde{g}_j \subset M$, since every monomial of degree $(\tilde{N}-1)(r+1)+1$ contains one of the $y_i^{\tilde{N}}$ as a factor. So $(y_0, \ldots, y_r)^{(\tilde{N}-1)(r+1)+1} \tilde{M} \subset M$ and

$$\tilde{M}_m = M_m \text{ for } m \geq m_0 = (\tilde{N}-1)(r+1) + 1 + \max\{\deg \tilde{m}_j\}.$$

□

This completes the proof of Theorem 7.4.2.

Remark 7.4.6. If M is a free S-module, then $M = M' = \tilde{M}$.

Definition 7.4.7. The homogeneous coordinate ring $R = \Bbbk[x_0, \ldots, x_n]/I(X)$ of a variety X is called **Cohen-Macaulay**, if there exists a Noether normalization $\Bbbk[y_0, \ldots, y_r] \hookrightarrow R$, which makes R into a free $\Bbbk[y_0, \ldots, y_r]$-module. The variety $X \subset \mathbb{P}^n$ is then called **arithmetically** or **projectively Cohen-Macaulay**.

Remark 7.4.8. The Cohen-Macaulay property of a homogeneous coordinate ring can be deduced from the initial ideal, if the homogeneous coordinates are adapted to the Noether normalization by the same criterion as in the affine case, compare Theorem 3.3.11.

The attribute arithmetic is used to avoid the conflict with the notion of a Cohen-Macaulay variety, which means, that all local rings $\mathcal{O}_{X,p}$ are Cohen-Macaulay rings. See Eisenbud Eisenbud (1995) for a definition of Cohen-Macaulay rings in general.

Exercise 7.4.9. Let $C \subset \mathbb{P}^3$ be a curve, whose homogeneous ideal $I(C) = \langle f, g \rangle$ is generated by two elements of degree d and e. Compute the degree and the arithmetic genus of C. Prove that R_C is Cohen-Macaulay by a Gröbner basis argument analogous to the proof of Theorem 4.3.18.3. □

Corollary 7.4.10. *If $C \subset \mathbb{P}^n$ is a smooth arithmetically Cohen-Macaulay curve then hypersurfaces of degree m cut out a complete linear system for every degree m.*

Example 7.4.11. Hypersurfaces do not always cut out a complete linear series as the following example shows. Let $C \subset \mathbb{P}^1 \times \mathbb{P}^1 \subset \mathbb{P}^3$ be a smooth curve of bidegree $(2, 4)$. We prove, that the quadrics in $\mathbb{P}^3$ do not cut out the complete system $|2H|$ on C. There is a 10-dimensional vector space of quadrics in $\mathbb{P}^3$, one of them is the equation of $\mathbb{P}^1 \times \mathbb{P}^1$, which vanishes on C. Thus quadric on $\mathbb{P}^3$ give a 9-dimensional subspace of $L(2H)$.

By Exercise 6.4.29, C has degree $\deg C = 6$ and arithmetic genus $p_a = 3$. A quadric $Q \neq \mathbb{P}^1 \times \mathbb{P}^1$ intersects C in 12 points. The residual intersection E of C with a cubic through these points consists of 6 points. These points impose at most 6 conditions on cubics in $\mathbb{P}^3$. So, there is a $20-6 = 14$-dimensional vector space of cubics through these points of which a four-dimensional subspace consists of multiples of the equation of $\mathbb{P}^1 \times \mathbb{P}^1$. Thus we obtain that $L(2H) = L(3H - E)$ has dimension at least $14 - 4 = 10$, which is greater than 9.

Theorem 7.4.12 (Riemann's inequality). *Let D be a divisor on a smooth projective curve of arithmetic genus p_a. Then*

$$\ell(D) \geq \deg D + 1 - p_a.$$

Proof. Let $C \subset \mathbb{P}^r$ be an embedding, and $p_C(t) = dt + 1 - p_a$ the Hilbert polynomial of this embedding. Suppose $D = D_1 - D_2$ is the difference of two effective divisors. We consider a sufficient large degree m, such that $h_C(m) =$

$p_C(m)$ and $(I_{D_1})_m \supsetneq (I_C)_m$. Let H_m be a hypersurface of degree m, which contains D_1 but not C. Then

$$(H_m) = D_1 + R$$

with $\deg R = md - \deg D_1$. Now we consider all hypersurfaces H'_m of degree m which pass through R and D_2. These hypersurfaces cut on C a system of dimension $\geq dm + 1 - p_a - (\deg R + \deg D_2)$. There intersection with C is $C.H'_m = D' + R + D_2$. Since $(H'_m) - (H_m)$ is a principal divisor, we obtain $D' + D_2 \equiv D_1$, ie. $D' \equiv D$. Thus, $\ell(D) \geq dm + 1 - p_a - (\deg R + \deg D_2) = dm - (dm - \deg D_1) - \deg D_2 + 1 - p_a = \deg D + 1 - p_a$ by Bézout's Theorem 6.4.33. □

Corollary 7.4.13. *The arithmetic genus is independent from the embedding* $C \hookrightarrow \mathbb{P}^r$.

Proof. Let $p_a^{(1)}$ and $p_a^{(2)}$ be the arithmetic genera for different embeddings of C. We want to prove $p_a^{(1)} = p_a^{(2)}$. By the completeness of hypersurface systems of large degree (Theorem 7.4.2), the complete linear system $|mH^{(1)}|$ is cut out by hypersurfaces of degree m, if m sufficiently large. So we have

$$\ell(mH^{(1)}) = m \deg H^{(1)} + 1 - p_a^{(1)}$$

for large m. Riemann's inequality based on the second embedding gives

$$\ell(mH^{(1)}) \geq m \deg H^{(1)} + 1 - p_a^{(2)}.$$

Thus $p_a^{(2)} \geq p_a^{(1)}$. The opposite inequality follows by the same argument. □

Proposition 7.4.14. *There exists a constant $d_0 = d_0(C)$ such that*

$$\ell(D) = \deg D + 1 - p_a$$

holds for every divisor of degree $\deg D \geq d_0$.

Proof. Let $C \subset \mathbb{P}^r$ be an embedding of the smooth curve, H the corresponding hyperplane class. By Theorem 7.4.2 we know, that there exists an $m > 0$, such that $|mH|$ is cut out by hypersurfaces of degree m and $\ell(mH) = \deg mH + 1 - p_a$. Consider any divisor D such that $\ell(D) \geq \deg D + 1 - p_a > \deg mH$. Then D is linear equivalent to a divisor $D' \geq mH$, because containing mH imposes at most $\deg mH$ conditions. By the same argument, we have $\ell(mH) \geq \ell(D') - (\deg D' - \deg mH)$. Since $\ell(mH) = \deg mH + 1 - p_a$, we obtain $\deg D' + 1 - p_a \geq \ell(D')$, and Riemann's inequlity is an equality for $D \equiv D'$. Thus, $d_0 = \deg mH + p_a - 1$ is sufficient. □

Corollary 7.4.15. *Every linear system $|D|$ of degree $d_0 + 2$ is very ample.*

□

Corollary 7.4.16. *Let C be a smooth irreducible projective curve and $p_1, \dots, p_r \subset C$ a finite collection of points. Then $C \setminus \{p_1, \dots, p_r\}$ is affine. In particular, $C \setminus \{p\}$ is affine for any point $p \in C$*

Proof. Let $d \geq \frac{d_0+2}{r}$. The divisor $D = dp_1 + \ldots + dp_r$ is a very ample and $1 \in L(D)$ corresponds to the equation of a hyperplane $H \subset \mathbb{P}(L(D))$, which intersects C only in $p_1, \dots, p_r$ (with multiplicity d in each point). □

Remark-Definition 7.4.17. A divisor D, where Riemann's inequality is strict is called **special**, the other divisors are called **non-special**.

The characterization of special divisors is the content of the Riemann-Roch Theorem 8.3.2, which will also give the precise value d_0.

The techniques so far give an algorithm to compute a complete linear system, provided we know, how to choose m_0. To obtain an honest algorithm for the computation of complete linear series on smooth space curves, we need to calculate m_0. The Riemann-Roch Theorem 8.3.2 below implies that $p_a = g$ and

$$\ell(D) = \deg D + 1 - g$$

for every divisor of degree $\deg D \geq 2g - 1$. Hence, $d_0 = 2g - 1$ and we can take m_0, such that

1. $m_0 \geq \frac{2p_a - 1}{d}$
2. $h_C(m) = dm + 1 - p_a$ for $m \geq m_0$.

Note, that $p_a(C)$ and the smallest m_0 satisfying 2.) can be calculated from the Betti numbers of a free resolution via the formula

$$h_C(m) = \sum_i (-1)^i \sum_j \beta_{ij} \binom{r + m - j}{r}.$$

Exercise 7.4.18. Implement an algorithm with the following spezification in your favorite computer algebra system.

Algorithm 7.4.19 (Computation of complete linear systems). .
Input: A smooth curve $C \subset \mathbb{P}^r$ and an effective divisor D on C given by their ideal in $\mathbb{P}^r$. Output: A basis for $L(D)$.

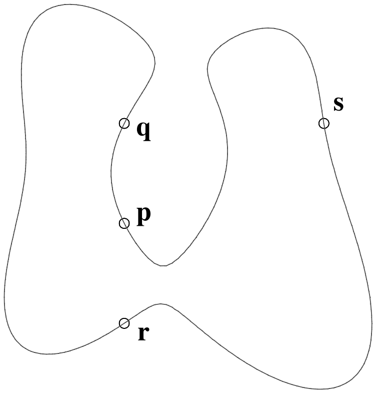
□

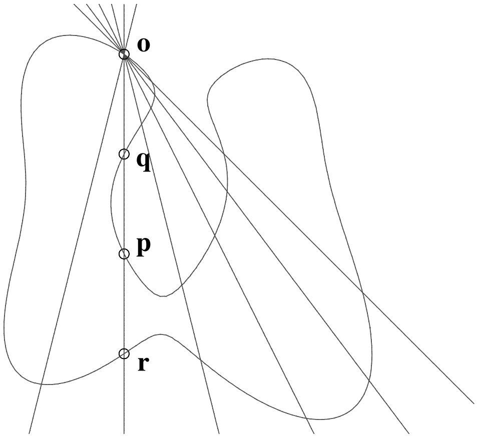
Example 7.4.20. Consider a smooth plane quartic C with 3 points. What is the dimension of $L(p_1 + p_2 + p_3)$?

To give a concrete example, we take C, the curve with the affine equation

$$2x(x^2-1)(x-2)+x^2y+y(y^2-1)(y+2)=0$$

and the points $p=(0,-1)$, $q=(0,0)$, $r=(0,-2)$ and $s=(2,0)$.



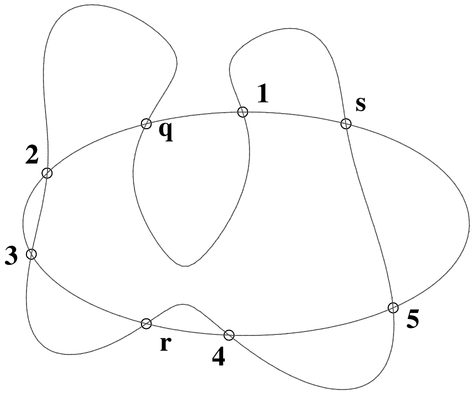


The space of rational functions $L(p+q+r)$ is two-dimensional, because the 3 points lie on a line. The pencil of lines through the forth intersection point $o=(0,1)$ cut out the complete linear system and $L(p+q+r)$ is spanned by 1 and $\frac{y-1}{x}\in\Bbbk(C)$

On the other hand, $L(q+r+s)$ is only one-dimenensional: The conic E defined by the affine equation

$$(x-1)^2+4(y+1)^2=5$$

intersects C in five further points, and any conic, which passes through these points, coincides with E by Bézout's theorem. Thus, $L(q+r+s)=\Bbbk$.



Exercise 7.4.21. Let D be a divisor of degree $\deg D=d-1$ on a smooth plane curve of degree d. Prove:

(1) $2\geq \ell(D)\geq 0$.

(2) $\ell(D)=2$ holds iff D is linearly equivalent to $d-1$ points on a line. □

Exercise 7.4.22. Let $C\subset\mathbb{P}^1\times\mathbb{P}^1\subset\mathbb{P}^3$ be a smooth curve of bi-degree $(3,3)$. Describe all special effective D and their dimensions $\ell(D)$ on C. □

Exercise 7.4.23. Compute m_0 for a smooth curve of bi-degree (a, b) on $\mathbb{P}^1 \times \mathbb{P}^1 \subset \mathbb{P}^3$. □

7.5 The δ-Invariant of Curve Singularities

In this section we study the δ-invariant of a curve singularity both from an geometric and arithmetic point of view, and prepare the proof that these invariants geometric and arithmetic delta invariant coincides. We start with the geometric δ-invariant.

If $C' \subset \mathbb{P}^2$ is an irreducible curve of degree d with non-ordinary singularities then the geometric genus might be strictly smaller than $\binom{d-1}{2} - \sum_{p\in C'} \binom{r_p}{2}$. In a sense, a non-ordinary singularity of multiplicity r_p can contribute more than $\binom{r_p}{2}$ to the difference $\binom{d-1}{2} - g$. We can calculate the precise contribution of a singularity $p \in C'$ as follows.

Consider an embedded resolution of the singularity $p \in C' \subset \mathbb{P}^2$

$$X^{(n)} \to X^{(n-1)} \ldots \to X^{(1)} \to X^{(0)} = \mathbb{P}^2$$

by a sequence of blow-ups.

A point $q \in X^k$, which is mapped to p, is called **infinitesimally near** to p. More precisely, points on the exceptional curve E of $X^1 \to X^0$ are called **infinitesimally near** to the **first order**. Inductively, a point on the exceptional divisor of the blow-up of an infinitesimally near point of k-th order are called infinitesimally near to the $k+1$-st order. p itself is called infinitesimally near to p of order 0. The full contribution of p is then

$$\delta_p := \sum_{\substack{q \text{ infinitesimally} \\ \text{near } p}} \binom{r_q}{2},$$

where r_q denotes the multiplicity of the strict transform of C' at q. We call δ_p the (geometric) **delta invariant** of the singularity (C', p). Our proof of resolution of singularities based on Theorem 7.2.23, shows the following:

Corollary 7.5.1. *Let $C' \subset \mathbb{P}^2$ be an absolutely irreducible curve of degree d with arbitrary singularities. The geometric genus is*

$$g = \binom{d-1}{2} - \sum_{p\in C'} \delta_p = \binom{d-1}{2} - \sum_{\substack{\text{all } q, \text{ including} \\ \text{inf. near points}}} \binom{r_q}{2}.$$

Exercise 7.5.2. Prove: $\delta_p = 1$ iff (C', p) is an ordinary node or an ordinary cusp. □

In the remaining part of this section we will give an arithmetic description of the delta invariant, i.e. an interpretation in terms of functions. This also leads to an interpretation of intersection multiplicities of plane curves in terms of vanishing orders on the desingularization of one of the curves. To start we note that normalization yields a second approach towards desingularization of curves.

**** **fix** ****perhaps move to Chapter 4, with some of the remarks. Theorem Normalization of curves should stay here.**** **fix** ****

Definition 7.5.3. The normalization of a reduced ring R is the integral closure $\widetilde{R}$ in its total fraction ring $Q(R)$. R is normal if $R = \widetilde{R}$

For a affine ring, being normal is a local property.

Theorem 7.5.4. *Let R be a reduced ring, and $Q(R)$ its total ring of fractions. Then the following are equivalent*

1. *R is normal.*
2. *$R_{\mathfrak{p}}$ is normal for all prime ideals $\mathfrak{p}$.*
3. *$R_{\mathfrak{m}}$ is normal for all maximal ideals $\mathfrak{m}$.*

Theorem 7.5.5 (Splitting of normalization). *Let R be a reduce noetherian ring, $\langle 0 \rangle = \mathfrak{p}_1 \cap \ldots \cap \mathfrak{p}_s$ the primary decomposition into the minimal primes. Then the normalization*

$$\widetilde{R} \cong \widetilde{R/\mathfrak{p}_1} \oplus \ldots \oplus \widetilde{R/\mathfrak{p}_r}$$

splitts into the sum of the normalizations of the components of R. Moreover, $\widetilde{R/\mathfrak{p}_j}$ coincides with normalization of $R/\mathfrak{p}_i$ in the total quotient ring $Q(R)$.

Theorem 7.5.6. *A 1-dimension local ring R is normal, iff R is a discrete valuation ring. In particular, normal 1-dimensional rings are domains.*

Corollary 7.5.7. *The coordinate ring of a smooth affine curve is normal.*

Proof. By Corollary **??** the local ring of a smooth curve at a point is discrete valuation ring. Hence the Corollary follows by combining Theorem 7.5.4 and Theorem 7.5.6. □

Theorem 7.5.8 (Normalization of curves). *Let $\eta \colon \widetilde{C} \to C$ be a resolution of singularities and $U \subset C$ and affine subset. Then $\widetilde{U} = \eta^{-1}(U)$ is affine as well, and $\Bbbk[\widetilde{U}]$ is the integral closure of $\Bbbk[U]$ in the total fraction ring $\Bbbk(C) = \Bbbk(C_1) \oplus \ldots \oplus \Bbbk(C_r)$ of C, where $C = C_1 \cup \ldots \cup C_r$ denotes the irreducible components of C. In particular, $\Bbbk[\widetilde{U}]$ is a finite $\Bbbk[U]$-module.*

Proof. Let $C_1, \ldots, C_s$ be the irreducible components of C which intersect U, and let $\widetilde{C}_1, \ldots, \widetilde{C}_s$ denote their desingularizations. Then $C_i \setminus U$ consists of finitely many points, and $D_i = \widetilde{C}_i \setminus \widetilde{U}$ is a finite collection of points, which we may regard as an effective divisor on C_i of positive degree. By Corollary 7.4.16

the intersections $\widetilde{C}_i \cap \widetilde{U}$ are affine. Hence the disjoint union $\widetilde{U} = \bigcup_{i=1}^{s} \widetilde{C}_i \cap \widetilde{U}$ is affine as well, and its coordinate ring

$$\mathbb{k}[\widetilde{U}] = \bigoplus_{i=1}^{s} \mathbb{k}[\widetilde{C}_i \cap \widetilde{U}]$$

is a finite $\mathbb{k}[U]$-module by Theorem **??**. In particular $\mathbb{k}[\widetilde{U}]$ is contained in the normalization $\widetilde{\mathbb{k}[U]}$ of $\mathbb{k}[U]$. Since by $\mathbb{k}[\widetilde{U}]$ is normal by Corollary 7.5.7, we have equality, and the desired result follows with Theorem 7.5.5 □

Remark 7.5.9. 1. The integral closure $\widetilde{R}$ of an integral domain R in its quotient $K = Q(R)$ is not always a finite R-module, even for Noetherian rings, see Nagata **?**. However for an affine domain the integral closure is always an affine domain again. This Theorem due to Emmy Noether, is a nice application of Noether normalization and Galois theory, and the reason, why we attach the name Noether instead of Hilbert to the concept of Noether normalization.

2. The normalization $\widetilde{X}$ of an affine algebraic set X is defined to be the affine algebraic set associated to the direct sum of the normalizations of its irreducible components. For an projective algebraic set, can define the normalization, either by gluing the normalization of its affine charts, or by taking the projective algebraic set, associated to the normalization of its homogeneous coordinate ring, which, as one can show, lies in some weighted projective space. A variety is called normal, if the affine coordinate rings for a affine covering are normal.
3. Various algorithm are known to compute the normalization of affine rings, or weighted homogeneous rings. `SINGULAR` has an implementation of the algorithm of Theo de Jong **?**
4. The singularities $\operatorname{sing} X$ of a normal variety X have codimension at least 2. As a consequence, for an irreducible subvariety V of X all local rings $\mathcal{O}_{X,V}$, defined as $\mathbb{k}[X \cap U]_{\mathrm{I}(V \cap U)}$ for an affine chart $U \subset X$ which intersects V, are discrete valuation rings. This allows to define the divisor of poles and zeroes for a rational function $f \in \mathbb{k}(X)$ as

$$(f) = \sum_{\substack{V \subset X \text{ subvariety} \\ \operatorname{codim}_X V = 1}} v_V(f) \,,$$

where $v_V : \mathbb{k}(X) \to \mathbb{Z}$ denotes the discrete valuation associated to $\mathcal{O}_{X,V} \subset \mathbb{k}(X)$. Hence concept of linear equivalence of divisors can be extended $\operatorname{Div}(X)$. We denote with

$$Cl(X) = \operatorname{Div}(X)/\equiv$$

the divisor class group of a normal variety X.

Exercise 7.5.10. Prove,

$$Cl(\mathbb{P}^n) \cong \mathbb{Z}$$

with the isomorphism induced by the degree of a hypersurface, i.e. induced by the map

$$\mathrm{Div}(\mathbb{P}^n) \to \mathbb{Z},\, H = \sum_{k=1}^{s} \mu_k H_k \mapsto \sum_{k=1}^{s} \mu_k \deg H_k$$

Let C be a curve, let $p \in C$ be a point, and let $\eta : \widetilde{C} \to C$ be a desingularization. Let U be an affine neighbarhood of $p \in C'$ such that $U \setminus \{p\}$ is smooth, and let $\widetilde{U}$ be the preimage of U in C. Then by Theorem **??** $\Bbbk[\widetilde{U}]$ is a finite $\Bbbk[U]$-module, and the inclusion $\Bbbk[U] \hookrightarrow \Bbbk[\widetilde{U}]$ is an isomorphism at every point except possibly at p. Hence the cokernel $\Delta_p = \Bbbk[\widetilde{U}]/\Bbbk[U]$ is finitely generated $\Bbbk[U]$ module with support at p. In particular Δ_p is a module of finite length. Localizing at p we find the exact sequence

$$0 \to \mathcal{O}_{C,p} \to \widetilde{\mathcal{O}_{C,p}} \to \Delta_p \to 0$$

which we take as a definition of Δ_p entirely in terms of the local ring $\mathcal{O}_{C,p}$.

Definition 7.5.11. Let $p \in C$ be a point on a curve. Then we call

$$\mathrm{length}\, \Delta_p$$

the **arithmetic δ-invariant**

Remark 7.5.12. The arithmetic δ-invariant is zero, iff C is smooth at p.

Theorem 7.5.13 (Structure of artinian rings). *Let R be an artinian ring and let $\mathfrak{m}_1, \dots, \mathfrak{m}_r$ denote its finitely many maximal ideals. Then*

$$R \cong \bigoplus_{j=1}^{r} R_{\mathfrak{m}_j}$$

Exercise 7.5.14. Compute the arithmetic δ-invariant for the following curve singularities

1. The A_k singularity $y^2 - x^{k+1}$ over an field with $\mathrm{char}\, \Bbbk \neq 2$.
2. The singularity of the space curve defined by the ideal $\langle xy, xz, yz \rangle$

Proposition 7.5.15. *Let $\eta : \widetilde{C} \to C$ be the desingularization of a possibly reducible projective curve. Then*

$$p_a(\widetilde{C}) = p_a(C) + \sum_{p \in C} \mathrm{length}\, \Delta_p.$$

Proof. Since the arithmetic genus of a (possibly singular) curve does not change under the Veronese embeddings by Exercise 6.4.28, we may assume, that each component C_i of $C \subset \mathbb{P}^n$ has degree $\deg C_i \geq d_0(\widetilde{C_i}) + 2$ with the bound $d_0(\widetilde{C}_i)$ from Corollary 7.4.15. Then

$$\widetilde{C} \to \mathbb{P}(\bigoplus_{i=1}^{s} L(H_i)) = \mathbb{P}^N$$

is an embedding, where $H_i = \eta|_{C_i}^{*} H$ denotes the pullback divisor to C_i of a general hyperplane. Moreover, the map $\eta : C \to C'$ is now induced by a projection $\mathbb{P}^N \dashrightarrow \mathbb{P}^n$ from a center disjoint from C. Choose a hyperplane $H \subset \mathbb{P}^n$ which intersects C' transversally in smooth points. Then $C_{\text{aff}} = C \setminus H \cap C'$ and $\widetilde{C}_{\text{aff}} = \eta^{-1}(C_{\text{aff}})$ are affine, and $\Bbbk[\widetilde{C}_{\text{aff}}]$ is a finite $\Bbbk[C_{\text{aff}}]$-module. Choose $m \gg 0$ large enough, such that:

1. The Hilbert function and Hilbert polynomial of C take the same value, i.e. $h_C(m) = dm + 1 - p_a(C)$.
2. The same for $\widetilde{C}$, i.e. and $h_{\widetilde{C}}(m) = dm + 1 - p_a(\widetilde{C})$.
3. The composition
$$\Bbbk[\mathbb{A}^N]_{\leq m} \to \Bbbk[\widetilde{C}_{\text{aff}}] \to \varDelta = \bigoplus_{p \in C} \varDelta_p$$
from the space of polynomials of degree $\leq m$ is surjective.

Then the degree m piece $(S_C)_m$ of the homogeneous coordinate rings of C has codimension in $(S_{\widetilde{C}})_m$ equal to the length $\varDelta$. Hence

$$md + 1 - p_a(\widetilde{C}) - \operatorname{length} \varDelta = md + 1 - p_a(C).$$

□

Exercise 7.5.16. A singularity is called **uni-branched**, if $p \in C'$ has a single preimage $q \in C$ in a desingularization $\eta : C \to C'$. Prove that

1. $\varGamma = \{v_q(g) \mid g \in \mathcal{O}_{C',p}\} \subset \mathbb{N}$ is a submonoid of $(\mathbb{N}, +)$,
2. The length $\varDelta_p$ is the number of gaps $\mathbb{N} \setminus \varGamma$.

We call $\varGamma$ the **value monoid** of p □

Exercise 7.5.17. Give an example of a value monoid $\varGamma \subset \mathbb{N}$ which cannot occur for a uni-branched plane curve. Prove, that the number of generators of the value monoid $\varGamma$ of a uni-branched plane curve singularity can be arbitrarily large. □

Exercise 7.5.18. A **monomial space curve** is the image of

$$\mathbb{P}^1 \to \mathbb{P}^3,\ [s:t] \mapsto [s^d : s^a t^{d-a} : s^b t^{d-b} : t^d]$$

with $gcd(a, b, d) = 1$. Compute the arithmetic genus and the delta invariants in examples. What is the maximal arithmetic genus possible for fixed d? □

The resolution of singularities gives another way to compute intersection multiplicites of hypersurfaces with curves.

Theorem 7.5.19. *Let $C \subset \mathbb{A}^n$ be a curve, and let $p \in C$ be a point. Let $\eta : \widetilde{C} \to C$ be a resolution of the singularity of C at p. For $q \in \eta^{-1}(p)$, let v_q be the valuation on the function field of the irreducible component of C, which contains q. Let g be a square free polynomial and $H = \mathrm{V}(g) \subset \mathbb{A}^n$ the corresponding hypersurface. Then*

$$\imath(C, H; p) = \sum_{q \in \eta^{-1}(p)} v_q(g).$$

Proof. We may assume, that g does not vanish on any component of C which passes through p, because otherwise both sides give ∞. **??**. The Snake Lemma **??** gives the following diagram of exact sequences

$$\begin{tikzcd}
 & & & 0 \arrow[d] & \\
 & 0 \arrow[d] & 0 \arrow[d] \arrow[r] & A \arrow[d] \arrow[dddll, rounded corners] & \\
0 \arrow[r] & \mathcal{O}_{C,p} \arrow[r] \arrow[d] & \widetilde{\mathcal{O}_{C,p}} \arrow[r] \arrow[d] & \Delta_p \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathcal{O}_{C,p} \arrow[r] \arrow[d] & \widetilde{\mathcal{O}_{C,p}} \arrow[r] \arrow[d] & \Delta_p \arrow[r] \arrow[d] & 0 \\
 & \mathcal{O}_{C,p}/g\mathcal{O}_{C,p} \arrow[r] \arrow[d] & \widetilde{\mathcal{O}_{C,p}}/g\widetilde{\mathcal{O}_{C,p}} \arrow[r] \arrow[d] & B \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

We assume that the ground field $\Bbbk \cong \mathcal{O}_{C,p}/\mathfrak{m}_p$ is algebraically closed. Then length coincides with $\dim_{\Bbbk}$ of vector spaces. By Exercise **??** and length $\Delta_p < \infty$ we have

$$\operatorname{length} A = \operatorname{length} B$$

and hence

$$\imath(C, H; p) = \operatorname{length} \mathcal{O}_{C,p}/g\mathcal{O}_{C,p} = \operatorname{length} \widetilde{\mathcal{O}_{C,p}}/g\widetilde{\mathcal{O}_{C,p}}$$

Now $\widetilde{\mathcal{O}_{C,p}}$ is an semi-local ring, whose finitely many maximal ideals are in bijection with the points $q \in \eta^{-1}(p)$, because our ground filed is algebraically closed. The residue ring $R = \widetilde{\mathcal{O}_{C,p}}/g\widetilde{\mathcal{O}_{C,p}}$ is artinian and its maximal ideals are still in bijection with the points of $\eta^{-1}(p)$. By Structure Theorem of Artinian rings

$$R = \bigoplus_{q \in \eta^{-1}(p)} R_{\mathfrak{m}_q}.$$

Since localization commutes with taking residue rings we have

$$R_{\mathfrak{m}_q} = \mathcal{O}_{\widetilde{C},q}/g\mathcal{O}_{\widetilde{C},q}.$$

Hence

$$\operatorname{length} R = \sum_{q\in\eta^{-1}(p)} \operatorname{length} \mathcal{O}_{\widetilde{C},q}/g\mathcal{O}_{\widetilde{C},q} = \sum_{q\in\eta^{-1}(p)} v_q(g).$$

□

The local condition in Noether's $AF + BG$ Theorem 5.5.1 likewise allows sometimes a formulation using valuations on the desingularization.

Proposition 7.5.20. *Let $C \subset \mathbb{P}^2$ be a plane curve with defining equation* $\mathrm{i}(C) = \langle f\rangle$ *and let $\eta : \widetilde{C} \to C$ be a desingularization. Let g, h be two further forms which do not vanish on a component of C. Suppose that $p \in C$ is an ordinary r-fold point. Then*

$$h_p \in \langle f_p, g_p\rangle \mathcal{O}_p$$

if $v_q(h) \geq v_q(g) + r - 1$ for all $q \in \eta^{-1}(p)$.

Proof. Let $C = C_1 \cup \ldots \cup C_s$ be the decomposition into irreducible components. We may assume that every component of C passes through p, so that

$$\mathcal{O}_{C,p} \subset \widetilde{\mathcal{O}_{C,p}} \subset \Bbbk(C_1) \oplus \ldots \oplus \Bbbk(C_s).$$

Let $\eta_i : \widetilde{C}_i \to C_i$ the normalization of C_i. We claim that

$$\bigoplus_{i=1}^{s} \bigcap_{q\in\eta_i^{-1}(p)} \{a \in \Bbbk(C_i) \mid v_q(a) \geq r-1\} \subset \mathcal{O}_{C,p}.$$

This implies $h/g \in \mathcal{O}_{C,p}$ as desired.

Since $\varDelta_p$ finite length we have that $\bigoplus_{i=1}^{s} \bigcap_{q\in\eta_i^{-1}(p)} \mathfrak{m}_{C_i,q}^N \subset \mathcal{O}_{C,p}$ for $N \gg 0$. So it suffices to find for each $q \in \eta_i^{-1}(p)$ and each n with $r - 1 \leq n \leq N$ functions $a \in \mathcal{O}_{C,p}$ such that $v_q(a) = n$ and $v_{q'}(a) \geq N$ for all $q' \neq q$.

We can approximate the smooth branch corresponding to q by a zero-locus of a polynomial $a_q \in \mathcal{O}_{\mathbb{P}^2,p}$, which is non-singular at p to arbitrary high order, i.e. there are a_q's with $v_q(a_q) \geq N$. Note that $v_{q'}(a_q) = 1$ for $q' \neq q$, because C has an ordinary singularity at p. Now let $x \in \mathfrak{m}_p$ be a linear form, whose zero-locus is not tangent to C at p. Then the function

$$a = x^{n-r+1} \prod_{q'\neq q} a_{q'} \in \mathcal{O}_{C,p},$$

has the desired vanishing property. □

In Corollary 8.2.6 we will prove $g = p_a$ for smooth irreducible curves. Using this, we deduce the following Proposition.

Proposition 7.5.21. *Let $p \in C$ be a plane curve singularity. Then*

$$\delta_p = \operatorname{length} \varDelta_p$$

Proof. First, note that both values δ_p and $\operatorname{length} \varDelta_p$ depend only on the completion $\widehat{\mathcal{O}}_{C,p}$ and, even better, only on $\mathcal{O}_{C,p}/\mathfrak{m}_{C,p}^N$ for an $N \gg 0$ depending on $\widehat{\mathcal{O}}_{C,p}$. For δ_p this is clear, because the location of the infinitesimally near points and their multiplicities depend only on the the equation $f \mod \mathfrak{m}_p^N$ for $N \gg 0$. An explicit value for N can be deduced from the embedded resolution (7.2.13). For $\operatorname{length} \varDelta_p$ this holds, because, as we just saw, the locations of the points $q \in \eta^{-1}(p) \subset C$ depend only on $\mathcal{O}_{C,p}/\mathfrak{m}_{C,p}^N$, and, because $\varDelta_p$ is annihilated by a power $\mathfrak{m}_{C,p}^N$. Thus, to compare arithmetic and geometric delta invariant, we may add a general polynomial in $\mathfrak{m}_p^N$ of degree $e \gg N$ to the affine equation f of C. The altered curve C'' is irreducible and has p as its only singularity by Bertini's Theorem 6.6.1, applied to the image of $\mathbb{P}^2$ under the rational map defined by $L(e; Np) \oplus \Bbbk f$, and by the irreducibility of a general hyperplane section (see 6.7.15 for the case $\operatorname{char} \Bbbk = 0$, for $\operatorname{char} \Bbbk > 0$ see Remark 7.5.22 below). The geometric genus of the new curve C'' is $g = \binom{e-1}{2} - \delta_p$. On the other hand the arithmetic genus of a plane curve of degree e is $\binom{e-1}{2}$, hence $p_a(C'') = \binom{e-1}{2} - \operatorname{length} \varDelta_p$ by Proposition 7.5.15. From $g = p_a(C'')$, which we will prove in Corollary 8.2.6, we obtain $\delta_p = \operatorname{length} \varDelta_p$. □

****** fix ******Improve–move parts to other sections, e.g. the upper bound on the local contribution

Remark 7.5.22. In the special case of the proof above, one can deduce the irreducibility of C'' for $e \gg N \gg 0$ as follows: Suppose C'' is reducible, say $C'' = C''_a \cup C''_b$ of curves degree $a + b = e$. Then the number of intersection points of two pieces is $a \cdot b \geq e - 1$ counted with multiplicities. We argue that some of the intersection points would be different from p, contradicting the smoothness of C'' away from p. Indeed, the contribution $\imath(C''_a, C''_b, p)$ to the intersection is at most $N_a \cdot N_b \leq (N/2)^2$ because the degree N part of the affine equation of C'' at p is general. Since $e \gg N$ we have $e - 1 > (N/2)^2$, and we would have further intersection points.

Chapter 8

Riemann-Roch and Applications

The Riemann-Roch Theorem 8.3.2 is the starting point for the study of geometric properties of algebraic curves in their various embeddings. As we have seen in the last Chapter, the Riemann-Roch spaces govern the study of maps $C \to \mathbb{P}^n$. The Riemann-Roch Theorem identifies the difference $\ell(D)-(\deg D+1-g)$ in Riemann's inequality 7.4.12 with the dimension of the space of global differential forms on C with zeroes in D. Thus, in Section 8.1 we study differential forms. In Section 8.2 we compute the canonical divisor class, i.e. the class of the divisor of poles and zeroes of a rational differential form, in terms of a plane model of the curve and prove the completeness of adjoint systems. Section 8.3 contains the proof of the Riemann-Roch Formula and the most basic applications.

In the remaining section we illustrate some applications. Section 8.4 contains Hurwitz' formula, which relates for a non-constant morphism $C \to E$ between smooth projective curves, the genus of C and E with the number of ramification points. This gives usually the easiest method to compute the geometric genus. As a consequence we prove Lüroth's Theorem, which says that uni-rational curves are rational, and the famous Plücker relations between the numerical invariants of a plane curve and their dual. In Section 8.5, we compute the number of Weierstrass points on a curve of genus g and deduce the finiteness of the automorphismen group in case $g > 1$. In the next Section we make Riemann's intuitive count: curves of genus $g > 1$ depend on $3g-3$ moduli. Of course, the count gives a rigorous argument, only if one establishes the existence of the Picard group $\operatorname{Pic}^0(C)$ and the moduli space as algebraic varieties, for which we do not have the techniques.

In Section 8.7 we study the equations of curves in their canonical embedding at some length. We prove Max Noether's and Petri's Theorem, and comment on Green's Conjecture about the connection between the syzygies of canonical curves and their Clifford index.

In the final Section 8.8 of our book we present Stepanov's proof of the Hasse-Weil Formulas for curves over finite fields.

Throughout this chapter C will denote an absolutely irreducible smooth projective curve.

8.1 Differential Forms

Let C be a smooth irreducible projective curve over an algebraically closed field $\mathbb{k}$. We want to introduce differential forms on C.

In case of $\mathbb{k} = \mathbb{C}$, we can think of C as a Riemann surface, on which we have the notion of meromorphic differential forms. In particular for any rational function f, the rational differential form df is defined. Guided by this and the obvious product rule, we define differential forms in general.

Remark-Definition 8.1.1. Consider the $\mathbb{k}(C)$-vector space with a basis denoted by the symbols "$[f]$" for each $f \in \mathbb{k}(C)$ and the subspace generated by the expressions

1. $[fg] - g[f] - f[g]$ for any pair $f, g \in \mathbb{k}(C)$,
2. $[f+g] - [f] - [g]$ for any pair $f, g \in \mathbb{k}(C)$,
3. $[\lambda f] - \lambda[f]$ for any pair $\lambda \in \mathbb{k}$ and $f \in \mathbb{k}(C)$.

The quotient vector space

$$\Omega(C) := \langle [f] \rangle / \langle \text{relations (1),(2),(3)} \rangle$$

is the space of **rational differential forms**. The natural map

$$d \colon \mathbb{k}(C) \to \Omega(C),\ f \mapsto df,$$

where df denotes the class of $[f]$ in the quotient space, is $\mathbb{k}$-linear due to the relations (2) and (3). However, d is not $\mathbb{k}(C)$-linear although both spaces are $\mathbb{k}(C)$-vector spaces. Instead d satisfies the product rule $d(fg) = gdf + fdg$ due to (1).

The quotient rule

$$d(\frac{f}{g}) = \frac{gdf - fdg}{g^2}$$

follows from the product rule applied to $f = \frac{f}{g}g$. Other rules familiar from caculus follow as well: $d(1) = d(1^2) = d(1) + d(1) \Rightarrow d(1) = 0$, and $\mathbb{k}$-linearity implies

$$d\lambda = 0 \quad \forall \lambda \in \mathbb{k}$$

Repeated application of the product rule gives the chain rule

$$dF = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i} dx_i$$

for a polynomial expression $F(x_1 \ldots, x_n) \in \mathbb{k}[x_1, \ldots, x_n]$ in the coordinate functions $x_i \in \mathbb{k}(C)$.

Proposition 8.1.2. *$\Omega(C)$ is a 1-dimensional $\Bbbk(C)$-vector space. If $\operatorname{char}\Bbbk = 0$ then df for any non constant rational function $f \in \Bbbk(C)$ is a basis. If $\operatorname{char}\Bbbk = p$ then df for any element in $\Bbbk(C) \setminus \Bbbk(C)^p\Bbbk$ generates. Here, $\Bbbk(C)^p\Bbbk$ denotes the $\Bbbk$-subspace generated by p^{th} powers.*

Proof. Consider an affine plane model C' of C given by an equation $F \in \Bbbk[x, y]$. Using the product and quotient rule repeatedly, we see, that dx, dy form a generating set. So $\Omega(C)$ is at most a 2-dimensional $\Bbbk(C)$ vector space. Since C' is smooth at a general point,

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy = 0 \in \Omega(C)$$

gives us a non-zero relation among dx and dy. So, $\Omega(C)$ is at most 1-dimensional. Thus, for the first statement it remains to prove, that $\Omega(C)$ is not the zero vector space. In case $\Bbbk = \mathbb{C}$ this is easy: There is a map

$$\Omega(C) \to \{\text{meromorphic differental forms on } C\},$$

because derivations of meromorphic functions satisfy $\mathbb{C}$-linearity and the product rule. Since image of df of a non-constant function f is non-zero we deduce that $\Omega(C) \neq 0$ in this case. We postpone the general case to the end of this section, Corollary 8.1.19 below.

To prove the second statement, we consider elements $x = f$ and a primitive element $y \in \Bbbk(C)$ of the field extension $\Bbbk(x) \subset \Bbbk(C)$, which exists since $\operatorname{char}\Bbbk = 0$. With the plane model defined by x and y, and the computation of in first part, we obtain that $dx = df$ is a $\Bbbk(C)$-basis of $\Omega(C)$. In case of characteristic p we note, that $df = 0$ for any $\Bbbk$-linear combination of p^{th} powers. On the other hand, if $f \notin \Bbbk(C)^p\Bbbk$ then $\Bbbk(f) \subset \Bbbk(C)$ is a finite separable extension, ****** fix ******ref or cite **??** or and there exists again a primitive element $y \in \Bbbk(C)$. □

Definition 8.1.3. Let $\omega = gdf$ be a non-zero rational differential form, let $p \in C$ a (smooth) point, and let $v_p \colon \Bbbk(C) \setminus \{0\} \to \mathbb{Z}$ be the discrete valuation with valuation ring $\mathcal{O}_{C,p}$. If $t \in \mathfrak{m}_p \subset \mathcal{O}_{C,p}$ is a local parameter of C at p and $\omega = hdt$ then

$$v_p(\omega) = v_p(h)$$

is the **vanishing order** of ω in p.

$$K := (\omega) := \sum_{p \in C} v_p(\omega)p$$

is called **a canonical divisor** on C.

Remark 8.1.4. 1. Every differential form ω can be written as $\omega = hdt$ for some rational function $h \in \Bbbk(C)$, because $\Omega(C)$ is 1-dimensional and dt a basis.

2. If $s \in \mathcal{O}_{C,p}$ is another local parameter then $s = (\lambda + g)t$ with $\lambda \in \Bbbk, \neq 0$ and $g \in \langle t \rangle$, because $\mathfrak{m}_p/\mathfrak{m}_p^2$ is one-dimensional. Hence, $ds = (\lambda + g)dt + tdg = (\lambda + g + tg')dt$ with $g + tg' \in \mathfrak{m}_p$. So $\frac{ds}{dt} = \lambda + g + tg'$ is a unit in $\mathcal{O}_{C,p}$ and the vanishing order of ω at p is independent of the choice of the local parameter.
3. The function $t - t(q)$ is a local parameter at q for q in a Zariski open neighborhood of p. Since $t(q) \in \Bbbk$, we have $dt = d(t - t(q))$. Hence, the local expression $\omega = hdt$ is valid in the whole affine neighborhood. In particular, the zeroes and poles of ω form a finite set, and (ω) is indeed a divisor.
4. Since $\Omega(C)$ is a 1-dimensional $\Bbbk(C)$ vector space we have that the ratio of two non-zero differential forms $\frac{\omega_1}{\omega_2} = f$ is a rational function. In particular $(\omega_1) = (\omega_2) + (f)$. Hence two canonical divisors are linearly equivalent. What is really canonical is the divisor class of K.

We will compute the degree of a canonical divisor in Section 8.2. In the rest of this section we establish $\Omega(C) \neq 0$ for arbitrary algebraically closed fields $\Bbbk$. In case $\Bbbk = \mathbb{C}$, we reduced the statement to the local computation of the derivative of a meromorphic function. The proof in the general case is a similar reduction to a local computation. First, we generalize the construction of differential forms.

Definition 8.1.5. Let R be a $\Bbbk$-algebra, and M an R-module. A **$\Bbbk$-derivation** of R with values in M is a $\Bbbk$-linear map

$$\partial : R \to M,$$

which satisfies the product rule

$$\partial(fg) = \partial(f)g + f\partial(g).$$

The R-module of all $\Bbbk$-derivations of R with values in M is denote by $Der_{\Bbbk}(R, M)$ or briefly by $Der(R, M)$, if the ground ring $\Bbbk$ is clear.

Example 8.1.6. For $R = \Bbbk[x_1, \ldots, x_n]$ the partial derivatives $\frac{\partial}{\partial x_i}$ is an element of $Der_{\Bbbk}(R, R)$. We will see below that these elements generate the whole module of derivations.

Derivations allow to define the Zariski tangent space in yet another way.

Proposition 8.1.7. *Let R be a $\Bbbk$-algebra and $\mathfrak{m} \subset R$ a maximal ideal with residue field $R/\mathfrak{m} \cong \Bbbk$. Then*

$$Der(R, R/\mathfrak{m}) \cong \mathrm{Hom}_{\Bbbk}(\mathfrak{m}/\mathfrak{m}^2, \Bbbk).$$

Proof. Let f_0 denote the image of f under the composition

$$R \to R/\mathfrak{m} \cong \Bbbk \hookrightarrow R.$$

The value $\partial(f)$ of a derivation $\partial \in Der_{\Bbbk}(R, R/\mathfrak{m})$ depends only on the class of $f - f_0$ in $\mathfrak{m}/\mathfrak{m}^2$, since the $\Bbbk$-linearity and the product rule imply $\partial(\Bbbk + \mathfrak{m}^2) = 0$. On the other hand, the composition $R \to \mathfrak{m} \to \mathfrak{m}/\mathfrak{m}^2$, defined by

$$f \mapsto f - f_0 \mapsto f - f_0 + \mathfrak{m}^2,$$

is already a derivation with values in $\mathfrak{m}/\mathfrak{m}^2$: $fg - f_0 g_0 = g_0(f - f_0) + f_0(g - g_0) + (f - f_0)(g - g_0) \equiv g_0(f - f_0) + f_0(g - g_0) \equiv g(f - f_0) + f(g - g_0) \mod \mathfrak{m}^2$. The result follows. □

Definition 8.1.8. The **universal derivation** $d : R \to \Omega_{R/\Bbbk}$ is an R-module $\Omega_{R/\Bbbk}$ together with a $\Bbbk$-derivation $d \in Der(R, \Omega_{R/\Bbbk})$, such that any $\Bbbk$-derivation $\partial \in Der_{\Bbbk}(R, M)$ is obtained from d by composing with a unique R-module homomorphism $\phi : \Omega_{R/\Bbbk} \to M$. $\Omega_{R/\Bbbk}$ is called the module of **Kähler differentials** of R over $\Bbbk$.

Uniqueness of $d : R \to \Omega_{R/\Bbbk}$ follows from the universal property. Existence follows by mimicking the construction in Definition 8.1.1.

Example 8.1.9. Let $S = \Bbbk[x_1, \ldots, x_n]$ be the polynomial ring. Applying the product rule repeately we see that $\Omega_{S/\Bbbk}$ is generated by $dx_1, \ldots, dx_n$. These elements are actually R-linearly independent. From the derivation $\frac{\partial}{\partial x_i} \in Der_{\Bbbk}(R, R)$ we obtain a R-linear map $\phi_i : \Omega_{R/\Bbbk} \to R$, such that $\frac{\partial}{\partial x_i} = \phi_i \circ d$. It follows $\phi_i(dx_j) = \delta_{ij}$, and any relation $g_1 dx_1 + \ldots + g_n dx_n = 0 \in \Omega_{R/\Bbbk}$ has trivial coefficients $g_i = \phi_i(g_1 dx_1 + \ldots + g_n dx_n) = 0 \in R$. Thus

$$\Omega_{S/\Bbbk} = \bigoplus_{i=0}^{n} S dx_i \cong S^n$$

and

$$Der(S, S) = \bigoplus_{i=0}^{n} S \frac{\partial}{\partial x_i}.$$

If R is a finitely generated $\Bbbk$-algebra, say $R = \Bbbk[x_1, \ldots, x_n]/I$ then the Kähler differentials $d\overline{x}_i$ generate $\Omega_{R/\Bbbk}$, as we can see from the product rule applied repeatedly. Moreover each element $f \in I$ defines a relation

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_j} d\overline{x}_j = 0 \in \Omega_{R/\Bbbk}.$$

Consider the maps

$$I \to \Omega_{S/\Bbbk} \otimes_S R, \ f \mapsto df \otimes 1$$

and

$$\Omega_{S/\Bbbk} \otimes_S R \to \Omega_{R/\Bbbk}, \ dx_i \otimes 1 \mapsto d\overline{x}_i$$

Proposition 8.1.10. *The sequence*

$$I/I^2 \to \Omega_{S/\mathbb{k}} \otimes_S R \to \Omega_{R/\mathbb{k}} \to 0$$

is an exact sequence of $R = S/I$ modules.

Proof. Consider the R-modul

$$\Omega = \frac{\Omega_{S/\mathbb{k}} \otimes_S R}{\text{image}\,(I)}$$

and the derivation $d : R \to \Omega$ defined by $d\overline{g}$ as the image of $dg \otimes 1$ for a representative $g \in \overline{g} = g + I$.

Then d defines a derivation $R \to \Omega$ and (Ω, d) is a universal derivation, since for any derivation $\delta \in Der(R, M)$ and any $f \in I$ the identity

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \delta(\overline{x}_i) = 0 \in M$$

is satisfied. Finally, the map $d : I \to \Omega_{S/\mathbb{k}} \otimes_S R$ has I^2 in the kernel, because $d(I^2) \subset I\Omega_{S/\mathbb{k}}$. □

Remark 8.1.11. Thus, for $I = \langle f_1, \ldots, f_r \rangle \subset S = \mathbb{k}[x_1, \ldots, x_n]$ and $R = S/I$ the module of Kähler differential is the cokernel

$$\Omega_{R/\mathbb{k}} \cong \text{coker}\,(R^r \to R^n)$$

where the map is defined by the jacobian matrix $(\frac{\partial f_j}{\partial x_i})$. In particular the module of Kähler differentials is defined over the field of definition of I.

Exercise 8.1.12. Compute $\Omega_{R/\mathbb{k}}$ and $Der_{\mathbb{k}}(R, R)$ for

1. $R = \mathbb{k}[x, y]/\langle y^2 - x^3 \rangle$,
2. $R = \mathbb{k}[x, y]/\langle y^2 - x^3 - x^2 \rangle$, and
3. $R = \mathbb{k}[x, y]/\langle y^2 - x^3 + x \rangle$.

Lemma 8.1.13. *Let U be a multiplicative subset of an $\mathbb{k}$-algebra R. The universal derivation $d : R \to \Omega_{R/\mathbb{k}}$ extends uniquely to a derivation $d : R[U^{-1}] \to \Omega_{R/\mathbb{k}}[U^{-1}]$ by the formula*

$$d(\frac{r}{t}) = \frac{tdr - rdt}{t^2}.$$

In particular, the formation of Kähler differentials commutes with localization: $\Omega_{R[U-1]/\mathbb{k}} \cong \Omega_{R/\mathbb{k}}[U^{-1}]$.

Proof. The map

$$R[U^{-1}] \to \Omega_{R/\Bbbk}[U^{-1}],\ \frac{r}{t} \mapsto \frac{tdr - rdt}{t^2}$$

is well-defined, because

$$\frac{r_1}{t_1} = \frac{r_2}{t_2} \in R[U^{-1}] \Leftrightarrow \exists t_3 \in U \text{ with } t_3(t_2r_1 - t_1r_2) = 0 \in R$$

gives $0 = t_3^2(t_2r_1 - t_1r_2) \in R \Rightarrow 0 = t_3^2(r_1dt_2 - r_2dt_1 + t_2dr_1 - t_1dr_2) \in \Omega_{R/\Bbbk} \Rightarrow 0 = t_3^2(t_1t_2(t_2dr_1 - t_1dr_2) - t_2^2r_1dt_1 + t_1^2r_2dt_2) \in \Omega_{R/\Bbbk} \Rightarrow$

$$\frac{t_1dr_1 - r_1dt_1)}{t_1^2} = \frac{t_2dr_2 - r_2dt_2}{t_2^2} \in \Omega_{R/\Bbbk}[U^{-1}].$$

The product rule and $\Bbbk$-linearity for the extension follow with a straight forward computation. Uniqueness of the extension holds, because the commutativity of the diagram

$$\begin{tikzcd}
R \arrow[r] \arrow[d] & \Omega_{R/\Bbbk} \arrow[d] \\
R[U^{-1}] \arrow[r] & \Omega_{R/\Bbbk}[U^{-1}]
\end{tikzcd}$$

implies

$$\frac{dr}{1} = d(\frac{t}{1}\frac{r}{t}) = \frac{r}{t}dt + td(\frac{r}{t})).$$

The universal property of $d : R[U^{-1}] \to \Omega_{R[U^{-1}]/\Bbbk}$ induces an morphism $\Omega_{R[U^{-1}]/\Bbbk} \to \Omega_{R/\Bbbk}[U^{-1}]$, whose inverse is induced by the map $dr \mapsto d(\frac{r}{1})$ □

A module M over a ring R is called **locally free** of rank r if $M_f \cong R_f^r$ holds for each elements $f \in T \subset R$ for a system T with $\langle T \rangle = R$.

Exercise 8.1.14. Let M be an R-module, and let r be a positive integer. The following are equivalent

1. M is locally free of rank r.
2. $M_\mathfrak{p} \cong R_\mathfrak{p}^r$ for each prime ideal $\mathfrak{p}$ or R.
3. $M_\mathfrak{m} \cong R_\mathfrak{m}^r$ for each maximal ideal $\mathfrak{m}$ or R.

An R-module M is projective, if for each exact sequence

$$N \to N'' \to 0$$

of R modules, the induced sequence

$$\mathrm{Hom}_R(M, N) \to \mathrm{Hom}_R(M, N'') \to 0$$

is exact as well. In other words, any morphism φ as below can be lifted to a morphism Φ:

$$\begin{tikzcd}
& R \arrow[dl, "\exists\,\Phi"'] \arrow[d, "\varphi"] & . \\
N \arrow[r] & N'' \arrow[r] & 0
\end{tikzcd}$$

Exercise 8.1.15. Let M be a projective module. Prove:

1. M is projective, iff M is a direct summand of a free module
2. If M is finitely generated module over the coordinate ring of affine variety, then M is projective iff M is locally free.

Exercise 8.1.16. Let $R = C^\infty(S^2, \mathbb{R})$ be the ring of C^∞-function on the 2-sphere and let M be the module of C^∞ vector fields on S^2. Prove that M is a locally free, but not a free R-module. Hint: Use some algebraic topology.

Remark 8.1.17. Inspired by the corresponding statement in the C^∞-category and holomorphic category Serre [195x] **?** conjectured that any locally free $\Bbbk[x_1, \dots, x_n]$ module is actually free. This was indepently proved by Quillen [197x] **?** and Suslin **?** [197x]. The commutative algebra book of Kunz Kunz (1985) culminates with a proof of this result.

Theorem 8.1.18. *Let $R = \Bbbk[X]$ be the coordinate ring of an affine variety X over an algebraically closed field $\Bbbk$. X is smooth of dimension d iff $\Omega_{R/\Bbbk}$ is locally free of rank d.*

Proof. Since R is finitely generated, $\Omega_{R/\Bbbk}$ is a finitely generated module. The images dx_i of $\Bbbk$-algebra generators x_i generate. The minimal number of generators of the localization of $\Omega_{R/\Bbbk}$ in a small Zariski neighbarhood $p \in X$ with maximal ideal $\mathfrak{m} = I_X(p)$ is $\dim_\Bbbk \Omega_{R/\Bbbk}/\mathfrak{m}\Omega_{R/\Bbbk}$. Since $Der_\Bbbk(R, R/\mathfrak{m}) \cong \operatorname{Hom}_R(\Omega_{R/\Bbbk}, R/\mathfrak{m}) = \operatorname{Hom}(\Omega_{R/\Bbbk}/\mathfrak{m}\Omega_{R/\Bbbk}, R/\mathfrak{m})$, the number of generators coincides with $\dim_\Bbbk \mathfrak{m}/\mathfrak{m}^2$ by Proposition 8.1.7. Thus, $\Omega_{R/\Bbbk}$ is locally minimally generated by $\dim_\Bbbk \mathfrak{m}/\mathfrak{m}^2 \geq \dim \mathcal{O}_p$ many elements, and it is locally free iff equality holds. □

Corollary 8.1.19. *Let X be an affine variety of dimension d over an algebraically closed field $\Bbbk$. Then $\Omega_{\Bbbk(X)/\Bbbk}$ is a $\Bbbk(X)$-vector space of dimension d.*

Proof. We may assume that X is affine. Then $\Omega_{\Bbbk(X)/\Bbbk} \cong \Omega_{\Bbbk[X]/\Bbbk}[U^{-1}]$, where $U = \Bbbk[X] \backslash \{0\}$ by Lemma 8.1.13. By Proposition 8.1.18 the modue $\Omega_{\Bbbk[X][f^{-1}]/\Bbbk}$ is free for rank d for a suitable $f \in \Bbbk[X]$. Hence, $\Omega_{\Bbbk(X)/\Bbbk}$ is a vectorspace of dimension d. □

Lemma 8.1.20. *Suppose R is a finitely generated $\Bbbk$-algebra. Then the formation of the derivation modules commutes with localization: If $U \subset R$ is a multiplicative subset then*

$$Der_\Bbbk(R, M)[U^{-1}] \cong Der_\Bbbk(R[U^{-1}], M[U^{-1}])$$

Proof. We have to prove

$$\mathrm{Hom}_R(\Omega_{R/\Bbbk}, M)[U^{-1}] \cong \mathrm{Hom}_{R[U^{-1}]}(\Omega_{R/\Bbbk}[U^{-1}], M[U^{-1}]).$$

Since R is finitely generated algebra over $\Bbbk$, the module of Kähler differentials $\Omega_{R/\Bbbk}$ is finitely generated as an R-module. Consider a finite free presentation

$$F_1 \to F_0 \to \Omega_{R/\Bbbk} \to 0$$

Then $\mathrm{Hom}_R(F_i, M)[U^{-1}] \cong \mathrm{Hom}_{R[U^{-1}]}(F_i[U^{-1}], M[U^{-1}]$ because the F_i are finite free modules. The exact sequence

$$0 \to \mathrm{Hom}(\Omega_{R/\Bbbk}, M) \to \mathrm{Hom}(F_0, M) \to \mathrm{Hom}(F_1, M)$$

localized at U yields an isomorphism

$$\begin{aligned}\mathrm{Hom}_R(\Omega_{R/\Bbbk}, M)[U^{-1}] &\cong \ker(\mathrm{Hom}(F_0[U^{-1}], M[U^{-1}] \to \mathrm{Hom}(F_1[U^{-1}], M[U^{-1}])\\ &\cong \mathrm{Hom}_{R[U^{-1}]}(\Omega_{R/\Bbbk}[U^{-1}], M[U^{-1}])\end{aligned}$$

□

Exercise 8.1.21. Let R be a $\Bbbk$-algebra and $r \in R$ an element, which is separable algebraic over $\Bbbk$. Prove $dr = 0 \in \Omega_{R/\Bbbk}$. □

Exercise 8.1.22. Consider $\Bbbk = \mathbb{F}_p(x)$ and $L = \Bbbk[t]/\langle t^p - x\rangle$. Prove that

$$\Omega_{L/\Bbbk} \cong Ldt \cong L,$$

although $L/\Bbbk$ an algebraic field extension. *Hint*: Observe that $\frac{\partial}{\partial t}$ gives a $\Bbbk$-derivation on L. □

Exercise 8.1.23. Let $\Bbbk$ be any field and $L/\Bbbk$ a finitely generated field extension of transzendence degree d. Prove that $\Omega_{L/\Bbbk}$ is an L-vector space of dimension $\geq d$, and equality holds iff $L/\Bbbk$ has a separable transzendence basis. □

8.2 Adjoint Curves

Let C be a smooth absolutely irreducible projective curve. Perhaps the simplest way to calculate the degree of a canonical divisor is with the help of a plane model. We choose a birational morphism $\eta : C \to C' \subset \mathbb{P}^2$ onto a plane curve with only ordinary singularities. Recall 7.2.29, that the geometric genus g of such plane model is defined as

$$g = \binom{d-1}{2} - \sum_{q \in C'} \binom{r_q}{2}.$$

Proposition 8.2.1. *Let C be an absolutely irreducible smooth projective curve. A canonical divisor on C has degree*

$$\deg K = 2g - 2.$$

In particular, the geometric genus g of C is well-defined, i.e. independent of the choice of the plane model with only ordinary singular points.

Proof. Let $C' \subset \mathbb{P}^2$ a plane model of C with only ordinary multiple points. We choose coordinates on $\mathbb{P}^2$ general, and compute the divisor of dx for $x = \frac{x_0}{x_2}$. Let $f(x_0, x_1, x_2) \in \Bbbk[x_0, x_1, x_2]$ be the homogeneous equation of C. The line at infinity $L = \mathrm{V}(x_2)$ intersects C' in $d = \deg C'$ points transversally, $o = [0 : 1 : 0] \notin C'$, and the point o does not lie on any tangent line to singular points, since we have a general coordinate system. Since $\Bbbk(C') \cong \Bbbk(C)$ we can regard rational functions as functions both on C and C'. We do this freely. Moreover for a smooth points p of C' we denote with the same letter p the corresponding smooth points on C, and we cidentify $\mathcal{O}_{C',p} \cong \mathcal{O}_{C,p}$.

The rational function $w = 1/x$ is a local parameter at all points p on the line at infinity, hence $dx = d(\frac{1}{w}) = \frac{-1}{w^2}dw$ has a pol of order 2 at every point at infinity and is regular elsewhere. At a smooth points $p \in C'$ the differential form dx has a zero iff the line $\mathrm{V}(x_0 - \lambda x_2)$ for $\lambda = \frac{x_0}{x_2}(p)$ is tangent to C', because these are the points where $x - \lambda \in \mathfrak{m}_{C,p} \subset \mathcal{O}_{C,p}$ is not a local parameter. From our proof of Bézout's Theorem on plane curves (Section 5.4.8) we know, that these points are among the intersections of C' with the curve defined by $f_{x_1} = \frac{\partial f}{\partial x_1}$. To compare multiplicities, we use affine coordinates $x = x_0/x_2$ and $y = x_1/x_2$ around such an intersection point $p = [a : b : 1] \in C'$. Then $x - a$ defines the tangent line and $y - b$ is a local parameter of C' at p. The affine equation $f^a(x, y) = f(x, y, 1)$ gives us the relation $dx = -(f^a_y/f^a_x)dy$. Since $f^a_y = f_{x_1}/x_2^{d-1}$ and $f^a_x(p) \neq 0$, we obtain

$$v_p(dx) = v_p(f^a_y/f^a_x) = v_p(f^a_y) = i(f, f_{x_1}; p).$$

$\mathrm{V}(f_{x_1})$ intersects C' also at singular points. Let $\eta\colon C \to C'$ be the birational morphism. For a point $p \in C$ we denote by $r_p = \mathrm{mult}(C', q)$ the multiplicity of C' in $q = \eta(p)$. Then

$$E = \sum_{p \in C} (r_p - 1)p$$

is called the **divisor** of **multiple points** of C'. Since C' has only ordinary singularities we have

$$\deg E = \sum_{q \in C'} (r_q - 1)r_q.$$

The final result of our computation will be

$$(dx) + 2\eta^*(x_2) = \eta^*(f_{x_1}) - E.$$

It remains to verify the formula at points $p \in C$ over singular points $q = \eta(p) = [a : b : 1]$ of C'. Since $\mathrm{V}(x-a)$ is not a tangent line at q, we have that $x-a$ is a local parameter and so $v_p(dx) = 0$. On the other hand $v_p(f_{x_1}) = r-1$ for $r = r_q$, because in a factorization of the tangent cone of f at q

$$f^a \equiv \prod_{j=1}^{r} l_j \quad \mathrm{mod}\ \langle x-a, y-b\rangle^{r+1}$$

precisely one factor l_j is in $\langle x-a\rangle^2 \mathcal{O}_{C,p}$. Indeed, say l_r is this factor, then

$$\frac{\partial f^a}{\partial y} \equiv \frac{\partial l_r}{\partial y} \prod_{j=1}^{r-1} l_j \quad \mathrm{mod}\ \langle x-a\rangle^r \mathcal{O}_p$$

and the desired multiplicity follows, since $\partial l_r/\partial y \in \Bbbk$ and $v_p(l_j) = 1$ for $j < r$.

The degree formula follows now, since $\deg(dx) = \deg \eta^*((f_{x_1}) - 2\deg(x_2)) - \deg E = d(d-1) - 2d - \sum_{q \in C'} r_q(r_q - 1)$. Hence $\deg K = 2g-2$, since $g = \binom{d-1}{2} - \sum_{q\in C'} \binom{r_q}{2}$ by the definition of the geometric genus of a plane model 7.2.29. □

As a corollary of the proof we note

Theorem 8.2.2 (Adjuction formula). *Let C be a smooth projective curve, $\eta : C \to C' \subset \mathbb{P}^2$ a birational map to a plane model of degree d with only ordinary singularities and $E = \sum_{p\in C}(r_p - 1)p$ the multiple point divisor. Let $H = \eta^* L$ for a general line $L \subset \mathbb{P}^2$, denote the hyperplane divisor. Then the linear equivalence class of the canonical divisor is*

$$K \equiv (d-3)H - E.$$

Proof. $\eta^*((f_{x_1}) - 2(x_2)) \in |(d-1)H|$ because $(f_{x_1}) - 2(x_2) \equiv (d-3)L$ on $\mathbb{P}^2$ □

Corollary 8.2.3.

$$\ell(K) \geq g$$

Proof. Let $q_1, \dots q_s \in C'$ denote the singular points. The divisor of a form $g \in L(d-3; (r_1-1)q_1, \dots, (r_s-1)q_s)$ pulls back to a divisor $\eta^*(g) = D + E$. Hence the part $D \in\mid (d-3)H - E \mid \cong \mid K \mid$. Hence from

$$\dim L(d-3; (r_1-1)q_1, \dots, (r_s-1)q_s) \geq \binom{d-1}{2} - \sum_{j=1}^{s} \binom{r_j}{2} = g$$

we obtain a g-dimensional subspace of $L(K)$. We will see later, that equality holds. In particular, it will follows then that the singular points points impose independent conditions on forms of degree $d-3$. □

Definition 8.2.4. Let $C' \subset \mathbb{P}^2$ be an irreducible plane curve of degree d with only ordinary multiple points $q_1, \dots, q_s$ of multiplicity $r_1, \dots, r_s$. An **adjoint curve** of C' is a curve $G \not\supset C'$ of degree $d-3+a$ with multiplicity $\geq r_j - 1$ in q_j for $j = 1, \dots, s$ and $a \geq 0$.

Every adjoint curve cuts on C a divisor of class $|K+aH|$ by the adjunction formula 8.2.2. Notice, that a curve $G \not\supset C'$, such that $(G)-E \geq 0$ is an adjoint curve. Indeed, assume the contrary, that is, assume that $\operatorname{mult}(G,q) < r_q - 1$ at some singular point q of C'. Then $(G)-E$ has a positive coefficient at every point $p \in \eta^{-1}(q) \subset C$ iff each tangent line of C' at q is also a tangent to G at q. This is impossible, since C' has r_q different tangent lines and $\operatorname{mult}(G,q) < r_q - 1 < r_q$.

Theorem 8.2.5 (Completeness of the adjoint systems). *Let C, C' and E be as above. Suppose $D \equiv D'$ are two effective linearly equivalent divisors. Suppose G is an adjoint curve of degree b such that $(G) = D+E+R$ for some effective divisor R. Then, there exists an adjoint curve G' of degree b, such that $(G') = D'+E+R$.*

In other words the linear system of adjoint curves with additional base points in R cuts out the complete linear system $|D|$.

Proof. By definition of the rational function field and linear equivalence there exist curves H, H' of the same degree, such that

$$D + (H) = D' + (H').$$

Then $(GH) = (H')+D'+E+R \geq (H')+E$. Let F be the equation of C'. We apply Noether's AF+BG Theorem 5.5.1 to F, H' and GH. The local condition is satisfied by the above inequality on divisors and Proposition 7.5.20. Thus, there exist A and B, such that $AF + BH' = GH$. So $G' = B$ is the desired adjoint curve of degree b, because $(B) = (GH) - (H') = D' + E + R$. □

Corollary 8.2.6. *The geometric genus of a plane model $C' \subset \mathbb{P}^2$ with only ordinary singularities and the arithmetic genus of a smooth projective model $C \subset \mathbb{P}^r$ coincide, i.e.*

$$g = p_a.$$

Proof. $|mH - E|$ is cut out by adjoint curves by the theorem. We bound $\ell(mH - E)$ for $m \geq d$

$$\begin{aligned} \ell(mH-E) &\geq \binom{m+2}{2} - \binom{m-d+2}{2} - \sum_{q \in C'} \binom{r_q}{2} \\ &= \deg(mH-E) + 1 - g \end{aligned}$$

with equality for m sufficiently large by Theorem 5.3.11, which says that multiple points impose independent conditions on forms of sufficiently high degree. On the other hand Proposition 7.4.14 says, that $\ell(mH-E) = \deg(mH-E) + 1 - p_a$ for m sufficiently large. So $g = p_a$. □

Corollary 8.2.7 (Riemann's inequality, second version).

$$\ell(D) \geq \deg D + 1 - g.$$

Exercise 8.2.8. Give a proof of the second version of Riemann's inequality, which uses only the adjunction formula 8.2.2. □

8.3 The Riemann-Roch Formula

Remark-Definition 8.3.1. Let D be a divisor on C. Then

$$\Omega(D) = \{\omega \in \Omega(C) | (\omega) + D \geq 0\}$$

denotes the $\mathbb{k}$-vector space of differential forms with poles up to order D. If $K = (\omega_0)$ is a fixed canonical divisor then

$$L(K + D) \equiv \Omega(D)$$

via $f \mapsto f\omega_0$. In particular, we see that $\Omega(D)$ is finite-dimensional with the dimension bounded by $0 \leq \ell(K + D) \leq \deg K + \deg D + 1$ by 7.3.10.

Theorem 8.3.2 (Riemann-Roch). *Let C be a smooth projective curve of genus g and D a divisor on C. Then*

$$\ell(D) - \ell(K - D) = \deg D + 1 - g$$

We will prove the Theorem at the end of this section. First we give some applications.

The information on the dimension and the degree of the canonical system is part of the Riemann-Roch theorem.

Corollary 8.3.3.

$$\ell(K) = g.$$

Proof. This follows from the case $D = 0$, since $L(0) = \mathbb{k}$. □

Thus, another way to define the genus is as the maximal number of $\mathbb{k}$ linearly independent regular differential forms.

Corollary 8.3.4.

$$\deg K = 2g - 2.$$

Proof. This follows from the case $D = K$ and $\ell(K) = g$. □

Corollary 8.3.5. ?? *Let D be a divisor on C. Then*

$$\ell(D) = \deg D + 1 - g$$

if $\deg D > 2g - 2$. *In particular,*

1. *$|D|$ is base point free, if* $\deg D \geq 2g$.
2. *$|D|$ is very ample, if* $\deg D \geq 2g+1$.

Proof. $L(K-D) = 0$, because $\deg(K-D) < 0$, c.f. 7.3.10. □

Corollary 8.3.6. *Let $\mathbb{k}$ be algebraically closed. Then*

(1) Every smooth projective curve of genus $g = 0$ is isomorphic to $\mathbb{P}^1$.

(2) Every smooth projective curve of genus $g = 1$ is isomorphic to a smooth plane cubic.

Proof. Since $\mathbb{k}$ is algebraically closed, we can find a $\mathbb{k}$-rational point on these curves. $|p|$ defines an isomorphism in case (1), $|3p|$ an isomorphism onto a plane cubic curve in case (2). □

Exercise 8.3.7. Let C be a curve of genus 0 over a not necessarily algebraically closed field. Prove:

1. C is isomorphic to a smooth plane conic by an isomorphism, which defined over the field of definition.
2. A conic is rational, i.e. isomorphic to $\mathbb{P}^1$ iff it contains a $\mathbb{k}$-rational point.

□

Corollary 8.3.8. *Let K be a canonical divisor on a smooth projective curve C of genus g. Then*

(1) $|K|$ is base point free, if $g \geq 1$.

(2) $|K|$ is very ample, unless there exists a morphism $C \to \mathbb{P}$ of degree 2 onto a curve $\mathbb{P}$ of genus 0.

Proof. (1) $L(p)$ is one dimensional, because otherwise we get a morphismen $C \to \mathbb{P}^1$ of degree 1, which is necessarily an isomorphism. This contradicts $g \geq 1$. Hence $L(K-p) \subset L(K)$ has codimension 1 for every $p \in C$. (2) The space $L(K-p-q) \subset L(K)$ has only codimension 1 iff $L(p+q)$ is two dimensional, equivalently, iff $|p+q|$ defines a morphism $C \to \mathbb{P}^1$ of degree 2. □

Exercise 8.3.9. Let C be a smooth projective curve, $r \geq 1$ and $p_1, \dots, p_r \in C$ finitely many points. Prove that $C \setminus \{p_1 \dots, p_r\}$ is an affine curve. □

Our proof of the Riemann-Roch Theorem is based on the following fundamental Lemma.

Lemma 8.3.10 (Noether's Reduction Lemma). *Let D be an effective divisor and p a point on a smooth curve C. Then*

$$\ell(K-D-p) < \ell(K-D)$$

implies

$$\ell(D+p) = \ell(D).$$

We give two proofs, an analytic one, which works for $\Bbbk = \mathbb{C}$, and an algebraic proof based on the completeness of the adjoint system, and hence on Noether's AF+BG theorem.

Proof (Proof in case $\Bbbk = \mathbb{C}$*).* Suppose $\ell(K-D-p) < \ell(K-D)$ and $\ell(D+p) > \ell(D)$. Consider a rational differential form $\omega \in L(K-D) \setminus L(K-D-p)$ and a rational function $f \in L(D+p) \setminus L(D)$. Then

$$f\omega \in L(K+p) \setminus L(K)$$

is a meromorphic differential form with a simple pole at p and otherwise holomorphic. But this is impossible. Integration around a small disc Δ around p gives

$$\int_{\partial \Delta} f\omega = 2\pi i \operatorname{Res}_p(fw) \neq 0$$

by the residue theorem. On the other hand,

$$\int_{\partial \Delta} f\omega = -\int_{C \setminus \Delta} d(f\omega) = 0$$

by Stokes, because $d(f\omega) = \bar{\partial}(f\omega) = 0$, since $f\omega$ is a holomorphic form on $C \setminus \Delta$. This is a contradiction. □

Proof (General case). Consider a plane model C', which has only ordinary singularities disjoint from the support of D and p, cf. Theorem 7.2.23. Let d denote the degree of the plane curve and $E = \sum_{p \in C}(r_p - 1)p$ be the divisor of multiple points. Let $H = C.L_1 = \sum_{i=1}^{d} p_i$ be a hyperplane section of C' through distinct smooth points. Since the statement depends only on the linear equivalence class of K, we may assume

$$K = (d-3)H - E$$

by the adjunction formula 8.2.4, and then $L(K-D) \subset L((d-3)H - E)$.

By assumption, there exists an $h \in L(K-D)$ with $h \notin L(K-D-p)$. The rational function h is a restriction

$$h = g/z_1^{d-3},$$

where g defines an adjoint curve and z_1 is the defining equation of L_1 by the completeness of the adjoint system 8.2.5. Then $(g) = D + E + A$ with $A \geq 0$ but $A \ngeq p$. We take a line $L_2 = V(z_2)$ through p, such that $L_2.C' = p + B$ consists of d distinct smooth points. Then

$$(z_2 g) = (D+p) + E + (A+B).$$

Given now $f \in L(D+p)$, we have to show $f \in L(D)$. If we write $(f) + D = D'$ then this means that we have to prove, that D' is effective. The only possible negative term in D' is the coefficient of p. Since $D+p \equiv D'+p$ and both are effective divisors, we can apply the completeness of the adjoint systems 8.2.5. There exists a curve M of degree $d-2$ with $M.C' = (D'+p)+E+(A+B)$. But B consists of $d-1$ collinear points and M is a curve of degree $d-2$. Thus by Bezout's Theorem 5.4.8, L_2 is a component of M. Thus, $(M-L_2).C' = D'+A+E$ is effective. Since $A-p$ and $E-p$ are not effective, the coefficient of p in D' is non negative, and D' is effective. This proves the Reduction Lemma. □

Proof of the Riemann-Roch theorem. We have to prove the equation

$$(*)_D \qquad \ell(D) = \deg D + 1 - g + \ell(K-D)$$

for every divisor D.

Case 1: $\ell(K-D) = 0$. (Non-special divisors). Induction on $\ell(D)$. If $\ell(D) = 0$ then Riemann's inequality 8.2.7 for D and $K-D$ gives $0 = \ell(D) \geq \deg D + 1 - g$ and $0 = \ell(K-D) \geq \deg(K-D)+1-g = 2g-2-\deg D+1-g \geq -(\deg D + 1 - g)$. For the second inequality we used $\deg K = 2g-2$ proved Proposition 8.2.1. Combining both inequalities give the the desired formula. If $\ell(D) = 1$, we may assume that D is effective. Then

$$\ell(K) \leq \ell(K-D) + \deg D = \deg D$$

by 7.3.10 and our assumption $\ell(K-D) = 0$. Moreover

$$g \leq \ell(K)$$

by 8.2.4. This implies $g \leq \deg D$. So Riemann's inequality $1 \geq \deg D + 1 - g$ is an equality. If $\ell(D) > 1$ then we can choose a point p such that $\ell(D-p) = \ell(D)-1$. By the Reduction Lemma, this implies $\ell(K-D+p) = \ell(K-D) = 0$ and the formula $(*)_D$ follows from $(*)_{D-p}$, which we know by induction.

Case 2: $\ell(K-D) > 0$ (special divisors). If $\ell(D) = 0$ then the formula follows from case 1 applied to $(*)_{K-D}$: $K-D$ is non-special, since $\ell(K-(K-D)) = \ell(D) = 0$. So $(*)_{K-D}$ holds, and $\ell(K-D) = \deg(K-D)+1-g = g-1-\deg D$. This is equivalent to $(*)_D$, because $\ell(D) = 0$. If $\ell(D) > 0$, we may assume that D is effective. We apply induction on $\ell(K-D)$ and the Reduction Lemma. Choose p such that $\ell(K-D-p) = \ell(K-D)-1$. Then $\ell(D+p) = \ell(D)$ by 8.3.10 and $(*)_{D+p}$, which we know by induction, implies $(*)_D$. □

Corollary 8.3.11. *Let D be any divisor and p any point. Then, either*

$$\ell(K-D-p) = \ell(K-D) - 1 \ \text{ and } \ \ell(D+p) = \ell(D)$$

or

$$\ell(K-D-p) = \ell(K-D) \ \text{ and } \ \ell(D+p) = \ell(D)+1$$

holds. □

8.4 Hurwitz's Formula

Definition 8.4.1. Let $\varphi : C \to E$ a morphism of curves of **degree** d, i.e. $[k(C) : k(E)] = d$. φ is called **separabel** respectively **purely inseparable** if the field extension $\Bbbk(C) \supset \Bbbk(E)$ is separable respectivley purely inseparable.

An arbitrary dominant morphism $C \to E$ factors through a separable morphism $C' \to E$ and a pure inseparable morphism $C \to C'$. This follows from Corollary 7.1.13, because the corresponding statement holds algebraic field extensions.

We want to calculate, how the genus changes under a morphism. We treat the separable case first. Along the computation of the degree of a canonical divisor we performed such a calculation in a special case of a projection $C \to \mathbb{P}^1$.

Definition 8.4.2. Let $\varphi : C \to E$ be a morphism. Let $p \in C$, $q = \varphi(p) \in E$ and $t \in \mathfrak{m}_q \subset \mathcal{O}_{E,q}$ a local parameter. The **ramification index** of φ at p is defined as

$$e_p = v_p(\varphi^* t).$$

φ is **unramified** at p if $e_p = 1$ and **ramified** otherwise. The ramification is called **tame** if $\operatorname{char} \Bbbk$ does not divide e_p and **wild** otherwise. If φ is separable then we define furthermore the **differential index** as

$$\rho_p = v_p(dt),$$

where we regard dt as a rational differential form on C, and we call

$$R = \sum_{p \in C} \rho_p p$$

the **ramification divisor** of φ. A point in $\operatorname{supp}(R)$ is called a **ramification point**, the image in E a **branch point**.

Remark 8.4.3. 1. e_p is independent from the choice of a generator $t \in \mathfrak{m}_q$. For a different generator, say $t' = ut$ with $u \in \mathcal{O}_q$ a unit, we have $v_p(\varphi^* t') = v_p(\varphi^* t)$, because $\varphi^* u \in \mathcal{O}_p$ is also a unit.

2. $\rho_p \geq e_p - 1$ and equality holds iff the ramification is tame. Indeed is $t = us^e$ with $u \in \mathcal{O}_p$ a unit, then $dt = (ues^{e-1} + u's^e)ds$ and $v_p(dt) = e - 1$ iff $\operatorname{char} \Bbbk \not| e$. Hence,

$$R = \sum_{p \in C} (e_p - 1)p,$$

if φ has only tame ramification.

3. Since φ is separable, R is indeed a divisor. $t \notin \Bbbk/C)^p\Bbbk$ because $\Bbbk(E) \subset \Bbbk(C)$ is separable, Hence, $v_{p'}(dt) = v_{p'} d(t - t(p')) = 0$ for all but finitely many points, and the assertion follows since $t - t(q')$ for $q' = \varphi(p')$ is a local parameter for points q' in a Zariski open neighborhood of q.

Exercise 8.4.4. Prove that the ramification divisor is defined over the field of definition of φ. □

It is convenienent to introduce pullback and pushforward of divisors.

Definition 8.4.5. Let $\varphi : C \to E$ be a morphism between smooth projective curves. Then we define

$$\varphi^* : \mathrm{Div}(E) \to \mathrm{Div}(C), \quad \sum_{q \in E} n_q q \mapsto \sum_{p \in C} e_p n_{\varphi(p)} p$$

and

$$\varphi_* : \mathrm{Div}(C) \to \mathrm{Div}(E), \quad \sum_{p \in C} n_p p \mapsto \sum_{p \in C} n_p \varphi(p).$$

Note, that $\varphi_* \varphi^* D = dD$ for every divisor $D \in \mathrm{Div}(E)$, because $\sum_{p \in \varphi^{-1}(q)} e_p = d$ for every point $q \in E$. By the same reason, we have $\deg \varphi^* D = d \deg D$ for $D \in \mathrm{Div}(E)$ and $\deg \varphi_* D' = \deg D'$ for $D' \in \mathrm{Div}(C)$.

Theorem 8.4.6 (Hurwitz's formula). *Let $\varphi : C \to E$ be a separable morphism of degree d between smooth irreducible curves of genus g_C and g_E. Then*

$$K_C \equiv \varphi^* K_E + R,$$

in particular

$$2g_C - 2 = d(2g_E - 2) + \deg R,$$

where $R = \sum_{p \in C} \rho_p p$ denotes the ramification divisor of φ.

Proof. Since φ is separable, a non-zero rational differential form $\omega = f dg$ on E pullsback to a nonzero rational diferential from $\varphi^* \omega$, defined as

$$\varphi^* \omega = (g \circ \varphi)\, d(f \circ \varphi).$$

We choose ω such that the support of the divisor (ω) contains no branch point. Then $\varphi^* \omega$ has zeroes and poles with the same multiplcity at each of the d point over a point in the support of the divisor (ω) and additional zeroes at the points p in the support of R with multiplicity ρ_p. Hence $K_C = \varphi^* K_E + R$ and $2g_C - 2 = \deg K_C = d(\deg K_E) + \deg R = d(2g_E - 2) + \deg R$. □

Hurwitz's Theorem allows us to give a purley topological interpretation of the genus of smooth projective curves over $\mathbb{C}$. This explains the attribute "geometric" in the notion "geometric genus".

Corollary 8.4.7. *Let C be a smooth projective curve over $\mathbb{C}$ and consider the underlying real 2-dimensional topological manifold, that is C equipped with the Euclidean topology. Then*

$$2 - 2g = euler(C)$$

where g is the geometric genus and $euler(C)$ denotes the Euler number of the underlying surface.

Proof. The Euler number is the alternating sum of the number of i-simplices in any triangulation. Euler observed the independence from the triangulation and calculated the value for S^2,

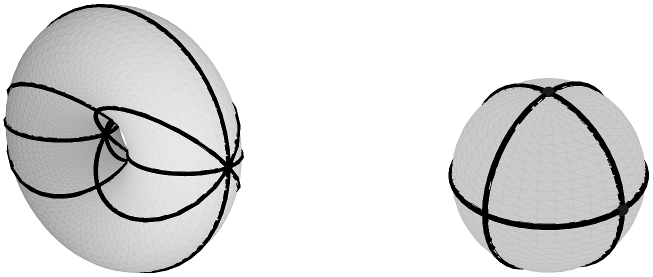
$$euler(S^2) = v - e + t = 2,$$

for v, e, t the number of vertices, edges and triangles in an arbitrary triangulation of S^2. Since $\mathbb{P}^1$ has genus 0 the formula holds for $\mathbb{P}^1$. In general, consider a dominant morphism $\varphi : C \to \mathbb{P}^1$. Refining the triangulation of S^2 we may assume, that every branch point is among the vertices, and that the preimage of each triangle is homoeomorphic to a union of triangles, which intersect at most in common ramification points. Then the preimages of the triangulation of S^2 give a triangulation of C. With $d = \deg \varphi$ we obtain $\tilde{t} = td$ triangles and $\tilde{e} = ed$ edges in the induced triangulation of C. However, the number of vertices is only $\tilde{v} = vd - \deg R$ because of the ramification. Hence, the Euler number is

$$euler(C) = \tilde{v} - \tilde{e} + \tilde{t} = (v - e + t)d - \deg R = 2d - \deg R = 2 - 2g,$$

by Hurwitz's formula. □

Example 8.4.8. Consider the elliptic curve $E = \mathrm{V}(y^2 - x^3 + x)$ and the projection onto the x-axis $\varphi : E \to \mathbb{P}^1$. Then φ has degree 2 and branch points in $\{0, 1, -1, \infty\}$. We triangulate $\mathbb{P}^1 \cong S^2$ like an octahedron.



The induced triangulation of E has 8 vertices, namely the 4 ramification points, which have now each valence 8 in the 1-skeleton, and the $2 \cdot 2$ preimages of $\pm i$, which remain vertices of valence 4.

Exercise 8.4.9. In the example above assume that $\infty \in \mathbb{P}^1$ correspond to the north pol, that $1 \in \mathbb{P}^1$ and its preimage in E correspond to the right most points. Describe a possibility for the identification of the ramification points on E in the illustration and the curve of real points. □

Corollary 8.4.10 (Lüroth's Theorem, first version). *Let $\varphi : \mathbb{P}^1 \to E$ be a separable morphism to a smooth curve E. Then $E \cong \mathbb{P}^1$.*

Proof. The only way how the left hand side in Hurwitz's formula can be negative, is when $g_E = 0$ and $\deg R = 2d - 2$. So E has genus 0, and since it contains a point, defined over of field of definition of φ, it is rational by Theorem 5.4.13. □

We now come to inseparable morphisms. Let $\Bbbk$ be a field of characteristic $p > 0$.

$$F = F_{\Bbbk} \colon \Bbbk \to \Bbbk,\ a \mapsto a^p$$

is a monomorphism of fields, which is called the **Frobenius morphism**. In case $\Bbbk$ is algebraically closed or $\Bbbk$ is finite, it is an automorphism, i.e. also surjective.

Inseparable morphisms $C' \to C$ of curves are closely related to the Frobenius morphism on the function field

$$F = F_{\Bbbk(C)} : \Bbbk(C) \to \Bbbk(C).$$

However, F does not correspond to a morphism of curves in the sense of Corollary 7.1.13, because F is not $\Bbbk$-linear. To make F $\Bbbk$-linear we change the $\Bbbk$-algebra structure on the right $\Bbbk(C)$. We define

$$\Bbbk \times \Bbbk(C) \to \Bbbk(C)$$

by

$$(a, f) \mapsto a^p f.$$

With this new structure $\Bbbk(C)$ becomes the function field of a new curve C_p, and the map F induces a morphism

$$F' : C_p \to C,$$

which we call the **geometric Frobenius morphism**. To describe equations of C_p, we consider the ring homorphism

$$F^{-1} : \overline{\Bbbk}[x_0, \dots, x_n] \to \overline{\Bbbk}[x_0, \dots, x_n],$$

defined by applying $F_{\overline{\Bbbk}}^{-1}$ to the coefficients. Suppose $C \subset \mathbb{P}^n$ is defined over $\Bbbk$ by the ideal $I = \langle f_1, \dots, f_r \rangle$. Then C_p has $\Bbbk^{1/p}$ as the field of definition and $F^{-1}(f_1), \dots, F^{-1}(f_r)$ as defining equations. More over, F' is the morphism induced by

$$\mathbb{P}^n \to \mathbb{P}^n,\ x_i \mapsto x_i^p,$$

on C_p. Indeed, $f_j(x_0^p, \dots, x_n^p) = (\,F^{-1}(f_j)(x_0, \dots, x_n)\,)^p$ vanishes on C_p for every defining equation $f_j \in \mathrm{I}(C)$. Similarly, on the level of function fields, we have $(F')^*(f) = (F^{-1}(f))^p \in \overline{\Bbbk}(C_p)$ for any $f \in \overline{\Bbbk}(C)$. Hence, F' is purely inseparable and $\overline{\Bbbk}(C_p) = \overline{\Bbbk}(C)^{1/p}$. Since the ramification index $e_q = e_q(F') = p$ and F' is bijective as a map on sets, we have $\deg F' = p$.

Theorem 8.4.11. *Let $\varphi : C' \to C$ be a purely inseparable morphism of smooth curves. Then φ is a composition of geometric Frobenius morphisms. In particular $g_{C'} = g_C$.*

Proof. $\deg \varphi = p^r$ is a p^{th} power, because the field extension is purely inseparable. Hence,

$$\overline{\Bbbk}(C')^{p^r} \subset \overline{\Bbbk}(C)$$

or equivalently,

$$\overline{\Bbbk}(C') \subset \overline{\Bbbk}(C)^{1/p^r}.$$

On the other hand, if we define inductively $C_{p^{j+1}} = (C_{p^j})_p$ then the composition

$$\begin{tikzcd} C_{p^r} \arrow[r, "F'"] & C_{p^{r-1}} \arrow[r, "F'"] & \cdots \arrow[r, "F'"] & C_p \arrow[r, "F'"] & C \end{tikzcd}$$

is another morphism of degree p^r and $\overline{\Bbbk}(C_{p^r}) = \overline{\Bbbk}(C)^{1/p^r}$. Hence $\overline{\Bbbk}(C') = \overline{\Bbbk}(C)^{1/p^r}$ for degree reasons. We conclude that $C' = C_{p^r}$ and that φ is the composition of the geometric Frobenius morphism from Corollary 7.1.13. □

Exercise 8.4.12. Suppose, C is defined over the finite field $\mathbb{F}_{p^r}$. Prove that $(F')^r$ is an automorphism of C. Is the converse true?

Hint: To answer the question, consider $\mathbb{F}_4 = \mathbb{F}_2[a]$ and the plane curve, defined by $ax^3 + (a+1)y^3 + z^3$. □

Theorem 8.4.13 (Lüroth's Theorem, final version). *Let $\Bbbk$ be a not necessarily algebraically closed field, and let $\Bbbk \subset L \subset \Bbbk(t)$ be a field of transzendence degree $\operatorname{trdeg}_{\Bbbk} L = 1$. Then $L \cong \Bbbk(s)$ is also a purely transzendental extension of $\Bbbk$.*

Proof. $L \subset \Bbbk(t)$ corresponds to a morphism $\mathbb{P}^1 \to C$ of curves defined over $\Bbbk$. We decompose $\mathbb{P}^1 \to C$ in a purely inseparable part $\mathbb{P}^1 \to C'$ and a separable part $C' \to C$. Then, we see $g_{C'} = 0$ from Theorem 8.4.11 and $g_C = 0$ by Corollary 8.4.10. Since C contains a $\Bbbk$-rational point, we obtain $C \cong \mathbb{P}^1$ over $\Bbbk$ and hence, $L = \Bbbk(C) \cong \Bbbk(s)$ for some transzendental element $s \in L$. □

Our second application of Hurwitz' formula is the proof of the Pücker's formulas: Let $\operatorname{char} \Bbbk = 0$, and let $C \subset \mathbb{P}^2$ be a smooth pojective curve of degree d. We want to compute the degree d^* of the dual curve $\check{C} \subset \check{\mathbb{P}}^2$. Intersection points of $\check{C} \cap L$ with a line L correspond to tangent lines of C, which pass through the point $p \in \mathbb{P}^2$ dual to L. Thus, applying the Hurwitz formula to the projection from p allows to compute d^*. The dual curve $\check{C}$ usually has double points and cusps corresponding to bitangents and flex tangents respectively. To arrive at formulas which are symmetric in the data from C and $\check{C}$, we allow ordinary double ppints and cusps also for C.

Theorem 8.4.14. *Let $C \subset \mathbb{P}^2$ and $\check{C} \subset \check{\mathbb{P}}^2$ a pair of dual curves with only ordinary nodes and cusps. Let δ, κ, b and f denote the number of nodes, cusps, bitangents and flexes of C, respectively. Then the corresponding numbers of*

$\check{C}$ are $\delta^* = b, \kappa^* = f, b^* = \delta$ and $f^* = \kappa$. The degrees d and d^* of C and C^* and their common geometric genus g are related by

$$g = \binom{d-1}{2} - \delta - \kappa = \binom{d^*-1}{2} - b - f,$$

$$d^* = d(d-1) - 2\delta - 3\kappa,$$

$$d = d^*(d^*-1) - 2b - 3f.$$

Moreover,

$$f = 3d(d-2) - 6\delta - 8\kappa,$$

$$\kappa = 3d^*(d^*-2) - 6b - 8f,$$

and

$$b = \frac{d(d-2)(d^2-9)}{2} - d(d-1)(2\delta+3\kappa) + \frac{(2\delta+3\kappa)^2}{2} + \frac{20\delta+25\kappa}{2}.$$

Proof. The first formula is the genus formula for curves with nodes and cusps. The second formula follows from Hurwitz 'formula by considering the projection from a point: The d^* intersection points of C^* with a general line $L_p \subset \check{\mathbb{P}}^2$ correspond to tangent lines of C passing through the corresponding point $p \in \mathbb{P}^2$. These are also ramification points of the morphism on the desingularization $\eta : \widetilde{C} \to C$ induced by the projection from p:

$$\pi_p \circ \eta : \widetilde{C} \to \mathbb{P}^1.$$

The cusps give further ramfication points, and no other points are ramified. Thus,

$$d^* = \deg R - \kappa = 2g - 2 + 2d = d(d-1) - 3\kappa - 2\delta.$$

To compute the number of flexes, we intersect C with the Hessian curve H and apply Proposition **??**. Recall from Exercise **??**, that the intersection multiplicity of the Hessian and C in a node is 6. In a cusp it is 8. Thus,

$$f = 3d(d-2) - 6\delta - 8\kappa.$$

The formulas for d and κ are dual to the preceeding ones. Finally, the formula for the bitangents follows from the previous formulas by a simple substitution. $2b = d^*(d^*-1) - d - 3f = (d(d-1) - 2\delta - 3\kappa)(d(d-1) - 2\delta - 3\kappa - 1) - d - 9d(d-2) + 18\delta + 24\kappa = d(d-2)(d^2-9) + \ldots$. □

8.5 Weierstrass Points and Automorphism

8.6 Riemann's Count

The Riemann-Roch Theorem allows to get a rough overview of all possible morphisms $C \to \mathbb{P}^n$ of a curve to some projective space. By Theorem **??**

such morphism corresponds to a base point free linear system $V \subset L(D)$ for some divisor class $D \in \mathrm{Div}(C)$, and linear equivalent divisors $D \equiv D'$ give isomorphic function spaces $L(D) \cong L(D')$ and equivalent embeddings. The main missing piece in our description is thus the quotient group

$$\mathrm{Pic}(C) = Div(C)/\equiv$$

called **Picard group** or **divisor class group** of C. The degree of a divisor induces a group homomorphism

$$\deg : \mathrm{Pic}(C) \to \mathbb{Z}.$$

Let $\mathrm{Pic}^d(C)$ denote the preimage of $d \in \mathbb{Z}$, i.e. the set of divisor classes of degree d. Not so obvious is, that each $\mathrm{Pic}^d(C)$ carries the structure of a projective variety defined over the field of C. We do not prove this fact in this book. However granted this, Riemann-Roch allows to compute its dimension.

Let $C^{(d)} = C \times \ldots C/S_d$ be the d-th symmetric product of C. Since C is 1-dimensional, the variety $C^{(d)}$ is smooth. To prove this fact, one uses the fundamental theorem on symmetric functions. For example, the symmetric product $(\mathbb{P}^1)^{(d)}$ is isomorphic to $\mathbb{P}^d$ via the map, which assigns to a collection of d points their defining equation up to a factor. In general we interprete $C^{(d)}$ as the variety of effective divisors $p_1 + \ldots + p_d$ of degree d on C. Consider the **Abel-Jacobi** map

$$u_d : C^{(d)} \to \mathrm{Pic}^d(C),$$

which associates to an effective divisor D its divisor class $[D] \in \mathrm{Pic}^d(C)$. The fiber of u_d over the point represented by a divisor D is the complete linear system $|D|$, which is a projective spaces of dimension $\ell(D) - 1$ or empty.

Thus, if C has genus g and $d \geq g$, then the map u_d is onto by Riemann's inequality, and for $d \geq 2g - 1$ all fibers are projective spaces of the same dimension $d - g$ by the Riemann-Roch Theorem 8.3.2. In particular, $\mathrm{Pic}^d(C)$ is irreducible, and the theorem on the fiber dimension **??** gives

$$\dim \mathrm{Pic}^d(C) = \dim C^{(d)} - \dim |D| = d - (d - g) = g,$$

for all $d \geq 2g - 1$. If $E \in \mathrm{Div}(C)$ is any divisor of degree $d - e$ then

$$D \mapsto D + E$$

induces an isomorphism

$$\mathrm{Pic}^e(C) \cong \mathrm{Pic}^d(C),$$

which is defined over the field of definition of E. We conclude

Proposition 8.6.1. *Let C be a smooth projective curve of genus g. Then $\dim \mathrm{Pic}^d(C) = g$ holds for every component of the Picard variety of C.*

Let M_g denote the set of isomorphism classes of smooth projective curves of genus g. It was used by Riemann intuitively, that M_g carries the structure of a variety. A rigorous construction of M_g and even better a geometric interpretation of a suitable projective closure was given by Mumford **?** in his fields medail winning work on geometric invariant theory.

The Riemann-Roch Theorem and Hurwitz's formula allow to compute the dimension of M_g at least if $\mathbb{k} = \mathbb{C}$.

By Riemann's inequality a general effective divisor $D \in C^{(d)}$ has $\ell(D) \geq 2$, if $d \geq g+1$. From any 2-dimensional subspace $W \subset L(D)$ we obtain a morphism

$$\varphi_W : C \to \mathbb{P}^1,\ p \mapsto [f_0(p) : f_1(p)],$$

once we choose a basis f_0, f_1 of W. Assume $d \geq 2g+1$. Then $|D|$ is base point free by Cororllary **??** and $\mathbb{P}(W) \subset |D|$ will be base point free as well for a general choice of $W \in \mathbb{G}(2, L(D))$ in the Grassmannian of 2-dimensional subspaces of $L(D)$. Thus, $\deg \varphi_W = d$ for general choice of W and the ramification divisor R has degree

$$\deg R = 2g - 2 + 2d$$

by Hurwitz's formula. Moreover, the map $\varphi_W : C \to \mathbb{P}^1$ for general choices has only simple branch points. The last fact is confirmed also by the dimension count below.

Let us consider

$$A_{d,g} = \{(C, [D], W) \mid C \in M_g, [D] \in \operatorname{Pic}^d(C), W \in \mathbb{G}(2, L(D))\}$$

for $d \geq 2g+1$ and

$$\widetilde{A}_{d,g} = \{(C, [D], W, f_0, f_1) \mid (C, [D], W) \in A_{d,g} \text{ and } f_0, f_1 \in W \text{ a basis}\}.$$

One can equip $A_{d,g}$ and $\widetilde{A}_{d,g}$ with the structure of a variety in a natural way. All fibers of the projection $\widetilde{A}_{d,g} \to A_{d,g}$ are isomorphic to $\mathrm{GL}(2, \mathbb{C})$, thus 4-dimensional. Let

$$M(\mathbb{P}^1, d)_g = \{(C, \varphi) \mid C \in M_g \text{ and } \varphi : C \to \mathbb{P}^1 \text{ of degree } \deg \varphi = d\}$$

denote the variety of morphisms of degree d from curves of genus g to $\mathbb{P}^1$. By Theorem 6.5.1 a general linear subspace of codimension 2 in $\mathbb{P}^{d-g}$ does not intersect the image of C under φ_D. Hence there is a dominant rational map

$$\widetilde{A}_{d,g} \dashrightarrow M(\mathbb{P}^1, d)_g$$

defined by

$$(C, [D], W, f_0, f_1) \mapsto (C, p \mapsto [f_0(p) : f_1(p)]).$$

The fibers of this morphism are all isomorphic to $\mathbb{C}^*$, because (f_0, f_1) and $(\lambda f_0, \lambda f_1)$ define the same map. We conclude

$$\dim \widetilde{A}_{d,g} = \dim M(\mathbb{P}^1, d)_g + 1.$$

To compute the dimension of $M(\mathbb{P}^1, d)_g$, we consider the map

$$M(\mathbb{P}^1, d)_g \to (\mathbb{P}^1)^{(2g-2+2d)}, (C, \varphi) \mapsto B = \varphi_* R,$$

which associates to a morphism φ the branch divisor B.

Proposition 8.6.2 (on branched coverings). *Suppose $\Bbbk = \mathbb{C}$. Given an effective divisor $B \in (\mathbb{P}^1)^{(2g-2+2d)}$, there are only finitely many isomorpohism classes of smooth projective curves C of genus g and maps $\varphi : C \to \mathbb{P}^1$, which have B as branch divisor. If B consists of $2g-2+2d$ distinct points, then the set of $\{(C, \varphi) \mid \varphi_* R = B\}$ is non-empty.*

The proof is based on topological and analytic arguments, which require that the reader is familiar with the notion of Riemann surfaces, and the following fundamental result:

Theorem 8.6.3 (Riemann). *For each compact connected Riemann surface C^{an} there exists a smooth irreducible projective algebraic curve C defined over $\mathbb{C}$, which has C^{an} as underlying complex analytic manifold. C is determined by C^{an} up to isomorphisms.*

Remark 8.6.4. The hardest part in the proof of this theorem is to establish that there exists a non-constant meromorphic functions on a compact Riemann surface. This is clearly satisfied in our applications below. We outline a proof of the theorem under this additional assumption in Exercises **??**-**??** below.

For higher dimensional compact complex manifold, the analogous result is not true. There are complex manifolds which have only constant meromorphic functions. A complex manifold might even differ considerably in its topology from any possible Euclidean topology of a complex projective variety. We refer as a start for further reading to Barth-Hulek-Peters-Van de Ven for the case of 2-dimensional complex manifolds.

Proof of the Proposition on branched coverings 8.6.2. Assume that $\infty \in \mathbb{P}^1$ is not a point of B and choose an ordering $b_1, \ldots, b_{r+1}$ of the points in the support of B such that the line segments $L_i = \{sb_i + (1-s)b_{i+1} \mid s \in [0,1]\} \subset \mathbb{A}^1(\mathbb{C}) \subset \mathbb{P}^1(\mathbb{C})$ form a polygon $L = L_1 \cup L_2 \cup \ldots \cup L_{r-1}$ with no self-intersections.

The underlying real 2-manifold of C is obtained from d copies, called sheets, of $\mathbb{P}^1 \setminus L$ by suitable gluings across the L_i. We fix an enumeration of the sheets. For each branch point b_i path, lifting to $C(\mathbb{C})$ of a small loop around $b_i \in \mathbb{P}^1(\mathbb{C})$ induces a permutation of the d sheets. Let $\sigma_i \in S_d$ be the corresponding element. Then we can recover the underlying Riemann surface of C as follows. Glue the d sheets across L_i according to the permutation $\sigma_i \circ \ldots \circ \sigma_1$. Thus, the underlying Riemann surface C^{an} depends only an B

and the finite set of additional data, given by the permutations σ_i. The first statement follows now from Riemann's Theorem 8.6.3.

For the second, we note that for given B, L and permutation σ_i the glueing leads to a connected Riemann surface iff the following conditions are satisfied:

1. $\sigma_r \circ \ldots \circ \sigma_2 \circ \sigma_1 = id$,
2. $\sigma_1, \ldots, \sigma_r$ generate a transitive subgroup of S_d.

Indeed, the first condition simply says, that a loop around the whole polygon induces the trivial permutation, the second is needed to obtain a connected surface. If these conditions are satisfied then we can give the surface the structure of a complex manifold: Away from the ramification points we take as local analytic coordinates preimage of coordinates on $\mathbb{P}^1$. At a ramification point p_{ij} over $b_i \in B$, we take the function $(z-b_i)^{1/e_{ij}}$ as a local coordinate, where e_{ij} is the ramification index, i.e. the length of the corresponding orbit of σ_i. Note, that although $(z-b_i)^{1/e_{ij}}$ is a multi-valued function on $\mathbb{P}^1$, it will be single-valued in a small neighbarhood of p_{ij}.

We obtain a reduced ramification divisor B iff each σ_i is a transposition. Thus, in case of $2g-2+2d$ ($\geq 2d$) simple branch points, we can choose for example the transposition $\sigma_{2i-1} = \sigma_{2i} = (i, i+1)$ for $i = 1, \ldots, d$ and $\sigma_{2i-1} = \sigma_{2i} = \tau_i$ an arbitrary transpositions for $i = d+1 \ldots d+g-1$. Then condition 1.) is satisfied because $\tau^2 = id$ holds for any transposition τ, and 2.) holds because $(1,2), \ldots, (d-1,d)$ generate S_d. □

Exercise 8.6.5. With the notion as in the first part of the proof above, show that the Galois group of $\mathbb{C}(C) \supset \mathbb{C}(\mathbb{P}^1) = \mathbb{C}(x)$ is isomorpic to the subgroup generated by $\sigma_1, \ldots, \sigma_r$. (c.f. Theorem 6.7.1 for a related topic). □

We are now ready to count the number of moduli of curves of genus g.

Theorem 8.6.6 (Riemann). *For $g \geq 2$ the moduli space M_g has dimension*

$$\dim M_g = 3g - 3,$$

i.e. curves of genus g depend on $3g-3$ moduli.

Proof. The result is actually true without any assumption on the ground field. However, we prove this only for $\Bbbk = \mathbb{C}$, because we will apply the result on branched coverings. By Theorem **??** any curve of genus $g \geq 2$ has a finite automorphism group. Thus, by the Proposition 8.6.2 on branched coverings $\dim M(\mathbb{P}^1, d)_g = 2g-2+2d$, and hence $\dim A_{d,g} = 2g-2+2d-3$. The fibers of

$$A_{d,g} \to \{(C,[D]) \mid C \in M_g,\, [D] \in \mathrm{Pic}^d(C)\}$$

are Grassmanians $\mathbb{G}(2, L(D))$ of dimension $2(\ell(D)-2) = 2(d+1-g-2)$ by Riemann-Roch and Exercise 6.3.39. Finally, the fibers of

$$\{(C,[D]) \mid C \in M_g,\, [D] \in \mathrm{Pic}^d(C)\} \to M_g$$

are g-dimensional by Proposition 8.6.1. Thus,

$$\dim M_g = \dim A_{d,g} - 2(d+1-g-2) - g = 3g-3.$$

□

Remark 8.6.7. If the genus $g = 1$ then $\dim \mathrm{Aut}(C) = 1$ with the connected component of the identity given the action of $\mathrm{Pic}^0(C)$ on $\mathrm{Pic}^1(C) \cong C$. This leads to $\mathrm{Hom}(1, \mathbb{P}^1)_d = 2g-2+2d+\dim \mathrm{Aut}(C)$ and we obtain $\dim M_1 = 1$. Similarly, $\dim M_0 = 0$, because $\mathrm{Aut}(\mathbb{P}^1) = \mathrm{PGL}(2, \mathbb{C})$ is 3-dimensional, which is consistent with the fact, that $\mathbb{P}^1$ is the only curve of genus 0 over $\mathbb{C}$.

Remark 8.6.8. A famous result of Bély says that a complex projective curve over $\mathbb{C}$ can be defined over an algebraic number field iff it can be described as a branched cover of $\mathbb{P}^1$ with only three branch points. Clearly, the proposition on branched coverings 8.6.2 says, that such curves do not depend on continuous parameters, because we can choose $0, 1, \infty$ as fixed branch points for such coverings.

Our next dimension count concerns the dimension of the space

$$H_{d,g} = \{C \subset \mathbb{P}^3 \mid C \text{ is a smooth curve of genus } g \text{ and degree } d\}.$$

Again, this set carries naturally the structure of an algebraic set.

Corollary 8.6.9. *The space of smooth curves of genus g and degree d in $\mathbb{P}^3$ has dimension*

$$\dim H_{d,g} = 4d$$

for $d \geq \max(2g+1, g+3)$.

Proof. To obtain a space curve we choose

1. a curve $C \in M_g$,
2. a divisor class $[D] \in \mathrm{Pic}^d(C)$,
3. a 4-dimensional subspace $W \subset L(D)$, which defines a very ample linear system, and
4. a basis of W up to a common scalar factor.

Thus,

$$\begin{aligned}\dim H_{d,g} &= \dim M_g + \dim \mathrm{Pic}^d(C) + \dim \mathbb{G}(4, L(D)) + \dim \mathrm{PGL}(3, \Bbbk) \\ &= 3g-3+g+4(d+1-g-4)+15 = 4d.\end{aligned}$$

□

Remark 8.6.10. The Proposition 8.6.9 is true in a larger range of degrees. For example also the space of lines in $\mathbb{P}^3$, i.e. the Grassmannian $\mathbb{G}(2,4)$, has dimension $4 = 4 \cdot 1$. However, for $3 \leq d \leq 2g-4$ the situation can be quite complicated because we do not know the dimensions and the behavior of the spaces of special divisors

$$W^r_d(C) = \{[D] \in \mathrm{Pic}^d(C) \mid \dim |D| \geq r\}$$

in this range precisely. Their dimensions vary with the curve upper semi-continuously. The following is known:

Theorem 8.6.11 (Brill-Noether,ACGH).

$$\dim W^r_d(C) \geq g - (r+1)(g+r-d),$$

and equality holds for general curves.

For special curves, the dimension $\dim W^r_d(C)$ can be larger. We refer to ACGH and HM for further reading on the subject of moduli of curves and special linear series.

Remark-Definition 8.6.12. The following notation is widely used in this context of studying special divisors: A g^r_d on a curve C denotes an r-dimensional linear system of divisors of degree d, possibly with base points.

Thus morphisms of C to $\mathbb{P}^3$ amounts to study g^3_d, while rational functions are related to g^1_d. Any g^1_d gives a non-constant rational function of degree $\leq d$. The degree might be smaller, because the g^1_d might have base points.

Exercise 8.6.13. Compute the dimension of the space of smooth curves of degree 2 in $\mathbb{P}^3$. □

Corollary 8.6.14. *A general smooth projective curve of genus g has no non-constant rational functions of degree*

$$d < \frac{g+2}{2}.$$

Proof. A rational function of degree d has at most $2g-2+2d$ branch points. Hence, by Proposition 8.6.2 the family of curves with such functions form a family of dimension at most $2g-2+2d-3$. The result follows from Riemann's count of moduli, Theorem 8.6.6, because

$$2g-2-2d-3 < 3g-3 \Leftrightarrow d < \frac{g+2}{2}.$$

□

Remark 8.6.15. The result is sharp. One can show, that every curve of genus g has a rational function of degree $\leq d = \lfloor \frac{g+3}{2} \rfloor$. To prove this by a dimension counting argument we would need the following:

Theorem 8.6.16 (Deligne-Mumford). *M_g is an irreducible variety.*

This combined with the following special piece of Brill-Noether theory would prove the assertion in Remark 8.6.15.

Proposition 8.6.17. *There exists a smooth projective curve C of genus g such that $W^1_d(C)$ for $d = \lfloor \frac{g+3}{2} \rfloor$ has dimension $= 0$ respectively $= 1$, if g is even respectively odd.*

However, both results are way beyond the techniques developed so far. Again we refer to [HM] and [ACGH] for further reading.

Definition 8.6.18. Let C be a smooth curve of genus g. The **gonality** of C is defined to be the smallest degree d of a non-constant rational function on C, in other words, the smallest number, such that $\overline{\mathbb{k}}(C)$ is an algebraic extension of $\overline{\mathbb{k}}(t)$ of degree

$$[\overline{\mathbb{k}}(C) : \overline{\mathbb{k}}(t)] = d.$$

The smallest possible gonality of a curve of genus $g \geq 2$ is two. These curves got an extra name:

Definition 8.6.19. A curve of genus $g \geq 2$ is **hyperelliptic**, if there exist a morphism

$$\pi : C \to \mathbb{P}^1$$

(possibly defined only over the algebraic closure of the ground field $\overline{\mathbb{k}}$).

Example 8.6.20. A smooth curve of bi-degree $(2, g+1)$ on $\mathbb{P}^1 \times \mathbb{P}^1$ is an hyperelliptic curve of genus g by Exercise 6.4.29.

Hyperelliptic curves of genus g are related to the study of hyperelliptic integrals

$$\int \frac{dx}{\sqrt{p(x)}}$$

with $p(x) = \prod_i (x - a_i) \in \mathbb{C}[X]$ a polynomial of degree $2g+1$ or $2g+2$ without multiple roots. (If $\deg p = 2g+1$ then the Riemann surface associated to the analytic function $\sqrt{p(x)}$ has at $\infty \in \mathbb{P}^1$ another branch point.)

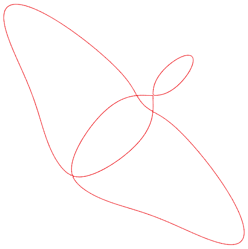
From an algebraic point of view, the gonality would seem to be the most basic invariant of a curve. Instead the genus plays this role. The genus stays constant in families of smooth projective curves, while the gonality is just upper semi-continues.

Example 8.6.21. Consider the family of curves

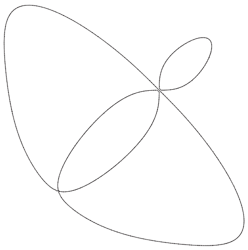
$$C_t = \mathrm{V}(f_t) \subset \mathbb{P}^2$$

defined by the affine equations $f_t = xy(x+y-2t) - 3(x^5+y^5) - 2(x^6+y^6) + t(12(x^4+y^4) + x^3y + xy^3 - 20x^2y^2 + 8(x^5+y^5) - 12(x^4y+xy^4) + 6(x^3y^2 +$

$x^2y^3) - 5(x^5y + xy^5) - 2(x^4y^2 + x^2y^4) + 14x^3y^3) + t^2(-12(x^3 + y^3) - 2(x^2y + xy^2) - 8(x^4 + y^4) + 24(x^3y + xy^3) - 44x^2y^2 + 10(x^4y + xy^4) + 20(x^3y^2 + x^2y^3) + 10(x^4y^2 + x^2y^4) + 24x^3y^3)$.
The curve C_t for a general value t has four ordinary double points in $p_0 = (0,0), p_1 = (2t, 0), p_2 = (0, 2t)$ and $p_3 = (-1,-1)$, and no further singularities.



$t = -0.1$



$t = 0$

However, the curve C_0 is irreducible with an ordinary tripel point in $p = (0,0)$ and a double point in $p_3 = (-1,-1)$. Hence, this is a family of curves of genus $g = 6$. Projection from the triple points yields a g^1_3 on C_0, hence, the curve C_0 is trigonal. On the other hand, the curve C_t for general values $t \neq 0$ is only 4-gonal. We will show this for the value $t = -\frac{1}{10}$ in Exercise 8.7.25 below. The projection from each of the 4 double points yields a g^1_4 on C_t. A fifth g^1_4 is obtained from the pencil of quadrics through the four double points. One can show that C_t has precisely five g^1_4.

8.7 Canonical Curves

Out of the many linear series on a curve only the canonical series is canonically given.

Theorem 8.7.1. *Let C be a smooth projective curve of genus $g \geq 2$. Then the canonical linear system $|K|$ is base point free. It is very ample unless C is hyperelliptic.*

Proof. This is an immediate application of the Riemann-Roch Theorem. $\ell(K - p) = \ell(K) - 1$ since $L(p) = \Bbbk$, as $C \not\cong \mathbb{P}^1$ has no rational function of degree 1. This gives base point freeness. Very ampleness means, that $\ell(K - p - q) = \ell(K) - 2$ for any pair of points on C. If this is not satisfied then $\ell(p + q) = 2$, which means that there exists a rational function of degree 2. $\square$

So the canonical morphism φ_K is either an embedding

$$\varphi_K : C \hookrightarrow \mathbb{P}^{g-1}$$

as a curve of degree $2g - 2$, or C is hyperelliptic and the map factors

$$C \xrightarrow{2:1} \mathbb{P}^1 \hookrightarrow \mathbb{P}^{g-1}$$

over the hyperelliptic map followed by the $(g-1)$-uple embedding of $\mathbb{P}^1$. In particular, we see that the hyperelliptic map is uniquely determined.

Example 8.7.2. A curve of genus $g = 2$ is always hyperelliptic, hence, a double cover of $\mathbb{P}^1$ branched at 6 points. Curves of genus 3 are either isomorphic to a plane quartic, or they are double covers of a smooth plane conic.

Note, that above we were bit sloppy. For arbitrary ground fields, the image of the canonical map is just a curve of genus 0, which is isomorphic to $\mathbb{P}^1$ over the ground field $\Bbbk$, iff it contains a $\Bbbk$-rational point.

Curves of genus 4 are either hyperelliptic or the intersection of a quadric and a cubic in $\mathbb{P}^3$, as we shall see below.

Exercise 8.7.3. Let C be an absolutely irreducible curve of genus 0. Prove that C is isomorphic to a plane conic over its field of definition.

Hint: Study the map given by the anticanonical system $|-K|$. □

Remark-Definition 8.7.4. A smooth **canonical curve** is an absolutely irreducible, smooth and non-degenerate curve of genus g and degree $2g-2$ in $\mathbb{P}^{g-1}$. On such a curve the hyperplanes cut out the complete canonical system.

Proof. To see $H \equiv K$, we note that since $\deg(H-K) = 0$, Riemann-Roch gives $\ell(H) = g$ iff $\ell(K-H) = 1$ iff $K \equiv H$. The $\ell(H) = g$ is satisfied, iff $C \subset \mathbb{P}^{g-1}$ is non-degenerate. □

Recall that an effective divisor D is special, if $L(K-D) \neq 0$. Since the hyperplanes cut out the complete linear series $|K|$ on a canonical curve, we can compute $|D|$ for an effective special divisor D as follows: Take a hyperplane $H \subset \mathbb{P}^{g-1}$ which passes through $D \subset C$ Consider the residual divisor $E = C.H - D$. Then the hyperplanes through E cut out $|D| + E = \{D' + E \mid D' \in |D|\}$.

Translated into a statement about secants to the canonical curves the Riemann-Roch formula takes an amusing form.

For D an effective divisor on a curve C, we denote by

$$\overline{D} = \bigcap_{\{H \mid \varphi_K^* H - D \geq 0\}} H \subset \mathbb{P}^{g-1}$$

the **linear span** of D in the canonical space.

Theorem 8.7.5 (Geometric version of Riemann-Roch). *Let D be an effective divisor on a curve of genus $g \geq 2$. Let $\overline{D} \subset \mathbb{P}^{g-1}$ be the linear span of D. Then*

$$\dim |D| = \deg D - 1 - \dim \overline{D}.$$

Proof. By the Riemann-Roch formula

$$\dim |D| + 1 = \ell(D) = \deg D + 1 - g + \ell(K - D).$$

The result follows by interpreting $\ell(K-D)$ as the codimension of $\overline{D}$ in $\mathbb{P}^{g-1}$. □

Example 8.7.6. If a canonical curve $C \subset \mathbb{P}^{g-1}$ has one trisecant line then C has infinitely many trisecants lines spanned by a pencil $|D|$ of divisor degree 3. In general a complete g^r_d with $d \leq g-1$ corresponds to an r-dimensional family of d-secant $(d-r-1)$-planes.

Our next goal is to prove, that canonical curves are projectively Cohen-Macaulay.

Lemma 8.7.7 (Base point free pencil trick). *Let C be a smooth projective curve and $P \subset |D|$ a base point free pencil spanned by two functions $f_1, f_2 \in L(D)$. Let H be a further divisor. The kernel of the map*

$$\mu : L(H) \oplus L(H) \to L(H+D),\ (g_1, g_2) \mapsto g_1 f_1 + g_2 f_2$$

is isomorphic to $L(H-D)$.

Proof. For $h \in L(H-D)$, we have $(f_2 h, -f_1 h) \in \ker \mu$ which proves one direction.

For the other inclusion, consider $D_i = (f_i) + D \in |D|$. By assumption, $\operatorname{supp}(D_1) \cap \operatorname{supp}(D_2) = \emptyset$. Moreover, choosing a different basis for P if necessary, we may assume in addition that the support of D_1 and D_2 is disjoint from the support of D. Now, if $(a_1, a_2) \in \ker \mu$, then $a_1 f_1 = -a_2 f_2$, and the assumption about the support of (f_1) and (f_2) implies that all points of D_2 are zeroes of a_1 with the appropriate multiplicity. Hence $\frac{a_1}{f_2} = -\frac{a_2}{f_1} \in L(H-D)$ as desired. □

Corollary 8.7.8. *Let $C \subset \mathbb{P}^r$ be a smooth irreducible curve. C is projectively normal iff C is arithmetically Cohen-Macaulay.*

Proof. Suppose C is projectively normal, i.e $(S/I_C)_m \cong L(mH)$ for every m. Consider a base point free pencil $P = \mathbb{P}(U) \subset L(H)$. The sequence

$$0 \to L((m-1)H) \to U \otimes L(mH) \to L((m+1)H)$$

is exact for all m, which implies, that for $\langle u_0, u_1 \rangle \in U \subset S_1$ the coordinate ring $S_C = S/\mathrm{I}_C$ is a free $\Bbbk[u_0, u_1]$-module. The converse implication follows from Remark **??**. □

Theorem 8.7.9 (Max Noether). *A canonical curve $C \subset \mathbb{P}^{g-1}$ is projectively normal.*

Proof. We have to prove that the map

$$\Bbbk[x_0, \ldots, x_{g-1}] \to \oplus_m L(mK)$$

is surjective. Equivalently, we have to prove that the multiplication maps

$$L(K) \otimes L((m-1)K) \to L(mK)$$

are surjective for all $m \geq 2$. The most difficult case is $m = 2$. Consider g general points $p_0, \ldots, p_{g-1}$ on $C \subset \mathbb{P}^{g-1}$. Then, these points will span $\mathbb{P}^{g-1}$. The sum of any $g-2$ of these points, say $E = p_1 + \ldots + p_{g-2}$, is a divisor, such that its span $\overline{E}$ intersects C only in the given $g-2$ points transversally. In other words, the linear system $L(K-E)$ is base point free.

(For example, we could take a $p_1, \ldots, p_{g-2}$ as $g-2$ points of a general hyperplane section of C by 6.7.9.)

Choose a basis $\omega_0, \ldots, \omega_{g-1}$ of $L(K)$ dual to these points, that is ω_i has a zero at p_j for $j \neq i$. Since $\deg 2K - E = 4g-4-(g-2) > 2g-1$, the subspace $L(2K-E) \subset L(2K)$ has codimension $\deg E = g-2$. Actually, $\omega_1^2, \ldots, \omega_{g-2}^2$ represent a basis of the quotient space $L(2K)/L(2K-E)$. Thus to prove that the multiplication map

$$L(K) \otimes L(K) \to L(2K)$$

is onto, it is enough to prove that

$$L(K-E) \otimes L(K) \to L(2K-E)$$

is surjective. By the base point free pencil trick, Lemma 8.7.7, the kernel of this map is isomorphic to $L(K-(K-E)) = L(E)$, which is one-dimensional. Hence the image has dimension $2g-1$. Since $\ell(2K-E) = 4g-4-(g-2)+1-g = 2g-1$ by Riemann-Roch, surjectivity follows. For arbitray $m \geq 3$, we argue similarly: $\omega_1^m, \ldots, \omega_{g-2}^m$ represent a basis of $L(mK)/L(mK-E)$ and $L((m-1)K) \otimes L(K-E) \to L(mK-E)$ is surjective, since the kernel $L((m-2)K+E)$ has dimension

$$\ell((m-2)K+E) = (m-2)(2g-2)-1,$$

which equals

$$2\ell((m-1)K) - \ell(mK-E) = (2m-1)(2g-2) - (m(2g-2)-2g+3).$$

□

Corollary 8.7.10. *The Hilbert function of a canonical curve of genus g takes values*

$$(1, g, 3g-3, 5g-5, \ldots, m(2g-2)+1-g, \ldots).$$

We go a bit further in the analysis of Max Noether and compute a Gröbner basis of the ideal I_C in coordinates $x_0, \ldots, x_{g-1}$ of $\mathbb{P}^{g-1}$, such that $x_i \mapsto \omega_i$. We use the reversed lexicographic order in the following order of our variables

$$x_1 > x_2 > \ldots > x_{g-1} > x_0.$$

Since $\omega_i \omega_j \in L(2K - E)$ for $1 \leq i \neq j \leq g-2$, they are contained in the image of $L(K-E) \otimes L(K)$.

This means, that I_C contains elements of type

$$F_{ij} := x_i x_j - \sum_{i=1}^{g-2} a_{ij}^r x_r - b_{ij}, \text{ for } 1 \leq i < j \leq g-2$$

with $a_{ij}^r, b_{ij} \in \Bbbk[x_{g-1}, x_0]$ of degree 1 and 2, respectively. Note $\mathbf{L}(F_{ij}) = x_i x_j$. Since $\dim(I_C)_2 = \binom{g+1}{2} - 3g + 3 = \binom{g-2}{2}$, these quadrics span $(I_C)_2$. However, these are not enough elements for a Gröbner basis for I_C, because $\dim\langle \mathbf{L}(F_{ij} | 1 \leq i < j \leq g-2 \rangle_3 = 6g-8 = 5g-5+g-3$. If we run Buchberger's test on the F_{ij}, we obtain a division expression

$$F_{ij,k} = x_k F_{ij} - x_j F_{ik} + \sum_{r \neq k} a_{ij}^r F_{rk} - \sum_{r \neq j} a_{ik}^r F_{rj} + a_{ij}^k x_k^2 - a_{ik}^j x_j^2 + \text{lower order terms.}$$

Let us take a closer look at the a_{ij}^k: All terms in F_{ij} vanish at p_k to the second order. Hence,

$$a_{ij}^k = \rho_{ijk} \alpha_k,$$

where $\alpha_k \in \langle x_{g-2}, x_0 \rangle$ defines the unique hyperplane in the pencil, which vanishes at p_k to order ≥ 2 and $\rho_{ijk} \in \Bbbk$. $\alpha_k x_k^2$ maps to an element of $L(3K - 2E)$.

Lemma 8.7.11. *The image of $L(K-E) \times L(2K-E)$ in $L(3K-2E)$ spans a space of codimension 1. Moreover, for a general choice of of $p_0, \ldots, p_{g-1}$ each of the elements $\alpha_k x_k^2$ represents a basis of this one dimensional cokernel*

Proof. Since $\langle x_{g-1}, x_0 \rangle = L(K-E)$ is base point free pencil, the trick 8.7.7 identifies the kernel of $L(K-E) \otimes L(2K-E) \to L(3K-2E)$ with $L(K)$. Hence the image has dimension

$$2\ell(2K-E) - \ell(K) = 2(3g-3-g+2) - g = 5g-5-2g+4-1.$$

This is codimension 1 in $L(3K-2E)$.

For the second statement, we note that if $\alpha_k x_k^2$ does not span the cokernel, then we find an equation of type $\alpha_k x_k^2 - \sum_{r=1}^{g-2} c_k^r x_r - d_k$ in the ideal I_C. Since there are precisely $g-3$ further Gröbner basis elements of degree three, at least one of the $g-2$ elements $\alpha_k x_k^2$ spans. But then all of them span for general choices: The spanning of $\alpha_k x_k^2$ is satisfied for the choice of $(p_0, \ldots, p_{g-1})$ in a Zariski open subset $U_k \subset C^g = C \times C \ldots \times C$. If one of these sets is non-empty, then by symmetry of the problem, all U_j are non-empty and Zariski dense. So we can choose $(p_0, \ldots, p_{g-1}) \in \bigcap_{j=1}^{g-2} U_j$, which is non-empty, because C^g is irreducible. □

Note also, that for general choices $\alpha_k \notin \langle x_0 \rangle$, so that $\mathbf{L}(\alpha_k)$ is a multiple of x_{g-1}. Adjusting α_k with a scalar multiple, we can achieve that the images of the $\alpha_k x_k^2$ in the one-dimensional space $L(3K-2E)/\operatorname{im}(L(K-E) \times L(2K-E))$ coincide. With this adjustment we get elements

$$G_{kl} = \alpha_k x_k^2 - \alpha_l x_l^2 - \sum_{r=1}^{g-2} c_{kl}^r x_r - d_{kl}$$

in the ideal, where $c_{kl}^r, d_{kl} \in \Bbbk[x_{g-1}, x_0]$. Note, that $G_{kl} + G_{ln} = G_{nl}$. This gives us $g-3$ cubics for the Gröbner basis, which have as initial forms scalar multiples of $x_{g-1}x_k^2$ for $k = 1, \ldots, g-3$. The last Gröbner basis element is a quartic,

$$H_{g-2} = \alpha_{g-2} x_{g-2}^3 + \text{ lower order terms}$$

obtained from any of the S-polynomials of the pairs $(G_{i,g-2}, F_{i,g-2})$ in Buchberger's test.

Proposition 8.7.12. *The polynomials $F_{ij}, G_{k,g-2}, H_{g-2}$ form a minimal Gröbner basis for I_C.*

Proof. The monomial ideal, which is generated by $x_i x_j, 1 \leq i < j \leq g-2$, $x_{g-1}x_k^2, k = 1, \ldots, g-3$, and $x_{g-1}x_{g-2}^3$, has the same Hilbert function as I_C. □

Theorem 8.7.13 (Petri). *Let $C \subset \mathbb{P}^{g-1}$ be a smooth irreducible canonical curve of genus $g \geq 4$. The homogeneous ideal of C is generated by quadrics and cubics. It is generated by quadrics alone, unless C is trigonal, or C is isomorphic to a smooth plane quintic, which has genus $g = 6$.*

In the exceptional cases, the ideal I_C needs in addition $g-3$ cubic generators, and the quadrics generate the ideal of a surface X of degree $g-2$, which is either the surface

$$X = \bigcap_{D \in g_3^1} \overline{D},$$

spanned by the tricecant lines, or the Veronese surface

$$\mathbb{P}^2 \hookrightarrow \mathbb{P}^5.$$

Proof. It is easy to see that the exceptional curves cannot have the ideal generated by quadrics: A curve with a g_3^1 has infinitely many trisecant lines, which will lie in the intersecion of all quadrics containing C by Bezout's Theorem 6.4.33. Similarly, a curve with a g_5^2 has infinitely many 5-secant 2-planes. Since 5 points in a plane lie on a unique conic, the intersection of the quadrics containing C will contain these conics.

In any case, by the Proposition on the Gröbner basis 8.7.12, the ideal is generated by quadrics and cubics, since the quartic in the Gröbner basis is a linear combination of these. Buchberger's test gives us syzygies

$$F_{ij,k} = x_k F_{ij} - x_j F_{ik} + \sum_{r \neq k} a_{ij}^r F_{rk} - \sum_{r \neq j} a_{ik}^r F_{rj} + \rho_{ijk} G_{kj}.$$

In particular, we see that for $a_{ij}^k = \alpha_k \rho_{ijk}$ and $a_{ik}^j = \alpha_j \rho_{ikj}$ the factors

$$\rho_{ijk} = \rho_{ikj}$$

coincide. Otherwise, we would get a contradiction to Lemma 8.7.11 above. So **Petri's coefficients** ρ_{ijk} are symmetric in all three indices. We argue now similarly as in the proof of the Lemma. By the symmetry of the situation and the observation, that the vanishing of a coefficient ρ_{ijk} defines an algebraic subset for choices of points $(p_0, \ldots, p_{g-1}) \in C^g$, we can assume that either all ρ_{ijk} are non-zero or all are zero, because C^g irreducible. If all are non-zero then quadrics generate the cubics by the syzygies above, and the ideal is generated by quadrics alone.

On the other hand, if all $\rho_{ijk} = 0$ then the additional $g - 3$ cubics of the Gröbner basis are needed for generation. The same syzygies as above tell us this time, that the F_{ij} form a Gröbner basis. Hence, they define a two-dimensional algebraic set X. Since C is irreducible and

$$\mathrm{V}(\{F_{ij} \mid 1 \leq i < j \leq g-2\}) \cap \mathrm{V}(x_{g-1}, x_0) = \{p_1, \ldots, p_{g-2}\} \subset C \subset \mathbb{P}^{g-1},$$

this algebraic set is an irreducible surface X of degree $g - 2$. By Bertini's classification of surfaces of minimal degree, see Exercise **??**, this is either a rational surface ruled by lines or the Veronese surface $\mathbb{P}^2 \hookrightarrow \mathbb{P}^5$. In the first case, the geometric version of Rieman-Roch 8.7.5 gives that the lines are trisecant to C: A rational family of 2-secant is exclude because otherwise the curve would be hyperelliptic, 4-secants are excluded, because the ideal of C is generated by quadrics and cubics. In the second case, we see that the preimage of C in $\mathbb{P}^2$ is a smooth quintic, canonically embedded by the adjoint system $|(5-3)L|$ on $\mathbb{P}^2$. □

Example 8.7.14. A non-hyperelliptic curve of genus $g = 4$ is the complete intersection of a quadric and a cubic in $\mathbb{P}^3$. The curve is trigonal and has either one or two g_3^1, depending on whether the quadric is a cone or isomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$.

A non-hyperelliptic curve of genus $g = 5$ is either a complete intersection of three quadrics, or it is trigonal, and $C \subset \mathbb{P}^4$ lies on a cubic scroll $X \subset \mathbb{P}^4$.

Petri's Theorem got much attention due to the conjectural generalization to higher syzygies of canonical curves due to Mark Green 1984. To explain Green' conjecture, we introduce a short notation for the numerical type of a free resolution. Let

$$F_0 \leftarrow F_1 \leftarrow \ldots \leftarrow F_c \leftarrow 0$$

be a free complex of graded S-modules with $F_i = \bigoplus_j S(-j)^{\beta_{ij}}$. We summarize the numerical information in a table indexed by i and $k = i - j$:

	0	1	...	i	...	c
0	β_{00}	β_{11}		β_{ii}	...	
1	β_{10}	β_{12}		β_{ii+1}		
⋮	⋮	⋮		⋮		⋮
k	β_{0k}	β_{1k+1}		β_{ii+k}		
⋮				⋮		
m		...		β_{ii+m}		β_{cc+m}

which we call the **Betti table** of the complex F_*.

A few invariants of the numerical data have special names. The integer

$$m + 1 = \min\{k \mid \beta_{ij} = 0 \forall i, j \text{ with } j \geq i + k\}$$

is called the **Castelnouvo-Mumford regularity** of the complex. For the minimal free resolution F of a ring S/I as an S-module, this is the number of rows occuring, and we say that S/I is $m+1$-regular. The largest homological degree c is called the **length of the complex**.

Example 8.7.15. The twisted cubic curve $\mathbb{P}^1 \hookrightarrow \mathbb{P}^3$ is defined by three quadric equations, among which there are two linear syzygies. Hence, the corresponding Betti table of the coordinate ring is

	0	1	2
0	1	-	-
1	-	3	2

The length of this resolution is $c = 2$ and S/I is 2-regular.

An easy encription rule for the Betti table runs as follows: Two numbers, standing in the same row next to each other, correspond to map between free modules of corresponding rank with linear entries. Entries in the corners of a square correspond to a map with quadratic entries.

The minimal complex of Example **??** 2.86 has Betti table

	0	1	2	3
0	1	-	-	-
1	-	5	5	-
2	-	-	-	1

that is, the complex has maps with quadratic entries in the end and in the beginning, and a map with linear entries in the middle.

Exercise 8.7.16. Let $\mathbb{P}^1 \hookrightarrow \mathbb{P}^d$ be the rational normal curve of degree d. Prove that its coordinate has the following Betti table

	0	1	2	…	i	…	$d-1$
0	1	-	-	…	-	…	-
1	-	$\binom{d}{2}$	$2\binom{d}{3}$	…	$i\binom{d}{i+1}$	…	$d-1$

□

Here is what is known about the Betti numbers of the minimal free resolution of canonical curves of genus $5 \leq g \leq 9$. To include the case of hyperelliptic curves, we consider the resolution of $R = \oplus_d L(dK)$ as an $S = \Bbbk[x_0, \dots, x_{g-1}]$-module. The tables below includes all possible Betti tables of smooth canonical curves defined over field of characteristic $\operatorname{char} \Bbbk = 0$, and the interpretation in terms of existence of special linear series.

$g = 5$

general					$\exists\, g^1_3$					$\exists\, g^1_2$				
	0	1	2	3		0	1	2	3		0	1	2	3
0	1	-	-	-	0	1	-	-	-	0	1	-	-	-
1	-	3	-	-	1	-	3	2	-	1	-	6	8	3
2	-	-	3	-	2	-	2	3	-	2	3	8	6	-
3	-	-	-	1	3	-	-	-	1	3	-	-	-	1

$g = 6$

general						$\exists\, g^1_3$ or g^2_5						$\exists\, g^1_2$					
	0	1	2	3	4		0	1	2	3	4		0	1	2	3	4
0	1	-	-	-	-	0	1	-	-	-	-	0	1	-	-	-	-
1	-	6	5	-	-	1	-	6	8	3	-	1	-	10	20	15	4
2	-	-	5	6	-	2	-	3	8	6	-	2	4	15	20	10	-
3	-	-	-	-	1	3	-	-	-	-	1	3	-	-	-	-	1

$g = 7$

general							$\exists\, g^1_4$							$\exists\, g^2_6$						
	0	1	2	3	4	5		0	1	2	3	4	5		0	1	2	3	4	5
0	1	-	-	-	-	-	0	1	-	-	-	-	-	0	1	-	-	-	-	-
1	-	10	16	-	-	-	1	-	10	16	3	-	-	1	-	10	16	9	-	-
2	-	-	-	16	10	-	2	-	-	3	16	10	-	2	-	-	9	16	10	-
3	-	-	-	-	-	1	3	-	-	-	-	-	1	3	-	-	-	-	-	1
$\exists\, g^1_3$							$\exists\, g^1_2$													
	0	1	2	3	4	5		0	1	2	3	4	5							
0	1	-	-	-	-	-	0	1	-	-	-	-	-							
1	-	10	20	15	4	-	1	-	15	40	45	24	5							
2	-	4	15	20	10	-	2	5	24	45	40	15	-							
3	-	-	-	-	-	1	3	-	-	-	-	-	1							

$g = 8$

general

	0	1	2	3	4	5	6
0	1	-	-	-	-	-	-
1	-	15	35	21	-	-	-
2	-	-	-	21	35	15	-
3	-	-	-	-	-	-	1

∃ g_4^1

	0	1	2	3	4	5	6
0	1	-	-	-	-	-	-
1	-	15	35	25	4	-	-
2	-	-	4	25	35	15	-
3	-	-	-	-	-	-	1

∃ g_6^2

	0	1	2	3	4	5	6
0	1	-	-	-	-	-	-
1	-	15	35	35	14	-	-
2	-	-	14	35	35	15	-
3	-	-	-	-	-	-	1

∃ g_3^1

	0	1	2	3	4	5	6
0	1	-	-	-	-	-	-
1	-	15	40	45	24	5	-
2	-	5	24	45	40	15	-
3	-	-	-	-	-	-	1

∃ g_2^1

	0	1	2	3	4	5	6
0	1	-	-	-	-	-	-
1	-	21	70	105	84	35	6
2	6	35	84	105	70	21	-
3	-	-	-	-	-	-	1

$g = 9$

general

	0	1	2	3	4	5	6	7
0	1	-	-	-	-	-	-	-
1	-	21	64	70	-	-	-	-
2	-	-	-	-	70	64	21	-
3	-	-	-	-	-	-	-	1

∃ g_5^1

	0	1	2	3	4	5	6	7
0	1	-	-	-	-	-	-	-
1	-	21	64	70	4	-	-	-
2	-	-	-	4	70	64	21	-
3	-	-	-	-	-	-	-	1

∃ two g_5^1

	0	1	2	3	4	5	6	7
0	1	-	-	-	-	-	-	-
1	-	21	64	70	8	-	-	-
2	-	-	-	8	70	64	21	-
3	-	-	-	-	-	-	-	1

∃ three g_5^1

	0	1	2	3	4	5	6	7
0	1	-	-	-	-	-	-	-
1	-	21	64	70	12	-	-	-
2	-	-	-	12	70	64	21	-
3	-	-	-	-	-	-	-	1

∃ g_7^2

	0	1	2	3	4	5	6	7
0	1	-	-	-	-	-	-	-
1	-	21	64	70	24	-	-	-
2	-	-	-	24	70	64	21	-
3	-	-	-	-	-	-	-	1

∃ g_4^1

	0	1	2	3	4	5	6	7
0	1	-	-	-	-	-	-	-
1	-	21	64	75	24	5	-	-
2	-	-	5	24	75	64	21	-
3	-	-	-	-	-	-	-	1

∃ $g_4^1 \times g_5^1$

	0	1	2	3	4	5	6	7
0	1	-	-	-	-	-	-	-
1	-	21	64	75	44	5	-	-
2	-	-	5	44	75	64	21	-
3	-	-	-	-	-	-	-	1

∃ g_6^2

	0	1	2	3	4	5	6	7
0	1	-	-	-	-	-	-	-
1	-	21	64	90	64	20	-	-
2	-	-	20	64	90	64	21	-
3	-	-	-	-	-	-	-	1

∃ g_3^1

	0	1	2	3	4	5	6	7
0	1	-	-	-	-	-	-	-
1	-	21	70	105	84	35	6	-
2	-	6	35	84	105	70	21	-
3	-	-	-	-	-	-	-	1

∃ g_2^1

0	1	2	3	4	5	6	7
1	-	-	-	-	-	-	-
-	28	112	210	224	140	48	7
7	48	140	224	210	112	28	-
-	-	-	-	-	-	-	1

Remark 8.7.17. The proof, that these are all possible cases and that they have the corresponding interpretation, can be found in Schreyer [1986] for $g \leq 8$ and Sagraloff [2005] for $g = 9$. The interpretation in case of genus 9 in terms of the existence of several g_5^1's has to be taken as a count with multiplicity. It is true, that a general curve in this strata defined by these Betti numbers have that many g_5^1's precisely.

The length $c = g - 3$ and regularity $m = 4$ of the free resolution of a canonical curve can be deduced from the Gröbner basis 8.7.12 and the syzygy algorithm 2.8.11. The fact, that the alternating sums $\sum_i (-1)^i \beta_{ij}$ for fixed j depends only the Hilbert function follows from the formula

$$h_C(t) = \sum_{i=0}^{c} (-1)^i \sum_j \beta_{ij} \binom{t+j+g-1}{g-1},$$

relating the Betti numbers to the Hilbert function, c.f. the proof of Theorem **??**.

The symmetry of the Betti table of a canonical curve

$$\beta_{i,j} = \beta_{g-2-i,g+1-j}$$

follows from the fact, that R_C is a so called **Gorenstein ring**. More general, the following is true

Proposition 8.7.18. *Let $C \subset \mathbb{P}^n$ be a smooth arithmetically Cohen-Macaulay curve, such that the canonical divisor class is a multiple of the hyperplane class $K \equiv kH$ for some integer k. Then S_C is a Gorenstein ring and the minimal free resultion F_* is self-dual:*

$$\operatorname{Hom}(F_i, S(-n-1)) \cong F_{n-2-i}(k)$$

and the matrices in the resolutions can be choosen to be transposed of each other possibly upto a sign.

For a proof and the definition of Gorenstein rings, we refer to Eisenbud [1995].

No one has yet formulated a precise statement of how much a configuration of extra special linear series on a canonical curve of genus g contribute to extra syzygies, and to answer the question, whether the Betti numbers depend only on the configuration of the special linear series.

However Mark Green [1984] formulated a precise conjecture concerning the range non-zero β_{ij}. We need the notion of the Clifford index of a curve.

Theorem 8.7.19 (Clifford). *Let D be a special divisor of degree $\deg D > 0$ on a smooth irreducible projective curve of genus g. Then*

$$d \geq 2r,$$

where $r = \dim |D|$, and equality holds, iff either $D \equiv K$ or C is hyperelliptic and $D \equiv kH$ for $|H|$ the hyperelliptic pencil and k an integer $1 \leq k \leq g-2$.

We need a Lemma.

Lemma 8.7.20. *Let D,E be effective divisors on a smooth irreducible projective curve C. Then*

$$\dim |D| + \dim |E| \leq \dim |D+E|.$$

Proof. Consider the multiplication map

$$L(D) \times L(E) \to L(D+E).$$

We may interprete this map as an $\ell(D) \times \ell(E)$ matrix φ with entries in the vector space $L(D+E)$. Since the multiplication of two non-zero rational functions on a irreducible curve is non-zero, the matrix φ has no zero entry for arbirtrary choices of bases for $L(D)$ and $L(E)$. Matrices of linear forms with this property are called 1-generic. They are studied for example in Eisenbud et al. . From a different point of view, we may interprete φ as a linear subspace $\mathbb{P}^n$ of the Segre space $\mathbb{P}^{\ell(D)\ell(E)-1}$, containing $\mathbb{P}^{\ell(D)-1} \times \mathbb{P}^{\ell(E)-1}$ of codimension $\leq \ell(D+E)$. The condition on 1-genericity says that this $\mathbb{P}^n$ does not intersect $\mathbb{P}^{\ell(D)-1} \times \mathbb{P}^{\ell(E)-1}$. So

$$\ell(D) - 1 + \ell(E) - 1 \leq \ell(D+E) - 1$$

by the dimension bound on intersections 6.5.1. This is the desired formula. □

Proof of Theorem 8.7.19. If $\ell(D) = 0$ then there is nothing to prove. So we may assume, that $\ell(D) \geq 1$ and $\ell(K-D) \geq 1$. By Lemma 8.7.20, we have

$$\ell(D) - 1 + \ell(K-D) - 1 \leq \ell(K) - 1 = g - 1.$$

Adding this inequality to the Riemann-Roch formula

$$\ell(D) - \ell(K-D) = \deg D + 1 - g,$$

we obtain the desired

$$2(\ell(D) - 1) \leq \deg D.$$

This proves the first statement.

It is clear, that in case $D \equiv K$ equality holds and that no other divisor class of degree $\geq 2g-2$ achieves equality. So suppose, equality holds for a divisor D of necessarily even degree d with $2 \leq d \leq 2g-4$. We have to prove, that C is hyperelliptic. If $d = 2$ then D itself is the hyperelliptic pencil. In general, we proceed by induction on d. Suppose $d \geq 4$ and hence $\dim |D| \geq 2$. Choose a divisor $E \in |K-D|$ and two points $p, q \in C$ with $p \in \operatorname{Supp} E$ and $q \notin \operatorname{Supp} E$. Since $\dim |D| \geq 2$, we can choose the divisor $D \in |D|$, such that $D - p - q$ is effective. Consider now $D' = min(D, E)$, the divisor whose coefficients are the minimum of those of D and E. We will show that we can apply to D' the induction hypothesis.

Since $Q \in \operatorname{Supp} D$ and $Q \notin \operatorname{Supp} E$, we have $\deg D' < \deg D$. On the other hand, $\deg D' > 0$, because $p \in \operatorname{Supp} D'$.

By definition of of D', we have an exact sequence

$$0 \to L(D') \to L(D) \oplus L(E) \to L(D+E-D'),$$

defined by the diagonal inclusion and subtraction. Hence,

$$\ell(D) + \ell(E) \leq \ell(D') + \ell(D+E-D').$$

Since $E \equiv K - D$, the left hand side has dimension $g+1 = \ell(K)+1$ by Riemann-Roch and the induction hypothesis $\ell(D) - 1 = \frac{1}{2} \deg D$. Since $D + E - D' \equiv K - D'$ holds as well, we obtain from Lemma 8.7.20

$$\ell(D') + \ell(K-D') \leq \ell(K) + 1 = g+1.$$

Thus, equality holds everywhere and $\deg D' = \frac{1}{2} \dim |D'|$ holds as well. With the induction hypothesis, we conclude that C is hyperelliptic.

Finally, we prove $D \equiv kH$ for H the hyperelliptic series and $k = \frac{d}{2}$. By the Lemma $\dim|D| + \dim|K-kH| \leq \dim|K-kH+D|$. Since the left hand side is $g-1$, we conclude $\ell(K-kH+D) \geq g$. Since $\deg D - kH = 0$, this is only possible if $D \equiv kH$. □

Definition 8.7.21. Let C of genus $g \geq 3$. The Clifford index of a divivsor D is

$$Cliff(D) = d - 2r,$$

where $d = \deg D$ and $r = \dim |D|$. The Clifford index of C is

$$Cliff(C) = \min\{Cliff(D) | \dim|D| \geq 1 \text{ and } \dim|K-D| \geq 1\}.$$

Divisor classes with $\dim|D| \geq 1$ and $\dim|K-D| \geq 1$ are said to contribute to the Clifford index.

Thus,

$$Cliff(C) = 0 \Leftrightarrow C \text{ is hyperelliptic.}$$

Furthermore

$$\begin{aligned}
Cliff(C) = 1 \Leftrightarrow\ & C \text{ is trigonal, or} \\
& C \text{ is isomorphic to a smooth plane quintic and } g = 6, \\
Cliff(C) = 2 \Leftrightarrow\ & C \text{ is 4-gonal,or} \\
& C \text{ is isomorphic to a smooth plane sextic and } g = 10, \\
Cliff(C) = 3 \Leftrightarrow\ & C \text{ is 5-gonal, or} \\
& C \text{ is isomorphic to a smooth plane septic and } g = 15, \text{ or} \\
& C \text{ is isomorphic to a smooth complete intersection of} \\
& \text{two cubics in } \mathbb{P}^3 \text{ and } g = 10, \\
& \vdots
\end{aligned}$$

For more information on the Clifford index we refer to [ELMS]. From [CM] it is known, that

$$gon(C) - 3 \leq Cliff(C) \leq gon(C) - 2.$$

Thus, we may regard the Clifford index as a slight modification of the gonality of a curve.

Conjecture 8.7.22 (Green, 1984). Let C be a smooth projective curve defined over $\mathbb{C}$ of genus g. Let β_{ij} be the Betti numbers of $R = \oplus_n L(nK)$ as $S = \mathbb{C}[x_0, \ldots, x_{g-1}]$-module, and let $p \geq 0$ be an integer. Then

$$\beta_{pp+2} \neq 0 \Leftrightarrow Cliff(C) \leq p.$$

Thus, the conjecture says, that a curve of Clifford index p has a non-zero Betti numbers in the following range:

	0	1	...	p-1	p	...	...	g-2-p	g-1-p	...	g-3	g-2
0	1	-	-	-	-	-	-	-	-	-	-	-
1	-	β_{12}	...	β_{p-1p}	β_{pp+1}	...	...	β_{pp+2}	-	-	-	-
2	-	-	-	-	β_{pp+2}	...	...	β_{pp+1}	β_{p-1p}	...	β_{12}	-
3	-	-	-	-	-	-	-	-	-	-	-	1

Remark 8.7.23. We may interprete Noether's Theorem as the case $p = 0$. Petri's Theorem is the case $p = 1$. The case $p = 2$ was proved in Voisin [198x] for $g \geq 11$ and in Schreyer [1991] in general. The direction $\Leftarrow$, that is from geometry to syzygies, which is the easier one, was established by Green and Lazarsfeld [1984].

For the other direction, Voisin [2002], [2004] proved, that a general curve of genus g satisfies Green's conjecture using the theory of $K3$-surfaces and more. Her result combined with the work of Teixidor [199x] implies, that the conjecture holds for a general curve of any given gonality. (The set of curves of gonality $p+2$ and genus g can be equipped with the structure of a variety, baptised Hurwitz scheme. In particular, they form one irreducible family.)

Note that the conjecture for curves of odd genus $g = 2k+1$ says, that a for a general curve the middle matrix in the free resolution has only quadratic entries, since such a curve has no g^1_{k+1}, (a fact, which is known from Brill-Noether theory). It follows from Voisin result, that for odd genus, curves with extra syzygies form a pure codimension 1 algebraic subset in M_g. Hirschowitz and Ramanan [199x] proved, that this subset coincide with the closure of curves with g^1_k. This is a very strong evidence for the conjecture.

On the other hand, the conjecture does not hold for curve defined over a field of char($\Bbbk$) > 0, the first case occurs for curves of genus $g = 7$ over a field $\Bbbk$ of characteristic char($\Bbbk$) $= 2$, in which case $\beta_{24} = \beta_{34} = 1$ holds for a general curve. Other examples are $g = 9$, char($\Bbbk$) $= 3$, C general, where $\beta_{35} = \beta_{45} = 6$ holds, and $g = 10$, char($\Bbbk$) $= 3$, C general, where $\beta_{35} = \beta_{56} = 1$. The last example was only established experimentally, a rigorous proof is still waiting. For this and further examples see Schreyer [2003].

Remark 8.7.24. Let $C \subset \mathbb{P}^2$ be a absolutely irreducible curve of geometric genus g and degree d with only ordinary singularities $p_1, \ldots, p_s$ of multiplicity $r_1, \ldots, r_s$. Then the canonical curve is contained in the image of the rational map $\mathbb{P}^2 \dashrightarrow \mathbb{P}^{g-1}$ defined by the adjoint system $L(d-3; (r_1-1)p_1, \ldots, (r_s-1)p_s)$.

Exercise 8.7.25. Compute the image of the curves in Remark 8.6.21 under the canonical map for $t = 0$ and $t = -1/10$ using a Computer algebra system. Conclude, that the curve for $t = 0$ is trigonal, and that the curve for $t = -1/10$ is not. □

Exercise 8.7.26. Compute the canonical image and the syzygies of the following curves of genus $g = 9$ in examples constructed by yourself.

1. An irreducible octic plane curve with 12 ordinary double points.
2. An irreducible octic plane curve with 1 ordinary triple point and 9 ordinary double points.
3. An irreducible octic plane curve with 2 ordinary triple points and 6 ordinary double points.
4. An irreducible octic plane curve with 3 ordinary triple points and 3 ordinary double points.
5. An irreducible septic plane curve with 6 ordinary double points.

Can you find a curve of genus $g = 9$ with three g^1_5? □

Exercise 8.7.27. Compute the dimension of the family of curves in Exercise 8.7.26 and compare this with the dimension of

1. the space of all curves of genus 9,
2. the space of 5-gonal curves of genus 9,
3. the space of curves of genus 9 with two g^1_5,
4. the space of curves of genus 9 with a g^2_7.

□

8.8 The Hasse-Weil Formulas

In this section $\mathbb{F}_q$ denotes a finite field with q elements. Given a variety X defined over $\mathbb{F}_q$, we ask how many $\mathbb{F}_q$-rational points are contained in X.

Example 8.8.1. For the projective space we have

$$|\mathbb{P}^n(\mathbb{F}_q)| = q^n + q^{n-1} + \ldots + q + 1 = \frac{q^{n+1} - 1}{q - 1}.$$

Indeed, the vector space $\mathbb{F}^{n+1}$ has $q^{n+1} - 1$ non-zero elements and precisely $q - 1$ vectors represent the same point in $\mathbb{P}(\mathbb{F}_q)$.

Exercise 8.8.2. How many $\mathbb{F}_q$ rational points are contained in the Grassmannian $\mathbb{G}(k, n)$? □

The number of points might vary for different curves of the same genus.

Exercise 8.8.3. Show by examples that a smooth projective curve of genus $g = 4$ defined over $\mathbb{F}_3$ can have any number of $\mathbb{F}_3$-rational points between 0 and 10 points. Prove, that the bound is sharp using the structure of the canonical curve 8.7.2 *Hint:* The ideal of all $\mathbb{F}_q$-rational points in $\mathbb{P}^n$ is given by the 2×2 minors of the matrix

$$\begin{pmatrix} x_0 \; x_1 \; \ldots \; x_n \\ x_0^q \; x_1^q \; \ldots \; x_n^q \end{pmatrix}.$$

For very small q, this ideal is small enough for explizit computations. □

To get an idea about the range for curves in general, Hasse compared the coordinate ring $R = \mathbb{F}_q[x_1, \ldots, x_n]/I(C)$ of a (affine) curve C defined over $\mathbb{F}_q$ with the ring of integers $\mathcal{O}_K$ of a number field K. Both rings are one-dimensional Dedekind domains.
In number theory, the Dedekind zeta-functions

$$\zeta(K, s) = \sum_{\mathfrak{a}} N(\mathfrak{a})^{-s}$$

encodes fundamental information about the number field. Here the sum runs over all non-zero ideals $\mathfrak{a} \subset \mathcal{O}_K$, and $N(\mathfrak{a}) = |\mathcal{O}_K/\mathfrak{a}|$, the **absolute norm**, denotes the number of elements in the residue ring. The series converges for complex $s = x + iy$ with $x = Re\, s > 1$. The fundamental facts about this function are the following, see e.g. [Neukirch, 199x]:

1. **Euler product.** The zeta-function has a product expansion
$$\zeta(K, s) = \prod_{\mathfrak{p}} (1 - N(\mathfrak{p})^{-s})^{-1}.$$
2. **Functional equation.** $\zeta(K, s)$ has a meromorphic continuation to the whole complex plane with a single simple pole at $s = 1$ and satisfies the functional equation
$$\zeta(K, 1-s) = |d_K|^{s-\frac{1}{2}} (\cos \frac{\pi s}{2})^{r_1+r_2} (\sin \frac{\pi s}{2})^{r_2} (2(2\pi)^{-s} \Gamma(s))^{r_1+2r_2} \zeta(K, s),$$
where r_1, r_2 denote the number of real respectively pairs of complex embedding $\sigma_i : K \hookrightarrow \mathbb{C}$, d_K denotes the discriminant of the number field, and $\Gamma(s) = \int_0^\infty e^{-x} x^s \frac{dx}{x}$ the Gamma-function. In particular, $\zeta(K, s)$ has zeroes at $s = -2, -4, -6, \ldots$, i.e. at the even strictly negative integers, since $\Gamma(s)$ has simple poles at these points. These zeroes are called trivial zeroes of $\zeta(K, s)$.
3. **The class number formula.** $\zeta(K, s)$ has at $s = 1$ a simple pole with residue
$$\lim_{s \to 1} (s-1) \zeta(K, s) = \frac{2^{r_1} (2\pi)^{r_2}}{w |d_k|^{1/2}} Rh,$$

where w is the number of roots of unity in K, and R the **regulator** is the size of a fundamental mesh of the lattice of units, $\operatorname{im}(\mathcal{O}_K^* \to \mathbb{R}^{r_1+r_2-1})$, under the map

$$u \mapsto (\log |(\sigma_1(u)|, \ldots, \log |(\sigma_{r_1+r_2}(u)|) \in \{x \in \mathbb{R}^{r_1+r_2} \mid \sum x_i = 0\}.$$

Finally, h is the **class number** of K, i.e. the order of the group

$$Cl(\mathcal{O}_K) = \{ \text{ fractional ideals } \}/\{ \text{ principal ideals } \}.$$

4. **The generalized Riemann Hypothesis.** The only nontrivial zeroes of the zeta function lie on the critical line $Re\, s = 1/2$

Remark 8.8.4. The Euler product in case of $\mathbb{Q}$

$$\sum \frac{1}{n^s} = \prod_p \frac{1}{(1-p^{-s})}$$

implies unique factorization for the integers. The class number h measures the deviation from unique factorization in $\mathcal{O}_K$. The (generalized) Riemann hypothesis is perhaps the most famous conjecture im mathematics. The hypothesis in case of $K = \mathbb{Q}$ implies the optimal assymptotic expansion for the density of primes

$$|\{p \text{ prime number} \mid p \leq x\}| = \int_2^x \frac{dt}{\log t} + O(x^{1/2} \log x),$$

as Riemann pointed out. The much weaker statement, that the nontrivial zeroes of $\zeta(\mathbb{Q}, s)$ lie in the critical strip $\{s \mid 0 < Re\, s < 1\}$, already implies the **Prime Number Theorem**

$$|\{p \text{ prime number} \mid p \leq x\}| \sim \frac{x}{\log x},$$

see for example Patterson [1988].

Guided by the analogy

number theory	curves over finite fields
K	$\mathbb{F}_q(C)$
ideals $\mathfrak{a}$ in $\mathcal{O}_K$	divisors $D \geq 0$ on C defined over $\mathbb{F}_q$
valuations of K	prime divisors defined over $\mathbb{F}_q$
$N(\mathfrak{a})$	$N(D) = q^{\deg D}$

Hasse made the following definition.

Definition 8.8.5. Let C be a smooth projective curve defined over a finite field $\mathbb{F}_q$. The function

$$\zeta(C,s) = \sum_{D\geq 0} N(D)^{-s},$$

where the sum runs over all effective divisors on C defined over $\mathbb{F}_q$, is called the (congruence) **ζ-function** of C over $\mathbb{F}_q$.

Our first goal is to prove, that the sum converges absolutely for $Re\, s > 1$. Let $e > 0$ be the smallest integer, such that there exists a divisor $D \in \operatorname{Div} C$ defined over $\mathbb{F}_q$. We will see later that $e = 1$. For the moment, we note that $e|2g-2$, because there are canonical divisors defined over $\mathbb{F}_q$.

Lemma 8.8.6. *Let C be a smooth absolutely irreducible projective curve of genus g defined over $\mathbb{F}_q$.*

1. *Let $D_1, D_2 \in \operatorname{Div}^d(C)$ be two divisor classes of the same degree defined over $\mathbb{F}_q$. If $D_1 \equiv D_2$ then there exists a rational function $f \in \mathbb{F}_q(C)$ defined over $\mathbb{F}_q$, such that $(f) = D_1 - D_2$.*
2. *There exist only finitely many divisor classes $[D] \in Pic^0(C)$ defined over $\mathbb{F}_q$.*

Proof. By definition $D_1 \equiv D_2 \Leftrightarrow L(D_1 - D_2) \neq 0$. Now, if D is defined over $\mathbb{F}_q$ then $L(D) = \{f \in \overline{\mathbb{F}_q}(C) \mid (f) + D \geq 0\}$ is defined over $\mathbb{F}_q$, which means, that it has a basis of functions $f \in \mathbb{F}_q(C)$. This implies the first statement. For the second, consider the integer e as defined above and some divisor E of degree e defined over $\mathbb{F}_q$. Let D be a divisor of degree 0. Consider an integer $n > 0$ such that $ne > g-1$. Then, by Riemann-Roch $L(D+nE) \neq 0$. Choose a divisor $D' \in |D+nE|$ defined over $\mathbb{F}_q$. Then $D \equiv D' - nE$. The second assertion follows, because there are only finitely many divisors defined over $\mathbb{F}_q$ for any given degree d, in this case degree $d = ne$. Indeed, an effective divisor of degree d defined over $\mathbb{F}_q$ is simply a $\mathbb{F}_q$-rational point of the symmetric product $C^{(d)}$. □

Definition 8.8.7. The order h of the group $\operatorname{Pic}^0(C)(\mathbb{F}_q)$ of divisor classes of degree 0 defined over $\mathbb{F}_q$ is called the **class number** of C over $\mathbb{F}_q$.

Remark 8.8.8. Note, that $h = 1$ for curves of genus $g = 0$. Indeed, by Riemann-Roch any two divisors of the same degree are linearly equivalent.

Proposition 8.8.9 (Rationality of the congruence ζ-function, preliminary version). *Let C be an absolutely irreducible smooth projective curve of genus g defined over $\mathbb{F}_q$. Let e be the minimal degree of strictly effective divisors on C, and let h denote the class number. The ζ-function*

$$\zeta(C,s) = \sum_{D\geq 0} N(D)^{-s}$$

converges absolutely in the domain $Re\, s > 1$ and is in this domain a rational function of q^{-s} of the form

$$\zeta(C,s) = F(q^{-s}) + \frac{hq^{1-g}q^{(1-s)\max(0,2g-2+e)}}{(q-1)(1-q^{e(1-s)})} - \frac{h}{(q-1)(1-q^{-es})},$$

where F is a polynomial of degree at most $2g-2$.

Proof. We only treat the case $g \geq 1$ leaving the easier case $g = 0$ as an exercice. We partition the effective divisors by their degree $\deg D = ne$ refined by the h divisor classes $[nE + H_1], \ldots, [nE + H_h]$, where $H_1, \ldots H_h$ represent the different $\mathbb{F}_q$-rational divisor classes of degree zero. Since

$$|D| = \mathbb{P}^{\ell(D)-1} = \mathbb{P}^{\deg D - g}$$

for any of these divisors, $|D|$ contains precisely

$$\frac{q^{\ell(D)}-1}{q-1}$$

different $\mathbb{F}_q$-rational divisors by Example 8.8.1. Thus, we have the difference

$$\sum_{D\geq 0}(q^{-s})^{\deg D} = \sum_{n\geq 0}\sum_{j=1}^{h}\frac{q^{\ell(nE+H_j)-sne}}{q-1} - \sum_{n\geq 0}\sum_{j=1}^{h}\frac{q^{-nes}}{q-1}.$$

The second sum is a geometric series, which has the limit

$$\sum_{n\geq 0}\sum_{j=1}^{h}\frac{q^{-nes}}{q-1} = \frac{h}{(q-1)(1-q^{-es})}$$

as $|q^{-s}| < 1$. The first sum becomes a geometric series, if we ignore the beginning terms, since once $ne > 2g-2$ we know the value $\ell(nE + H_j) = ne + 1 - g$ precisely. This gives a contribution

$$\sum_{n>(2g-2)/e}\sum_{j=1}^{h}\frac{q^{ne+1-g-nes}}{q-1} = \frac{hq^{1-g}q^{(2g-2+e)(1-s)}}{(q-1)(1-q^{e(1-s)})},$$

as $|q^{1-s}| < 1$ by our assumption $Re\, s > 1$. The remaining finite sum

$$F(q^{-s}) = \sum_{n=0}^{(2g-2)/e}\sum_{j=0}^{h}\frac{q^{\ell(nE+H_j)-sne}}{q-1} = \sum_{n=0}^{(2g-2)/e}\Big(\sum_{j=0}^{h}\frac{q^{\ell(nE+H_j)}}{q-1}\Big)(q^{-s})^{ne}$$

is a polynomial of degree $\leq 2g-2$ in q^{-s}. □

Exercise 8.8.10. Prove Proposition 8.8.9 in case $g = 0$. □

Corollary 8.8.11. $\zeta(C,s)$ *extends to a meromorphic function on the complex plane, which has with simple poles at the points* $s \in \{\frac{2\pi in}{e\log q} \mid n \in \mathbb{Z}\} \cup \{1 - \frac{2\pi im}{e\log q} \mid m \in \mathbb{Z}\}$ *and no other singularities.*

Our next goal is an Euler product.

Definition 8.8.12. Let C be an absolutely irreducible smooth projective curve defined over $\mathbb{F}_q$. An ($\mathbb{F}_q$-rational) **prime divisor** P is an $\mathbb{F}_q$-rational strictly effective divisor, which cannot be decomposed in a sum of strictly effective $\mathbb{F}_q$-rational divisors.

It is easy to see, how a prime divisor look like. Suppose $C \subset \mathbb{P}^n$. If $a \in C(\overline{\mathbb{F}_q})$ is a point in the support of P, say a is defined over $\mathbb{F}_{q^d}$, with d minimal. Then $\deg P = d$ and P constitutes of the orbit

$$P = \sum_{i=0}^{d-1} F^i(a)$$

of a under the Frobenius morphism

$$F : C \to C, a = [a_0 : \ldots : a_n] \mapsto [a_0^q : \ldots : a_n^q].$$

From another point of view, prime divisors are simply prime ideals of dimension one in the two-dimensional homogeneous coordinate ring $\mathbb{F}_q[x_0,\ldots,x_n]/\mathrm{I}(C)$.

Like the decomposition of positive integers into a product of primes, the decomposition of an effective divisor into a sum of prime divisors is unique.

Theorem 8.8.13 (Euler product). *Let C be an absolutely irreducible smooth projective curve defined over $\mathbb{F}_q$. Then*

$$\zeta(C,s) = \prod_P (1 - q^{-s\deg P})^{-1}$$

for $Re\, s > 1$.

Proof. For any given bound M, there are only finitely many prime divisors P with $\deg P \leq M$. Suppose, these are $P_1,\ldots,P_N$. Since $(1 - q^{-s\deg P}) = \sum_{n=0}^\infty q^{-s\deg nP}$, the product

$$\prod_{P,\deg P\leq M} (1 - q^{-s\deg P}) = \sum_{D'} q^{-s\deg D'}$$

where the last sum runs over all $\mathbb{F}_q$-rational effective divisors $D' = n_1P_1 + \ldots + n_NP_N$, i.e. divisors with no prime summand of degree $> M$. Thus, the absolute convergence of $\zeta(C,s) = \sum_D q^{-s\deg D}$ implies

$$\zeta(C,s) = \lim_{M\to\infty} \prod_{P,\deg P\leq M} (1 - q^{-s\deg P})^{-1},$$

which is the desired formula. □

Corollary 8.8.14. *Any absolutely irreducible smooth projective curve C over a finite field $\mathbb{F}_q$ has an $\mathbb{F}_q$-rational divisor E of degree $e = 1$. In particular,*

1. *if C has genus $g = 0$, then $C \cong \mathbb{P}^1$ over $\mathbb{F}_q$, and,*
2. *if C has genus 1, then C has an $\mathbb{F}_q$-rational point, and hence $C(\mathbb{F}_q)$ carries the structure of an abelian group.*

Proof. Suppose the minmal degree $e > 1$. We will derive a contradiction by comparing the the ζ-function $\zeta(C, s) = \zeta(C/\mathbb{F}_q, s)$ with the $\zeta(C/\mathbb{F}_{q^e}, s)$ of C as a curve defined over $\mathbb{F}_{q^e}$. Let P an $\mathbb{F}_q$-rational prime divisor of degree $\deg P = d$, say

$$P = \sum_{i=0}^{d-1} F^i(a),$$

where $a \in C(\overline{\mathbb{F}_q})$ is a point in the the support of P. By assumption $e|d$. Let $k = \frac{d}{e}$ denote the quotient. As $\mathbb{F}_{q^e}$-rational divisor, P decomposes into e distinct $\mathbb{F}_{q^e}$-rational prime divisors

$$P_j = \sum_{i=0}^{k-1} F^{ei+j}(a) \text{ for } j = 0, \dots, e-1$$

of degree k. Conversely, any $\mathbb{F}_{q^e}$-rational prime divisor P' gives rise to a single $\mathbb{F}_q$-rational prime divisor $P = \sum_{j=0}^{e-1} F^j(P')$. Hence, we have a bijection of a single Euler factor of $(1 - q^{-s \deg P})^{-1}$ of $\zeta(C/\mathbb{F}_q, s)$ with the product of e conjugated Euler factors $\prod_{j=0}^{e-1}(1 - (q^e)^{-s \deg P_j})^{-1}$ of $\zeta(C/\mathbb{F}_{q^e}, s)$. Since $e \deg P_j = \deg P$, we conclude that

$$\zeta(C/\mathbb{F}_{q^e}, s) = \zeta(C/\mathbb{F}_q, s)^e.$$

But this is impossible, since by Corollary 8.8.11 the meromorphic continuations of both ζ-functions have a simple pole at $s = 1$. Thus $e = 1$.

For the further statements, we note that for a curve C of genus 0, any $\mathbb{F}_q$-rational divisor E of degree 1 and any choice of a $\mathbb{F}_q$-rational basis of $L(E)$ gives an isomorphism

$$\varphi_E : C \to \mathbb{P}^1$$

over $\mathbb{F}_q$. For the second item, we note that a divisor E of degree $\deg E = 1 > 2g - 2 = 0$ on a curve of genus $g = 1$ has $\ell(E) = 1$ by Riemann-Roch. Hence, $|E| = \{p\}$ is non-empty, which gives us the $\mathbb{F}_q$-rational point p. □

Exercise 8.8.15. Prove by example, that a smooth cubic plane curve over $\mathbb{F}_q$ might contain only one $\mathbb{F}_q$-rational point, if $q + 1 - 2\sqrt{q} < 1 (\Leftrightarrow q < 4)$. □

Now knowing $e = 1$, we can simplify Proposition 8.8.9.

Theorem 8.8.16 (Rationality of $\zeta(C,s)$). *The congruence ζ-function $\zeta(C,s) = \sum_{D\geq 0} q^{-s\deg D}$ converges in the domain $Re\, s > 1$ absolutely towards a rational function of q^{-s} of type*

$$\zeta(C,s) = \frac{P(q^{-s})}{(1-q^{-s})(1-qq^{-s})},$$

where

$$P(t) = 1 - a_1 t + \ldots - a_{2g-1}t^{2g-1} + q^g t^{2g}$$

is a polynomial with integer coefficients of degree $2g$ with constant term $a_0 = 1$ and leading coefficient $a_{2g} = q^g$.

Proof. We treat only the case $g \geq 1$. From Proposition 8.8.9 and its proof we know

$$\zeta(C,s) = F(q^{-s}) + \frac{hq^{1-g}q^{(2g-1)(1-s)}}{(q-1)(1-q^{1-s})} - \frac{h}{(q-1)(1-q^{-s})},$$

where

$$\begin{aligned} F(q^{-s}) &= \sum_{j=1}^{h} \frac{q^{\ell(H_j)}}{q-1} \quad + \sum_{d=1}^{2g-3}\sum_{j=1}^{h} \frac{q^{\ell(dE+H_j)-sd}}{q-1} + \sum_{j=1}^{h} \frac{q^{\ell((2g-2)E+H_j)-(2g-2)s}}{q-1} \\ &= \frac{h-1+q}{q-1} \quad + \sum_{d=1}^{2g-3}\sum_{j=1}^{h} \frac{q^{\ell(dE+H_j)}}{q-1}q^{-sd} + \frac{(h-1+q)q^{g-1}}{q-1}q^{-(2g-2)s}, \end{aligned}$$

since precisely one class of degree 0 is effective, and only one class of degree $2g-2$, namely the canonical, is special. Taking all terms to the common denominator

$$\frac{1}{(1-q^{-s})(1-q^{1-s})},$$

we get as numerator a polynomial of degree $\leq 2g$ in q^{-s} with rational coefficients. The contribution to the constant term arizes from the first term of $F(q^{-s})$ and the principal part of $\zeta(C,s)$. More precisely,

$$a_0 = \frac{h-1+q}{q-1} - \frac{h}{q-1} = 1.$$

Similary,

$$a_{2g} = \frac{(h-1+q)q^{g-1}}{q-1}q + \frac{hq^{1-g}q^{2g-1}}{q-1}(-1) = q^g.$$

Finally, to see that all coefficients of P are integers, we note that the power series expansion of $\zeta(C,s)$ in the variable $t = q^{-s}$ has only integral coefficients, since it is the generating function for effective divisors $Z(C,t) = \sum_{D\geq 0} t^{\deg D}$. Hence all the coefficients $P(t)$ have to be integers as the denominator $(1-t)(1-qt)$ is an integer polynomial. □

Theorem 8.8.17 (Functional equation). *The congruence ζ-function of an absolutely irreducible smooth projective curve C defined over $\mathbb{F}_q$ satisfies the functional equation*

$$\zeta(C, 1-s) = q^{(g-1)(2s-1)}\zeta(C, s).$$

Proof. We decompose

$$\zeta(C, s) = \zeta_1(C, s) + \zeta_2(C, s) + \zeta_3(C, s)$$

into three summands which satisfy the functional equation individually. We take

$$\zeta_1(s) = \frac{h-1+q}{q-1}(1 + q^{g-1-(2g-2)s}) = \frac{h-1+q}{q-1}(1 + q^{-(g-1)(2s-1)})$$

$$\zeta_2(s) = \frac{h}{q-1}(\frac{q^{g-(2g-1)s}}{(1-q^{1-s})} - \frac{1}{(1-q^{-s})})$$

and

$$\zeta_3(s) = \sum_{d=1}^{2g-3}\sum_{j=1}^{h}\frac{q^{\ell(dE+H_j)}}{q-1}q^{-sd}.$$

For the first two summands the functional equation is an easy computation:

$$\begin{aligned}\zeta_1(1-s) &= \frac{h-1+q}{q-1}(1 + q^{-(g-1)(1-2s)})\\ &= q^{(g-1)(2s-1)}\frac{h-1+q}{q-1}(q^{-(g-1)(2s-1)} + 1) = q^{(g-1)(2s-1)}\zeta_1(s)\end{aligned}$$

and

$$\begin{aligned}\zeta_2(1-s) &= \frac{h}{q-1}(\frac{q^g q^{-(2g-1)(1-s)}}{1-q^s} - \frac{1}{1-q^{s-1}})\\ &= \frac{q^{(g-1)(2s-1)}h}{q-1}(\frac{1}{q^{-s}-1} - \frac{q^{1-s+(1-g)(2s-1)}}{q^{1-s}-1}) = q^{(g-1)(2s-1)}\zeta_2(s).\end{aligned}$$

The functional equation for $\zeta_3(s)$ follows from the Riemann-Roch Theorem: In $\zeta_3(s)$, the sum runs over all divisor classes with

$$1 \le \deg D \le 2g-3.$$

The classes of $K-D$ run through the same set. We compare the summand corresponding to $K-D$ in $\zeta_3(C, 1-s)$ with the summand corresponding to D in $q^{(g-1)(2s-1)}\zeta_3(C, s)$. The term in $(q-1)\zeta_3(C, 1-s)$ is

$$\begin{aligned} q^{\ell(K-D)-(1-s)\deg(K-D)} &= q^{\ell(D)-\deg D+g-1-(1-s)(2g-2-\deg D)} \\ &\quad \text{by Riemann-Roch,} \\ &= q^{(g-1)(2s-1)+\ell(D)-s\deg D}, \end{aligned}$$

which is the corresponding term in $(q-1)q^{(g-1)(2s-1)}\zeta_3(s)$. This proves the Theorem. □

Frequently, the ζ-function is studied via the substitution $t = q^{-s}$ as the generating function

$$Z(C,t) = \sum_{D\geq 0} t^{\deg D},$$

where the sum runs over all effective divisors on C defined over $\mathbb{F}_q$. Since $\zeta(C,s)$ converges for $Re\, s > 1$, the formal power series $Z(C,t)$ converges in the disc $|t| \leq q^{-1}$

This function enumerates also the number of points in C over all fields $\mathbb{F}_q^r$ as follows.

Proposition 8.8.18. *Let C be a smooth absolutely irreducible curve defined over $\mathbb{F}_q$. Let $N_r = |C(\mathbb{F}_{q^r})|$ denote the number of points on C defined over $\mathbb{F}_{q^r}$. Then*

$$\zeta(C,s) = \exp(\sum_{r\geq 1} N_r \frac{t^r}{r}),$$

where $t = q^{-s}$.

Proof. We use the Euler-product

$$\zeta(C,s) = \prod_{\mathfrak{p}} (1 - q^{-s\deg \mathfrak{p}})^{-1} = \prod_{\mathfrak{p}} (1 - t^{\deg \mathfrak{p}})^{-1}.$$

Taking the logarithm

$$\log Z(C,t) = \sum_{\mathfrak{p}} \sum_{r=1}^{\infty} \frac{t^{r\deg \mathfrak{p}}}{r},$$

we see that we have to prove

$$N_r = \sum_{\deg \mathfrak{p}|r} \deg \mathfrak{p}.$$

This holds, because any prime divisor $\mathfrak{p}$ defined over $\mathbb{F}_q$ with $\deg \mathfrak{p} \mid r$ gives rise to $\deg \mathfrak{p}$ points, with each point defined over $\mathbb{F}_{q^{\deg \mathfrak{p}}} \subset \mathbb{F}_{q^r}$. Conversely, $C(\mathbb{F}_q^r)$ is partitioned into the points of such prime divisor classes. □

The last formula for the $Z(X,t)$ makes sense for any variety over $\mathbb{F}_q$.

Definition 8.8.19. Let X be a variety defined over $\mathbb{F}_q$, and let $N_r = X(\mathbb{F}_r)$. The **Weil zeta-function** is defined as

$$Z(X,t) = \exp(\sum_{r\geq 1} N_r \frac{t^r}{r}).$$

As the definition stands, $Z(X,t) \in \mathbb{Q}[[t]]$ is just a formal power series. Weil conjectured, that $Z(X,t)$ is a rational function. This was proved by Dwork [] using p-adic analysis and later by Grothendieck using his étale cohomology [] again:

Theorem 8.8.20 (Dwork, Grothendieck). *If X is projective of dimension n, then*

$$Z(X,t) = \frac{P_1(t)P_3(t)\dots P_{2n-1}(t)}{P_0(t)P_2(t)\dots P_{2n}(t)}$$

for certain integral polynomials $P_i(t) \in \mathbb{Z}[t]$.

Exercise 8.8.21. Compute the Weil zeta-function of $\mathbb{P}^n$ and of $\mathbb{G}(k,n)$! □

The proof of Theorem 8.8.20 is far beyond the techniques of the book, however, we have established this for curves.

Corollary 8.8.22. *If C is a smooth absolutely irreducible curve of genus $g$defined over $\mathbb{F}_q$ then*

$$Z(C,t) = \frac{P(t)}{(1-t)(1-qt)}$$

where $P(t) = 1 - a_1 t + a_2 t^2 - \dots + q^g t^{2g} \in \mathbb{Z}[t]$.

Proof. This is just a reformulation of Theorem 8.8.16. □

With the Weil zeta-function, we obtain a nice closed formula for the number of rational points over $\mathbb{F}_{q^r}$. Let

$$P(t) = \prod_{j=1}^{2g}(1-\alpha_j t).$$

We call the algebraic intergers α_j the **reciprocial roots** of the numerator $P(t)$ of the Weil zeta-function.

Corollary 8.8.23. *Let α_j denote the reciprocal roots of the numerator of the Weil zeta-function of a smooth projective absolutely irreducible curve of genus g defined over $\mathbb{F}_q$. Then the number of points on C defined over $\mathbb{F}_{q^r}$ is*

$$N_r = q^r + 1 - \alpha_1^r \dots - \alpha_{2g}^r.$$

Proof. We take the logarithm on both sides in the formula

$$\exp(\sum_{r\geq 1} N_r \frac{t^r}{r}) = \frac{(1-\alpha_1 t)\dots(1-\alpha_{2g}t)}{(1-t)(1-qt)}$$

and obtain

$$\sum_{r\geq 1} N_r \frac{t^r}{r} = \log(1-qt) + \log(1-t) - \sum_{j=1}^{2g} \log(1-\alpha_j t).$$

Now use the power series expansion of the logaritm $\log(1-x) = \sum_{r=1}^{\infty} \frac{x^r}{r}$ and compare coefficients. □

Theorem 8.8.24 (Hasse-Weil). *Let C be a smooth projective absolutely irreducible curve of genus g defined over $\mathbb{F}_q$. The number N_r of $\mathbb{F}_{q^r}$-rational points on C is bounded by*

$$|N_r - q^r - 1| \leq 2g\sqrt{q}.$$

This theorem is, using Corollary 8.8.23, an immediate consequence of the following more precise result.

Theorem 8.8.25 (Analog of the Riemann hypothesis). *The reciprocal roots α_j of the numerator $P(t)$ of the Weil zeta-function of a smooth projective absolutely irreducible curve defined over $\mathbb{F}_q$ satisfy*

$$|\alpha_j| = \sqrt{q}.$$

In other words, the only zeroes of $\zeta(C,s) = Z(C,q^{-s})$ have real part $Re\, s = 1/2$. This explains the name of the theorem. To prove the theorem, we consider the logarithmic derivative of $Z(C,t)$ and subtract the parts coming from the denominator:

$$\frac{Z'(X,t)}{Z(X,t)} - \frac{1}{1-t} - \frac{q}{1-qt} = \sum_{j=1}^{2g} \frac{\alpha_j}{1-\alpha_j t}.$$

Thus, $|\alpha_j| \leq \sqrt{q}$ holds, iff the radius of convergence R of the power series

$$\frac{Z'(X,t)}{Z(X,t)} + \frac{1}{1-t} + \frac{1}{1-qt} = \sum_{r=1}^{\infty} (N_r - q^r - 1)t^{r-1}$$

is bounded from below by $R \geq 1/\sqrt{q}$. So an inequality of type

$$N_r - q^r - 1 \leq cq^{r/2},$$

holding for almost all r with a constant c independent from r, will imply half of the claim in Theorem 8.8.25. The other half, $|\alpha_j| \geq \sqrt{q}$, or equivalently, $\zeta(C,s)$ has no zeroes s with $Re\, s < 1/2$, will follow from the first half using the functional equation 8.8.17. We conclude, that basically the Hasse-Weil Theorem and the Theorem on the Riemann hypothesis of the congruence zeta-function are equivalent.

To prove the desired bound on the number of $\mathbb{F}_{q^r}$-rational points we follow an ingenious argument of Stepanov [1974]. In view of the formula for N_r rational points, Corollary 8.8.23, it suffices to prove such bound after passing from $\mathbb{F}_q$ to some larger field $\mathbb{F}_{q^\rho}$. So we may assume, that C has an $\mathbb{F}_q$-rational point and that $q = p^{2\tau}$ is an even power of the characteristic $p = \mathrm{char}(\mathbb{F}_q)$.

Lemma 8.8.26 (Stepanov). *Let C be a smooth projective absolutely irreducible curve of genus g defined over $\mathbb{F}_q$. Suppose, that C contains an $\mathbb{F}_q$ rational point, that $q = p^{2\tau}$ is an even power of $p = \mathrm{char}(\mathbb{F}_q)$ and that $q^r > (g+1)^4$. Then*

$$N_r \leq q^r + 1 + (2g+1)q^{r/2}.$$

Proof. Let $a \in C$ be an $\mathbb{F}_q$-rational point. The basic idea of Stepanov's proof is to construct a rational function f on C which has only a pole in a and which vanishes at all $\mathbb{F}_{q^r}$ rational points except a with fairly high multiplicity. The bound for N_r will follow from the fact, that the number of poles and the number of zeroes of a rational function counted with multiplicity is the same. We are thus led to study the Riemann-Roch spaces $L(ma)$. For $m > 2g-2$, this space has dimension $\ell = \ell(ma) = m+1-g$. There exist a basis $f_1 = f_1, \ldots, f_\ell$ of $L(ma)$, such that $0 = v_a(f_1) > v_a(f_2) > \ldots > v_a(f_\ell)$.

Each $f_j = f_j(x)$ is a rational function of degree zero in the homogeneous coordinates $x = [x_0 : \ldots : x_k]$ of some projective embedding $C \hookrightarrow \mathbb{P}^k$. Hence, $f_j(x^{q^r})$ is a rational function as well, with $v_a(f_j(x^{q^r}) = q^r v_a(f_j(x))$, and the same values on all points of $C(\mathbb{F}_{q^r})$, because $[x_0 : \ldots : x_k] = [x_0^{q^r} : \ldots : x_k^{q^r}]$ holds on the points of $\mathbb{P}^k(\mathbb{F}_{q^r})$. We will construct our desired function f as a linear combination

$$f(x) = u_1^{p^k}(x) f_1(x^{q^r}) + \ldots + u_\ell^{p^k}(x) f_\ell(x^{q^r}),$$

where the u_j are functions in $L(na)$ (for a suitable n), which we choose in such a way, that

$$u_1^{p^\tau}(x) f_1(x) + \ldots + u_\ell^{p^\tau}(x) f_\ell(x) = 0 \in L((np^\tau + m)a),$$

holds.

Claim 1. Suppose that $m, n > 2g-2$ and that

$$(n+1-g)(m+1-g) > np^k + m + 1 - g$$

holds. Then there exists a non-trivial expression

$$u_1^{p^k}(x) f_1(x) + \ldots + u_\ell^{p^k}(x) f_\ell(x) = 0 \in L((np^k + m)a)$$

Indeed, for each function $u_j \in L(na)$, we have q^{n+1-g} many choices, because $L(na)$ is a vector space of dimension $n+1-g$ defined over $\mathbb{F}_q$. So, there

are $q^{(n+1-g)(m+1-g)}$ many choices for the expression. On the other hand, $L(np^k+m)$ contains only $q^{np^k+m-1-g}$ functions defined over $\mathbb{F}_q$. So two of these expressions give the same function and their difference can be written in the same form, because $u_j^{p^k} - \tilde{u}_j^{p^k} = (u_j - \tilde{u}_j)^{p^k}$.
Claim 2. Suppose $np^k < q^r$ then any non-trivial expression

$$f(x) = u_1^{p^k}(x) f_1(x^{q^r}) + \dots u_\ell^{p^k}(x) f_\ell(x^{q^r})$$

is a non-trivial function in $L((mq^r + np^k)a)$.

Indeed, since $v_a(f_j(x)) \geq v_a(f_{j+1}(x)) + 1$, we have

$$v_a(f_j(x^{q^r}) - v_a(f_{j+1}(x^{q^r}) \geq q^r,$$

and the coefficients $u_j(x)^{p^k}$ cannot fill the gap to make the pole order at a of two consecutive summands to agree. Thus, the pole order of these expression is attained at the summand $u_j^{p^k} f_j(x^{q^r})$ with $j = \max\{i \mid u_i \neq 0 \in L(na)\}$. In particular, the expression is non-trivial.

Suppose the conditions of both Claims are satisfied. Then f of Claim 2 constructed with the coefficients u_j as in Claim 1 is a function with a pole of order $-v_a(f) \leq mq^r + np^k$ and, which vanishes in all other $\mathbb{F}_{q^r}$ rational points, because $f_j(x) = f_j(x^{q^r})$ holds for such points and because of the choice with Claim 1. Moreover, since $f_j(x^{q^r}) = (f_j(x))^{q^r}$ and $p^k \mid q^r$ the function $f(x)$ is the p^k-th power of the function $g(x) = u_1(x) f_1^{p^{2\tau r-k}}(x) + u_\ell(x) f_\ell^{p^{2\tau r-k}}(x)$. So the vanishing order in each of these points it at least p^k. We conclude, that

$$(N_r - 1)p^k \leq np^k + mq^r,$$

since the number of poles and zeroes of a rational function counted with multiplicities coincide.

Choose $p^k = q^{r/2}$, i.e. $k = \tau r$, and take $n = q^{r/2} - 1$ and $m = q^{r/2} + 2g$. The condition of Claim 2 $np^k = (q^{r/2}-1)q^{r/2} < q^r$ is satisfied. The inequality $(n+1-g)(m+1-g) = (q^{r/2}-g)(q^{r/2}+1+g) = (q^r - (g+1)^2 + m + 1 - g > nq^{r/2} + m + 1 - g = q^r - q^{r/2} + m + 1 - g$ holds, since $q^{r/2} \geq (g+1)^2$ holds by the assumption on r in the Lemma. Finally, $m = q^{r/2} + 2g > 2g - 2$ and $n = q^{r/2} - 1 \geq (g+1)^2 - 1 = g^2 + 2g > 2g - 2$ are also satisfied. We therefore may apply both Claims and obtain

$$(N_r - 1)q^{r/2} \leq (q^{r/2} - 1)q^{r/2} + (q^{r/2} + 2g)q^r \Leftrightarrow N_r \leq q^r + (2g+1)q^{r/2}.$$

This proves Stepanov's Lemma 8.8.26, and hence the Riemann hypothesis for the congruence ζ-function 8.8.25 and the Hasse-Weil Theorem 8.8.24. □

Inspired by the result on curves, Weil made a series of very important conjectures about the nature of the Weil zeta-function in general, which were proved later by Grothendieck and Deligne.

Theorem 8.8.27 (Grothendieck, Deligne). *Let X be a variety defined over a finite field $\mathbb{F}_q$, and let*

$$Z(X,t) = \exp(\sum_{r\geq 1} N_r \frac{t^r}{r})$$

with $N_r = |X(\mathbb{F}_{q^r})|$ be the corresponding zeta-function. Then

1. **Rationality.** *$Z(X,t)$ is a rational function*

$$Z(X,t) = \frac{P_1(t)P_3(t)\dots P_{2n-1}(t)}{P_0(t)P_2(t)\dots P_{2n}(t)}$$

with the $P_i(t)$ are integral polynomials with constant term 1.

2. **Topology.** *Suppose X is defined by reduction mod p of a system of equations with $\mathbb{Z}$ coefficients. Let $X(\mathbb{C})$ be the corresponding analytic variety equipped with the Euclidean topology. Then*

$$\deg P_i(t) = b_i = \dim \mathrm{H}^i(X(\mathbb{C}),\mathbb{Q})$$

is the i-th (topological) Betti number of $X(\mathbb{C})$.

3. **Functional equation.** *Suppose X is smooth projective and absolutely irreducible. Then*

$$Z(X,\frac{1}{q^n t}) = \pm q^{ne/2} t^e Z(X,t),$$

where $e = \sum_{i=0}^{2n}(-1)^i \deg P_i$. (The number e coincides with the Euler number of $X(\mathbb{C})$ if X arizes by reduction mod p.)

4. **Analog of the Riemann-Hypothesis.** *The reciprocal roots of $P_i(t) = \prod_{j=1}^{b_i}(1-\alpha_{ij}t)$ are algebraic integers of absolute value*

$$|\alpha_{ij}| = q^{i/2}.$$

Weil also outlined an approach to the proof of his conjecture. He explained, that if one has a cohomology theory for algebraic varieties, which coincides in case of the ground field $\mathbb{C}$ with ordinary cohomology, but which also works for finite ground fields, then one could get the number of N_r of $\mathbb{F}_{q^r}$-rational points, which are the fix points of the geometric Frobenius automorphism F^r via a Lefschetz fix point formula. The $P_i(t)$ could then be interpreted as the characteristic polynomials of the induced action of the Frobenius F on the i-th cohomology group.

It was quickly clear, that the Zariski topology is too rough for this purpose. Grothendieck developed his étale topology and ℓ-adic cohomology theory precisely for this purpose, and proved the first three parts of the conjecture. The final step, the proof of the analog of the Riemann Hypothesis was taken by Deligne [1974].

Current interest in arithmetic geometry is concerned with the question how the polynomials $P_i(t)$ vary with the prime p, if we consider the reduction mod p of a variety defined by equations with integer coefficients. This leads to the study of L-functions as already mentioned in Chapter 5.

8.9 What to read next?

The next step in mastering algebraic geometry after reading this book, is to get familiar with the concept of coherent sheaves and their cohomology. The standard text for this topic is Hartshorne book, Chapter II and III, for which this book is a good preparation.

We plan to write a different more computational approach to this topic in a further textbook.

References

Atiyah, M.F.; McDonald, I.G. (1969): *Introduction to commutative algebra.* Addison-Wesley.

Avramov, L.L. (1989): Modules of finite virtual projective dimension. *Invent. Math.* **96** (1989), 71–101.

Barth, W. (1996): Two projective surfaces with many nodes, admitting the symmetries of the icosahedron. *J. Algebraic Geom.* **5** (1996), 173–186.

Barth, W.; Hulek, K.; J.; Peters, C.A.M.; Van de Ven, A. (2004): *Compact Complex Surfaces.* Second enlarged edition. Springer-Verlag.

Bayer, D.; Stillman, M. (1987): A theorem on refining division orders by the reverse lexicographic orders. *Duke J. Math.* **55**, 321–328.

Bayer, D.; Stillman, M. (1988): On the complexity of computing syzygies. *J. Symb. Comput.* **6**, 135–147.

Böhm, J. (1999): *Parametrisierung rationaler Kurven.* Diplomarbeit, Universität Bayreuth.

Brieskorn, E.; Knörrer, H. (1986): *Plane algebraic curves.* Birkhäuser.

Bruns, W.; Herzog, J. (1993): *Cohen-Macaulay rings.* Cambridge Univ. Press.

Buchberger, B. (1965): *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenrings.* Dissertation, Universität Innsbruck.

Buchberger, B. (1970): Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems. *Aequationes mathematicae* **4**, 374–383.

Cartan, H; Eilenberg, S. (1956): *Homological algebra. With an appendix by David A. Buchsbaum.* Reprint of the 1956 original. Princeton Univ. Press.

Ciliberto, C; Miranda, R. (2000): Linear systems of plane curves with base points of equal multiplicity. *Trans. Amer. Math. Soc.* **352**, 4037–4050.

Clebsch, A. (1871): Über die Anwendung der quadratischen Substitution auf die Gleichungen 5^{ten} Grades und die geometrische Theorie des ebenen Fünfseits. *Math. Ann.* **4**, 284–345.

Clemens, C.H.; Griffiths, P.A. (1972): The intermediate Jacobian of the cubic threefold. *Ann. of Math. (2)* **95**, 281–356.

Cox, D.; Little, J.; O'Shea, D. (1997): *Ideals, Varieties, and Algorithms.* 2nd edition, Springer-Verlag.

Cox, D.; Little, J.; O'Shea, D. (1998): *Using Algebraic Geometry.* Springer-Verlag.

Decker, W., Greuel, G.-M., Pfister, G. (1999): Primary decomposition: algorithms and comparisons, In: B.H. Matzat et al (eds.), *Algorithmic algebra and number theory, Heidelberg 1997*, 187–220, Springer-Verlag.

Decker, W.; Lossen, C. (2006): *Computing in Algebraic Geometry. A Quick Start using SINGULAR* Springer-Verlag.

Decker, W.; Schreyer, F.O. (2001): Computational algebraic geometry today. In: C. Ciliberto et al (eds.): *Application of Algebraic Geometry to Coding Theory, Physics, and Computation,* 65–120, Kluwer.

Dummit, D.S.; Foote, R.M. (2003): *Abstract Algebra.* John Wiley & Sons.

Eisenbud, D. (1995): *Commutative algebra with a view toward algebraic geometry.* Springer-Verlag.

Eisenbud, D. (2005): *The geometry of Syzygies.* Springer-Verlag.

Eisenbud, D.; Grayson, D.R.; Stillman M.; Sturmfels, B. eds. (2002): *Computations in Algebraic Geometry with Macaulay 2.* Springer-Verlag.

Eisenbud, D.; Harris, J. (2000): *The geometry of schemes.* Graduate Texts in Mathematics, 197. Springer-Verlag.

Eisenbud, D.; Sturmfels, B. (1996). Binomial Ideals. *Duke J. Math.* **84**, 1–45.

van den Essen, E. (2000): *Polynomial Automotphisms.* Birkhäuser.

Flenner, H.; O'Carroll, L.; Vogel, W. (1999): *Joins and intersections.* Springer Monographs in Mathematics. Springer-Verlag.

Fulton, W. (1998): *Intersection theory.* Second edition. A Series of Modern Surveys in Mathematics. Springer-Verlag.

von zur Gathen, J.; Gerhard, J. (1999): *Modern computer algebra.* Cambridge Univ. Press.

Gekeler, E.-U. (2003): Frobenius distributions of elliptic curves over finite prime fields. *Int. Math. Res. Not.* **37**, 1999–2018.

Gelfand, I.M.; Kapranov, M.M.; Zelevinsky, A.V. (1994): *Discriminants, resultants and multidimensional determinants.* Birkhäuser Verlag.

Gordan, P. (1899): Neuer Beweis des Hilbertschen Satzes über homogene Funktionen. *Nachrichten König. Ges. der Wiss. zu Gött.*, 240–242.

Grauert, H. (1972): Über die Deformation isolierter Singularitäten analytischer Mengen. *Invent. Math.* **15**, 171–198.

Greuel, G.-M.; Pfister, G. (2002): *A SINGULAR introduction to commutative algebra.* Springer-Verlag.

Gröbner, W. (1951) Über den Multiplizitätsbegriff in der algebraischen Geometrie. *Math. Nachr.* **4**, 193–201.

Harris, J. (1992): *Algebraic Geometry.* Springer-Verlag.

Hartshorne, R. (1977): *Algebraic Geometry.* Springer-Verlag.

Hermann, G. (1926): Die Frage der endlich vielen Schritte in der Theorie der Polynomideale. *Math. Ann.* **95**, 736–788.

Hilbert, D. (1890): Über die Theorie der algebraischen Formen. *Math. Ann.* **36**, 473–534.

Hilbert, D. (1893): Über die vollen Invariantensysteme. *Math. Ann.* **42**, 313–373.

Hironaka, H. (1964): Resolution of singularities of an algebraic variety over a field of characteristic zero. *Annals of Math.* **79**. I: 109–203; II: 205–326.

Iskovskikh, V.A.; Manin, Yu.I. (1971): Three-dimensional quartics and counterexamples to the Lüroth problem . *Math. Sb.* **86**, 140–166 = *Math. USSR Sbornik.* **15**, 141–166.

de Jong, T. (1998): An algorithm for computing the integral closure. *J. Symb. Comput.* **26**, 273–277.

Jung, H.W.E. (1942): Über ganze birationale Transformationen der Ebene. *J. Reine Angew. Math.* **184**, 161–174.

Kaltofen, E. (1982): Polynomial factorization. In: B. Buchberger et al (eds.), *Computer algebra*, 95–113, Springer-Verlag.

Kaltofen, E. (1990): Polynomial factorization 1982-1986. In: I. Simon (ed.), *Computers in mathematics*, 285–309. Marcel Dekker, New York.

Kaltofen, E. (1992): Polynomial factorization 1987-1991. In: D.V. Chudnovsky and R.D. Jenks (eds.), *Proceedings of LATIN'92, Sao Paulo*, 294–313. Springer-Verlag.

Kaltofen, E. (2003): Polynomial factorization: a success story. In: J.R. Sendra (ed.), *Proc. ISSAC'03, Philadelphia*. ACM Press, 3–4.

Kline, M. (2000): *Mathematical Thought from Ancient to Modern Times*. Oxford University Press.

Koblitz, N. (1994): *A course in number theory and cryptography.* Second edition. Graduate Texts in Mathematics, 114. Springer-Verlag.

Kollár, J. (1999). Effective Nullstellensatz for arbitrary ideals. *J. Eur. Math. Soc.* **1**, 313–337.

van der Kulk, W. (1953): On polynomial rings in two variables. *Nieuw Archief Vor Wiskunde* **3**, 33–41.

Kunz, E. (1985): *Introduction to commutative algebra and algebraic geometry.* With a preface by David Mumford. Birkhäuser.

Macaulay, F. (1916): *The algebraic theory of modular systems.* Cambridge Univ. Press.

Macaulay, F. (1927): Some properties of enumeration in the theory of modular systems. *Proc. London Math. Soc.* **26**, 531–555.

Mac Lane, S. (1998). *Categories for the working mathematician.* 2nd edition, Springer-Verlag.

Matsumura, H. (1986): *Commutative ring theory.* Cambridge Univ. Press.

Mayr, E.; Meyer, A. (1982): The complexity of the word problem for commutative semigroups and polynomial ideals. *Adv. in Math.* **46**, 305–329.

Möller, H.M.; Mora, F. (1984): Upper and lower bounds for the degree of Gröbner bases. In: *Proceedings EUROSAM 84 (Cambridge, 1984)*, Lecture Notes in Comput. Sci. 174, 172–183, Springer-Verlag.

Mora, T. (1982): An algorithm to compute the equations of tangent cones. In: *Computer algebra, Proceedings EUROCAM '82, Marseille,* 158–165, Springer-Verlag.

Nagata, M. (1962): *Local rings.* Interscience Tracts in Pure and Applied Mathematics, No. 13. John Wiley & Sons.

Newman, P. (1972): *Integral matrices.* Academic Press.

Newton, I. (1710): Curves. In: Lexicon Technicum, John Harris, London. Reprinted in *The Mathematical Works of Isaac Newton,* Volume 2, Johnson Reprint Corporation 1967.

Noether, E. (1921): Idealtheorie in Ringbereichen. *Math. Ann.* **83**, 24–66.

Reid, M. (1988): *Undergraduate Algebraic Geometry.* Cambridge University Press.

Schreyer, F.-O. (1980): *Die Berechnung von Syzygien mit dem verallgemeinerten Weierstraßchen Divisionssatz und eine Anwendung auf analytische Cohen–*

Macaulay Stellenalgebren minimaler Multiplizität. Diplomarbeit, Universität Hamburg.

Schreyer, F.-O. (1991): A Standard Basis Approach to Syzygies of Canonical Curves. *J. Reine Angew. Math.* **421**, 83–123.

Serre, J.-P. (1965) *Algèbre locale. Multiplicités.* Seconde édition. Lecture Notes in Math. 11, Springer-Verlag.

Shafarevich, I.R. (1974): *Basic Algebraic Geometry.* Springer-Verlag.

Silverman, J.H. (1986): *The arithmetic of elliptic curves.* Graduate Texts in Mathematics, 106. Springer-Verlag.

Sturmfels, B. (1996): *Gröbner bases and convex polytopes.* University Lecture Series, 8. AMS, Providence, RI.

Zariski, O.; Samuel, P. (1975–1976). *Commutative Algebra.* Vols. I and II. Corr. 2nd printing of the 1958–1960 edition. Springer-Verlag.

Index