Wolfram Decker and Frank-Olaf Schreyer

Varieties, Gröbner Bases and Algebraic Curves

October 25, 2007

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

Preface

Algebraic Geometry is the study of the geometry of solution sets of algebraic systems of equations. It has a long tradition, being shaped by various schools in time, and there are marevelous text books on the subject availble such as Harris' [1992] and Hartshorne's [1977].

In applications, say coming from engeneering, quite often algebaic manipulation are used, so that one can understand the geometry of the solution set. The theory on the other hand is guided by the classification principle and very little computation with explicit equations occure in text books.

The rise of modern Computer Algebra systems can bridge this gap. The algebraic tools and algorithm origining from Hilbert's landmark paper [1890], work in practise for many interesting cases nowadays. The basic idea of this book is to develop the theory and algorithmic parts of Algebraic Geometry in parallel, so that the theoretically oriented student can explore the field through examples, and the students interested in examples, say coming from engeneering, will understand these better through guiding geometric theory beyond algorithms.

We aim to cover classical Algebraic Geometry up to the point where more advance tools like sheaves and cohomology are truly necessary. Roughly the scope of the book is the content of Fulton text on algebraic curve, whose influence on this book we gratefully acknowledge, with Computer Algebra added on.

What the reader will learn.

The first three Chapters cover Hilbert's landmark paper, which in our eyes is the real starting point of Algebraic Geometry. So Hilbert's Basis Theorem, the Nullstellensatz, the Syzygy Theorem and the Noether Normalization are the basic topics of these three sections. After these section the student will be able to answer the most fundamental questions about solutions of an algebraic system of equations:

Let $f_1, \ldots, f_r \in \mathbb{Q}[x_1, \ldots, x_n]$ *be polynomials and let*

$$V = \mathrm{V}(f_1, \dots, f_r) = \{a \in \mathbb{C}^n \mid f_j(a) = 0, \, j = 1, \dots, r\}$$

denote their common vanishing loci.

1. *Is V non-empty?*
2. *Is V finite? If yes, what is a bound on the number of solutions?*
3. *If V is infinite, what is the dimension of the solution set?*

The theoretical engridients to answer this are all present in Hilbert's paper, while the algorithmic work horse is Gröbner basis theory: The answers to all three question can be read from the initial ideal of a Gröbner basis. Moreover this compution give also a sufficient criterion for Cohen-Macaulayness and hence the equidimensionality. Our proof of the Nullstellensatz, will enable the student to construct explicitely points on V.

Since we want to perform explicit computations, we do not work over an algebracially closed field, but have a field of definition always present, in which we can compute via Computer Algebra. Fortunately all Gröbner basis computation are based on linear algebra, so no field extension are not necessary for these computations. On the other hand, absolute component and primary decomposition uses field extensions. We prove component decomposition as a typical example of Noetherian induction and touch upon primary decomposition. We do not present complete algorithms for these topics.

So in our treatment two fields $\Bbbk \subset \mathbb{K}$ are always present: a field $\Bbbk$, say $\mathbb{Q}$, for which we can use Computer Algebra systems and an algebraically closed field $\mathbb{K}$, say $\mathbb{C}$, over which we consider the solution sets. Our first object of study are not varieties over algebraic closed fields, as in many other text books, but rather algebraic sets defined over arbitrary fields.

After being able to compute the dimension, the distinction between smooth and singular points is next topic. In Chapter 4 we introduce the tangent space and the local ring of an algebraic set a point, and provide the tool to compute in this ring, Mora algorithm. As an application we define the local intersection multiplicities of plane curves algebraically. We prove that an algebraic set defined by c equations in affine space, has codimension at most c.

In Chapter 5 we introduce the projective space, which is needed to to make Bézout's Theorem become true: Two plane curve of degree d and e intersect in precisely de points counted with multiplicity, unless the curve have common component. A bound on the number of singular points for curves of degree d follows. Linear system of plane curves are used in an algorithm to parametize irreducible curves which achieve the bound. Finally we prove Noether famous AF+BG Theorem, which for example implies a fast generalization of Pascal's famous theorem on hexagons.

In Chapter 6 projetive geometry is introduced more systematically. The image of an algebraic set under a projective morphism is closed. This explains once more that projective geometry is easier then affine geometry. The conceptual extra load which one has to master to work in projectibe space is more than compensated by the availablity of the Hilbert polynomial. For example we can generalize the notion of the degree for hypersurface to arbitrary algebraic sets, and can prove a Bezout formula for intersections with hypersurfaces. Bertini Theorem allows to interprete the degree geometrically, and gives a dynamical interpretation of the algebraic intersection multiplicities of plane curves. Gröbner basis are used once more to prove the Theorem on fiber dimensions.

Chapter 7 deals with rational maps. We discuss the blow up of a point and prove resolution of singularities for plane curves with the help of Cremona transformations. We then introduce divisors and linear systems on curves, and explain there importants to obtain the various projective embeddings of a curve. From this point of view the computation of complete linear system is the fundamental tool to obtain new embeddings. We give an algorithm for this and observe the invariance of the arithmetic genus of a curve in different embeddings.

In Chapter 8 we finally prove the famous Riemann-Roch formula. As an application we obtain the equality of arithmetic and geometric genus of smooth curves, which by Hurwitz' formula has a purely topological interpretation. Lüroth theorem settles the questions which curves can be rationally parametrized: Necessary and sufficient is that the genus $g = 0$. The set of all smooth projective curves of genus g carries the structure of an algebraic variety, the moduli space M_g, which by the way plays an important role in modern theoretical physics. To prove this is far beyond the scope of this book. However believing this fact we can understand Riemann's count $\dim M_g = 3g - 3$ for the moduli space of curves of genus $g \geq 2$.

We study the canonical map $C \to \mathbb{P}^{g-1}$ and explain how its syzygies carry conjectually information about the configuration of special linear systems of the curve. For example we prove Petri's theorem, which characterize trigonal curves by their syzygies. The final section of the book contains Stephano's proof of the Hasse-Weil formula for number of points on curves over finite fields, which is a very tricky application of the Riemann-Roch formula.

Content for the experts

Chapter 1 establishes the Algebra - Geometry dictionary. We formulate the weak version of the Nullstellensatz whose proof is postphoned to Chapter 3. We deduce from the weak the strong version of the Nullstellensatz, and the correspondence between points or irreducible algebraic sets and maximal respectively prime ideal in $\overline{\mathbb{k}}[x_1, \ldots, x_n]$.

Chapter 2 introduces Gröbner basis. The solution of the ideal menbership problem gives as special case an algorithm to decide solvabiltity We go on to give further elementary application of Gröbner basis, such as computaion of syzygies, of intersections of ideals, of elimination ideals and of kernels of ring homomorphism. The last theme has application to implizitization, i.e. to compute equations of parametrized varieties. The coordinate ring and polynomial maps are introduced. After an algorithmic proof of Hilbert Syzygy Theorem, which will be used in Chapter 6 to define the Hilbert polynomial, we sketch the algorithmic computation of Ext and Tor groups over polynomial rings, however without much focus on their use and theory.

Chapter 3 which with a geometric proof of the Nullstellensatz. Motivated by this proof we introduce inetgral extension and discuss the Going-up and Going-Down Theorems of Cohen-Seidenberg. The Chapter concludes with Krull dimension.

Chapter 4 is devoted to the local study of algebraic sets. We introduce the tangent space and local ring of an algebraic set at a point, and establish most of the basic theory, ranging from Nakayama's Lemma to intersection multiplicities of plane curves. Mora's algorithm is included to compute in these rings, in particular to compute the intersection multiplicities. The local global principle is used in explaining how the Jacobian criterion can be used to establish that an equidimensional ideal is radical. We introduce artinian rings and prove Krull's principal Theorem as an application. Finally we introduce the analytic type and the tangent cone of a singularity.

Chapter 5 is devoted to plane projective curves. We introduce resultants and prove Bézout's Theorem for plane curves. We introduce the geometric genus of plane curve with ordinary singularities, and present the algorithm using linear system of plane curves with assigned singularities to compute a $\Bbbk$-rational parametrization implicitely given genus 0 curve in the presence of one smooth $\Bbbk$-rational point. Chapter 5 culminates with the proof of Nother's AF+BG theorem, which is used in Noether's proof of the Riemann-Roch theorem later on. The Chapter finishes with outlook on the geometry and arithmetic of elliptic curves.

Chapter 6 treats projective algebraic sets in general. Gröbner basis are used to compute the projective closure. Gröbner basis and a counting argument also provide a tool to prove that the homogeneous ideal of Segre and Veronese varieties are generated by the well-known quadrics. In most text books it is only proved that these quadrice generate the ideal locally, i.e. the 'ideal sheaf', leaving the question about the homogeneous ideal open. After the main theorem on elimination we introduce the Hilbert function, Hilbert polynomial and Hilbert series using Hilbert's syzygy theorem. We prove Bézout's theorem for the intersection of varieties with hypersurfaces as an application. Bertini's theorem and the semi-continuity of fiber dimension is treated, based on a Gröbner basis argument. Without the explicitly introducing the concept of flattness, what we basically do is, to use Gröbner basis to establish flatness stratification. Finally we tough upon the general position principle and monodromy arguments. Monodromy arguments play an important role in the Numerical Algebraic Geometry implemented in the package 'Bertini', which is under developement by Sommese, Wampler and coloberators.

Chapter 7 treats rational maps and birational transformations. We prove resolution of plane curve singularities via Cremona transformation, and characterise rational curves by their geometric genus. We establih the notion of divisors and linear system for curves. In a way, which would generalizes to higher dimension without further work, we prove the completeness of large degree hypersurface systems. This is used in an algorithm to compute complete linear systems of divisors (on curves). Thus the students can compute complete system before they learn about the Riemann-Roch Theorem. The Chapter concludes with a proof of Riemann's inequality. and a discussion of the δ-invariant of curve singularities.

In Chapter 8 we finally prove the Riemann-Roch theorem for curves, and give some of its application. For example the Hurwitz formula, Lüroth theorem, the Plücker formulas and Weierstrass points are treated. Then we present Riemann's count for the dimension of moduli space M_g, and prove the formula modulo the fact, that the Picard group, the moduli space and the Hurwitz scheme carry the structure of an algebraic sets. The formula $\dim H_d = 4d$ for many components of the Hilbert scheme of degree d curves in $\mathbb{P}^3$ is established similarly. In Section 8.6 we treat canonical curves and prove Noether's theorem on the normal generation as well as Petri's theorem. After the proof of Clifford theorem we give an outlook on Green's conjecture, which connects syzygy of canonical curve with special linear series. The final section of this book contains Stephanov's proof of the Hasse-Weil formula for the number of points for curves over finite fields.

Usage as a text book

The book can be used as an text book in several ways. Working through it linearly give material for a course of at least two terms. Little more than basic field theory, such as algebraic extension, transzendence degree is assumed. We recall basic notion like rings, ideals and modules, when we need them. The book should be accessable to undergraduate math and computer science students.

For students with strong background knowledge in commutative algebra, the book adds only computational skills in this area, and draws the connection to geometry. In this case much of the material of Chapter 1-3 can be presented quickly.

Exercises are scattered through out the text. They provide examples and further facts, which illustrate the material, whose solution, however is not needed immediately. Those exercises used in proofs of Theorems further on, are indicted by an asterisque, those, which need additional knowledge, such as Galois theory and/or the theory of analytic functions are indicated by an †. Two †† indicates an in our eyes truly difficult exercise.

The minimal path, which leads to a proof of Riemann-Roch including algorithms to compute linear systems, consists of the following sections:

Sections	page range	pages
1.1-1.9	1-31	31
2.1-2.5, 2.8	47-77, 91-97	38
3.1, 3.3	105-111, 122-129	15
4.1-4.4	137-176	40
5.1-5.4	203-221	19
6.1-6.3	235-246	21
7.1-7.4	273-296	24
8.1-8.4	303-320	18

These are 206 out of 356 pages.

Acknowledgement

We thank William Fulton, Oliver Labs, ….

Prerequisites and basic conventions

We suppose some basic knowledge of elementary topology, groups, rings, fields, and vector spaces. In particular, the reader should be familiar with integral domains and unique factorization domains (UFD's for short), and with the notion of transcendence degree for field extensions. If $\Bbbk \subset \mathbb{K}$ is such an extension, we will write $\operatorname{trdeg}_{\mathbb{K}} \Bbbk$ for the transcendence degree of $\mathbb{K}$ over $\Bbbk$. For some of the exercises, a little of Galois theory is needed.

All rings are commutative with multiplicative identity 1. All homomomorphisms of rings take 1 to 1. In an integral domain or a field, $0 \neq 1$.

For any nonempty set X, we write id_X for the identity map of X.

Saarbrücken,
April 2005

Wolfram Decker
Frank-Olaf Schreyer

Contents

Part I Affine Varieties

Part I

Affine Varieties

Chapter 1

The Geometry-Algebra Dictionary

This chapter will provide a first impression of the linkage between geometry and algebra. Our geometric objects of study will be *affine* algebraic sets, which are subsets of affine space $\mathbb{A}^n(\Bbbk)$ defined by polynomial equations, and which are closely related to ideals in the polynomial ring $\Bbbk[x_1, \ldots, x_n]$. In fact, by Hilbert's basis theorem, every ideal $I \subset \Bbbk[x_1, \ldots, x_n]$ defines an algebraic subset $\mathrm{V}(I) \subset \mathbb{A}^n(\Bbbk)$ – the common vanishing locus of all elements of I. Conversely, given an algebraic set $A \subset \mathbb{A}^n(\Bbbk)$, we may associate to A the ideal $\mathrm{I}(A)$ of all polynomials vanishing on A.

The relationship between algebraic sets and ideals is made precise by Hilbert's Nullstellensatz which allows one to set up a dictionary between geometric and algebraic statements. In developing the dictionary, we will study a number of natural geometric operations on algebraic sets together with their algebraic counterparts. Moreover, we will see examples of how to express properties of algebraic sets in terms of ideals $I \subset \Bbbk[x_1, \ldots, x_n]$ or, in turn, of quotient rings $\Bbbk[x_1, \ldots, x_n]/I$. 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, we will see that each algebraic subset $A \subset \mathbb{A}^n(\Bbbk)$ comes equipped with a ring of functions, the ring of polynomial functions on A, which is naturally isomorphic to the quotient ring $\Bbbk[x_1, \ldots, x_n]/\mathrm{I}(A)$. We will use the polynomial functions to define the natural maps between algebraic sets, and to relate these maps to ring homomorphisms on the algebraic side.

In presenting explicit examples, we will occasionally use pieces 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 $x_1, \ldots, x_n$ is a collection of variables, $R[x_1, \ldots, x_n]$ denotes the set of polynomials in n variables $x_1, \ldots, x_n$ with coefficients in R. To write the elements of $R[x_1, \ldots, x_n]$, we use multiindices. To begin with, a **monomial** in $R[x_1, \ldots, x_n]$ is a product $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, where $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$. A **term** in $R[x_1, \ldots, x_n]$ is an element of R times a monomial. Each polynomial $0 \neq f \in R[x_1, \ldots, x_n]$ can be uniquely expressed as the sum of finitely many nonzero terms involving distinct monomials. These terms (monomials) are called the **terms** (**monomials**) **of** $\boldsymbol{f}$.

With the usual algebraic operations, the set $R[x_1, \ldots, x_n]$ 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 $s_1, \ldots, s_n \in S$, there exists a unique homomorphism $\Phi : R[x_1, \ldots, x_n] \to S$ extending ϕ, and such that $\Phi(x_i) = s_i$ for all i. In fact, Φ is the map $f \mapsto f(s_1, \ldots, s_n)$, where the value $f(s_1, \ldots, s_n)$ 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 $R[s_1, \ldots, s_n]$ for its image in S.

The **degree of** of a monomial $\boldsymbol{x^\alpha} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ 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 $\deg 0 = -\infty$.

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

We usually write

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

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

Every nonzero polynomial $f \in R[x_1, \ldots, x_n]$ can be uniquely written as a sum $f = f_0 + f_1 + f_2 + \cdots + f_{\deg(f)}$, where the f_i are homogenenous of degree i. Given an extra variable x_0, the polynomial

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

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 $F \in R[x_0, x_1, \ldots, x_n]$ with respect to x_0 is defined to be the polynomial $F(1, x_1, \ldots, x_n) \in R[x_1, \ldots, x_n]$. 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 $\Bbbk[x]$ in one variable x over a field $\Bbbk$ 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, $\Bbbk(x_1, \ldots, x_n) = \mathrm{Q}(\Bbbk[x_1, \ldots, x_n])$ will denote the **field of rational functions** in $x_1, \ldots, x_n$ with coefficients in $\Bbbk$..

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 $f \in R[x_1, \ldots, x_n]$ 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 $\Bbbk$ is a field of characteristic $p > 0$, and $f \in \Bbbk[x_1, \ldots, x_n]$, then $\frac{\partial f}{\partial x_i} = 0$ iff $f \in \Bbbk[x_1, \ldots, x_i^p, \ldots, x_n]$. Conclude that all the $\frac{\partial f}{\partial x_i}$ are zero iff $f \in \Bbbk[x_1^p, \ldots, x_n^p]$. □

Formally similar to the construction of polynomial rings is the construction of formal power series rings:

Remark-Definition 1.1.4. Let $\Bbbk$ be a field. A **formal power series** in the variables $x_1, \ldots, x_n$ with coefficients in $\Bbbk$ is an expression of type

$$\sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha, \quad \text{with all} \quad a_\alpha \in \Bbbk.$$

These expressions form a ring, where the algebraic operations are defined as follows:

$$\sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha + \sum_{\alpha \in \mathbb{N}^n} b_\alpha x^\alpha = \sum_{\alpha \in \mathbb{N}^n} (a_\alpha + b_\alpha) x^\alpha \quad \text{and}$$

$$\sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha \cdot \sum_{\alpha \in \mathbb{N}^n} b_\alpha x^\alpha = \sum_{\gamma \in \mathbb{N}^n} \big(\sum_{\alpha+\beta=\gamma} a_\alpha b_\beta \big) x^\alpha.$$

This ring, denoted $\Bbbk[[x_1, \ldots, x_n]]$, is called the **ring of formal power series** in n variables $x_1, \ldots, x_n$ with coefficients $\Bbbk$. Note that $\Bbbk[x_1, \ldots, x_n]$ is naturally contained in $\Bbbk[[x_1, \ldots, x_n]]$ as a subring. The **multiplicity** of a formal power series $f = \sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha \in \Bbbk[[x_1, \ldots, x_n]]$, written $\operatorname{mult}(f)$, is defined to be

$$\operatorname{mult}(f) = \min\{m \mid \sum_{|\alpha|=m} a_\alpha x^\alpha \neq 0\}.$$

□

1.2 Algebraic Sets

Let $\Bbbk$ be any field. The **affine n-space** over $\Bbbk$ is the set

$$\mathbb{A}^n(\Bbbk) := \big\{(a_1, \ldots, a_n) \mid a_1, \ldots, a_n \in \Bbbk\big\}.$$

An element $p \in \mathbb{A}^n(\Bbbk)$ is called a **point**, and if $p = (a_1, \ldots, a_n)$, the a_i are called the **coordinates** of p. We call $\mathbb{A}^1(\Bbbk)$ and $\mathbb{A}^2(\Bbbk)$ the **affine line** and the **affine plane** over $\Bbbk$, respectively.

If $\Bbbk[x_1, \ldots, x_n]$ is the ring of polynomials in n variables with coefficients in $\Bbbk$, then each element $f \in \Bbbk[x_1, \ldots, x_n]$ defines a function

$$f : \mathbb{A}^n(\Bbbk) \to \Bbbk, \ (a_1, \ldots, a_n) \mapsto f(a_1, \ldots, a_n).$$

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** $x_i : \mathbb{A}^n(\Bbbk) \to \Bbbk, (a_1, \ldots, a_n) \mapsto a_i$.

Considering a polynomial $f \in \Bbbk[x_1, \ldots, x_n]$ 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. If $\mathbb{F}_2$ is the field with two elements, find a nonzero polynomial in $\mathbb{F}_2[x_1, \dots, x_n]$ involving all of $x_1, \dots, x_n$ and vanishing at every point of $\mathbb{A}^n(\mathbb{F}_2)$. □

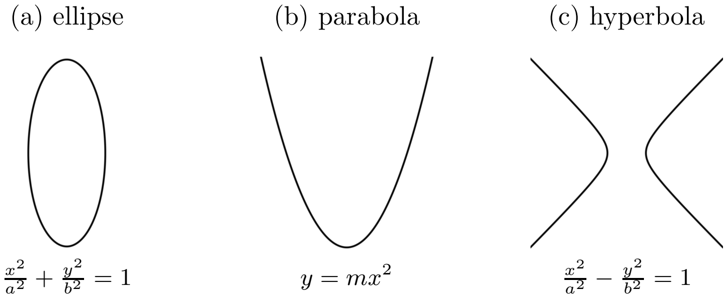
Definition 1.2.3. A subset $A \subset \mathbb{A}^n(\Bbbk)$ is called a **hypersurface** in $\mathbb{A}^n(\Bbbk)$ if $A = \mathrm{V}(f)$ for some nonconstant polynomial $f \in \Bbbk[x_1, \dots, x_n]$. In this case, we say that $f(x_1, \dots, x_n) = 0$ is a **defining equation** for A. □

A hypersurface in $\mathbb{A}^2(\Bbbk)$ is called an **affine plane curve**. We present some explicit examples, choosing $\Bbbk = \mathbb{R}$ 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:

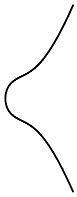


(a) ellipse (b) parabola (c) hyperbola

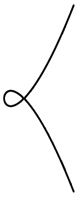
$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ $y = mx^2$ $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$

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

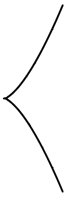
2. A **cubic curve** in $\mathbb{A}^2(\mathbb{R})$ is defined by a degree-3 equation. Such curves were systematically investigated by Newton (1666). Here are some particular examples:



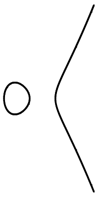
$y^2 = x^3 + x + 1$



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

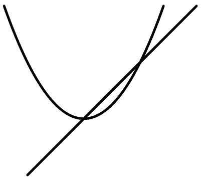


$y^2 = x^3 + x^2$



$y^2 = x^3 - x$

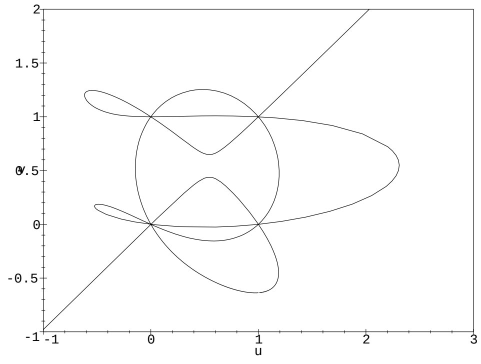
The cubic curve with equation $y^2 = xy + x^2y - x^3$ is the union of a parabola and a line:



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 $f(x(t), y(t)) = 0 \in \mathbb{R}(t)$. This shows that 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. □

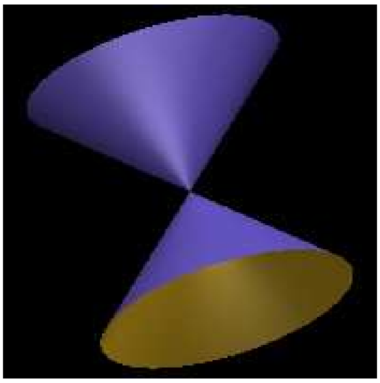
Hypersurfaces in affine 3-space provide our first examples of surfaces:

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

1. The surface

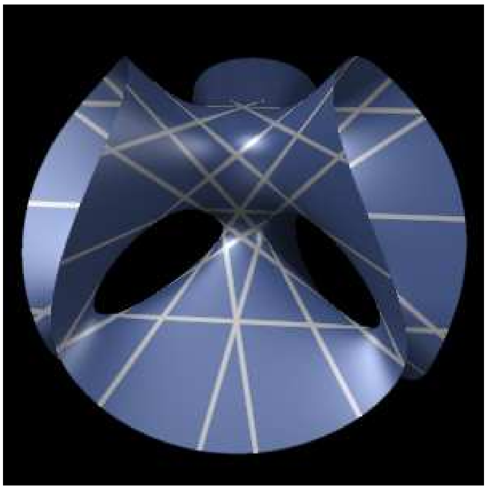
$$\mathrm{V}(x^2 + y^2 - z^2) \subset \mathbb{A}^3(\mathbb{R})$$

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):



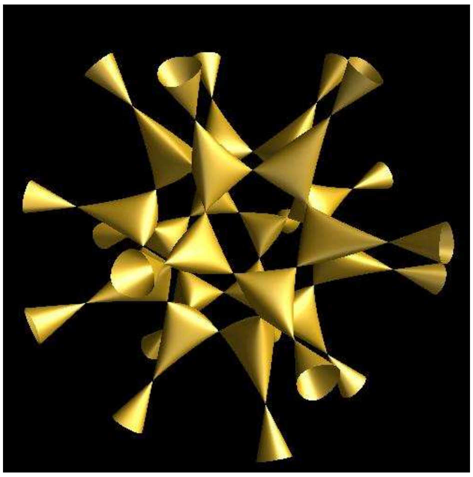
It is defined by the equation

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

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

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

□

In general, we are concerned with more than one polynomial equation. If $f_1,\dots,f_r \in \mathbb{k}[x_1,\dots,x_n]$, we write

$$\mathrm{V}(f_1,\dots,f_r)=\{p\in\mathbb{A}^n(\mathbb{k}) \mid f_1(p)=0,\dots,f_r(p)=0\}.$$

Definition 1.2.8. A subset $A\subset\mathbb{A}^n(\mathbb{k})$ is called **affine algebraic**, or simply **algebraic**, if $A=\mathrm{V}(f_1,\dots,f_r)$ for some polynomials $f_1,\dots,f_r\in\mathbb{k}[x_1,\dots,x_n]$. In this case, we say that

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

is a system of **defining equations** for A. □

Since $\mathrm{V}(f_1,\dots,f_r)=\bigcap_{i=1}^r \mathrm{V}(f_i)$, a subset of $\mathbb{A}^n(\mathbb{k})$ is algebraic iff it is the intersection of finitely many hypersurfaces.

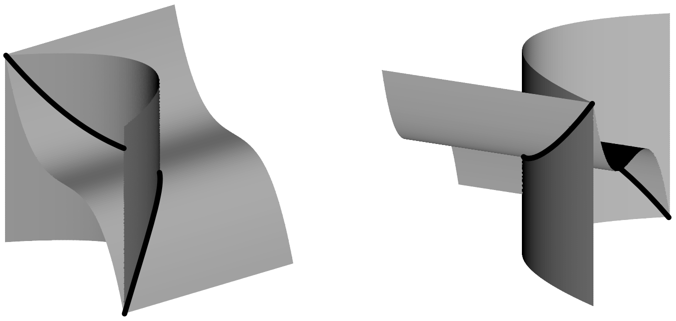
Example 1.2.9. If

$$f=a_1x_1+\dots+a_nx_n-b\in\mathbb{k}[x_1,\dots,x_n].$$

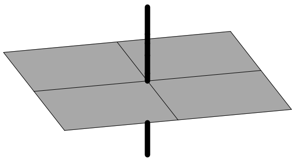
is a degree-1 polynomial, then $\mathrm{V}(f)\subset\mathbb{A}^n(\mathbb{k})$ is called a **hyperplane**. The intersection of finitely many hyperplanes is, then, the set of solutions of a system of linear equations as studied in linear algebra. We will refer to such a set as a **linear subvariety** of $\mathbb{A}^n(\mathbb{k})$. □

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

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**:



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:



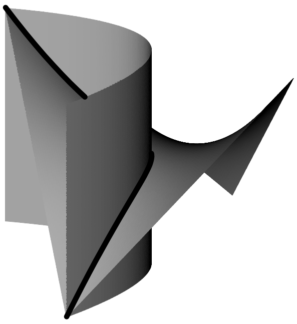
□

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

As we already 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, given an arbitrary system of polynomial equations

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

we consider linear combinations of the f_i 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 $\mathrm{V}(z-xy)$:



If A is the algebraic set defined by the vanishing of $f_1,\dots,f_r\in\Bbbk[x_1,\dots,x_n]$, then all $\Bbbk[x_1,\dots,x_n]$-linear combinations $g_1f_1+\dots+g_rf_r$ vanish on A, too. Thus, we may as well say that A is defined by the set of all these linear combinations.

Definition 1.2.12. If $I\subset\Bbbk[x_1,\dots,x_n]$ is any subset, its **locus of zeros** (or **vanishing locus**) in $\mathbb{A}^n(\Bbbk)$ is the set of common zeros of all elements of I, namely

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

Remark 1.2.13. If $\Bbbk\subset\mathbb{K}$ is a field extension such as $\mathbb{R}\subset\mathbb{C}$, then every subset $I\subset\Bbbk[x_1,\dots,x_n]$ is also a subset of $\mathbb{K}[x_1,\dots,x_n]$. We may, thus, speak of the locus of zeroes of I in $\mathbb{A}^n(\mathbb{K})$. □

Sets of $\Bbbk[x_1,\dots,x_n]$-linear combinations of polynomials as considered in the discussion above carry an algebraic structure whose definition we recall next.

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

If $I \subset R$ is an ideal, any subset X of I satisfying $I = \langle X\rangle$ 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 $I_1, \dots, I_s$ 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. □

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. □

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

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

Show:

1. If the I_k are pairwise coprime, then $I_1 \cdots I_s = \bigcap_{k=1}^s I_k$.
2. The map ϕ is surjective iff the I_k are pairwise coprime.
3. The map ϕ is injective iff $\bigcap_{k=1}^s I_k = \langle 0 \rangle$. □

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

To express geometric statements in algebraic terms, we will represent algebraic sets by ideals rather than by specific systems of defining equations. This fits well with the fact that every ideal of $\Bbbk[x_1, \ldots, x_n]$ defines an algebraic set:

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

Corollary 1.4.2. *If* $X \subset \Bbbk[x_1, \ldots, x_n]$ *is any subset, its locus of zeros in* $\mathbb{A}^n(\Bbbk)$ *is algebraic.*

Proof (of the corollary). Apply the basis theorem to the ideal generated by X in $\Bbbk[x_1, \ldots, x_n]$. □

All known proofs of the basis theorem itself proceed by induction on the number of variables, starting with the univariate case.

Remark 1.4.3. 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 landmark 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 2.8.9, 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 Hilbert's 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 1.4.4 below). From this condition, she derived the existence of primary decompositions (we will treat this in Section 1.8).

Exercise* 1.4.4. Prove 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 1.4.5. 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* 1.4.6 (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
$$\langle a_1 \rangle \subsetneq \langle a_1, a_2 \rangle \subsetneq \dots \subsetneq R.$$
□

Hilbert's basis theorem establishes a map V sending a subset I of $\Bbbk[x_1, \dots, x_n]$ to the algebraic subset $\mathrm{V}(I)$ of $\mathbb{A}^n(\Bbbk)$. We summarize some properties of V:

Proposition 1.4.7. *Let* $R = \Bbbk[x_1, \dots, x_n]$. *Then:*

1. $\mathrm{V}(0) = \mathbb{A}^n(\Bbbk)$. $\quad \mathrm{V}(1) = \emptyset$.
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*
$$\mathrm{V}(I) \cup V(J) = \mathrm{V}(I \cdot J) = \mathrm{V}(I \cap J).$$
In particular, the union of finitely many algebraic subsets of $\mathbb{A}^n(\Bbbk)$ *is algebraic.*

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

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

In particular, the intersection of any family of algebraic subsets of $\mathbb{A}^n(\Bbbk)$ is algebraic.

5. *If $a_1, \dots, a_n \in \Bbbk$, then*

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

Exercise* 1.4.8. Prove Proposition 1.4.7. □

Remark-Definition 1.4.9. By properties 1, 3, and 4 above, the algebraic subsets of $\mathbb{A}^n(\Bbbk)$ satisfy the axioms for the closed sets of a topology on $\mathbb{A}^n(\Bbbk)$, called the **Zariski topology** on $\mathbb{A}^n(\Bbbk)$. The open sets in this topology are of type

$$\begin{aligned}\mathbb{A}^n(\Bbbk) \setminus \mathrm{V}(f_1, \dots, f_r) &= \mathbb{A}^n(\Bbbk) \setminus (\mathrm{V}(f_1) \cap \dots \cap \mathrm{V}(f_r)) \\ &= \bigcup_{i=1}^r (\mathbb{A}^n(\Bbbk) \setminus \mathrm{V}(f_i)),\end{aligned}$$

where $f_1, \dots, f_r \in \Bbbk[x_1, \dots, x_n]$. In particular, the **distinguished open sets**

$$\mathrm{D}(f) := \mathbb{A}^n(\Bbbk) \setminus \mathrm{V}(f),\ f \in \Bbbk[x_1, \dots, x_n],$$

form a basis for the Zariski topology on $\mathbb{A}^n(\Bbbk)$. □

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, and 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 1.4.10. 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. □

1.5 Vanishing Ideals

On our way to explore the relationship between algebraic subsets of $\mathbb{A}^n(\Bbbk)$ and ideals of $\Bbbk[x_1, \dots, x_n]$, we so far have the surjective map

$$\{\text{algebraic subsets of } \mathbb{A}^n(\Bbbk)\} \xleftarrow[\mathrm{V}]{} \{\text{ideals of } \Bbbk[x_1, \dots, x_n]\}.$$

Now, we define a map I in the other direction.

Remark-Definition 1.5.1. If $A \subset \mathbb{A}^n(\Bbbk)$ is any subset, the set

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

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

Exercise 1.5.2. 1. Show that every polynomial $f \in \Bbbk[x, y, z]$ has a representation of type

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

where $g_1, g_2 \in \Bbbk[x, y, z]$ and $h \in \Bbbk[x]$.

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

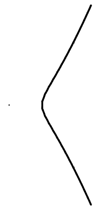
$$C = \{(a, a^2, a^3) \mid a \in \Bbbk\}.$$ □

The expression for f in terms of $y - x^2$ and $z - x^3$ in the exercise above is reminiscent of Euclidean division with remainder, except that we are dividing by two polynomials instead of one. In Exercise 2.2.15, we will recompute the expression in a more systematic way, using a generalized division algorithm.

Exercise 1.5.3. Let $\Bbbk = \mathbb{R}$, 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 the origin. □

Remark 1.5.4. 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 $\mathrm{I}(A)$. A method which in many cases of interest allows one to decide whether a given set of polynomials defining an algebraic set 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). See also Remark 2.4.12 (in conjunction with Hilbert's Nullstellensatz 1.6.2 and Section 2.7 on the role of the ground field). □

The following proposition summarizes some properties of I, and starts examining how V and I are related:

Proposition 1.5.5. *Let $R = \Bbbk[x_1, \dots, x_n]$. Then:*

1. $\mathrm{I}(\emptyset) = R$. *If* $\Bbbk$ *is infinite, then* $\mathrm{I}(\mathbb{A}^n(\Bbbk)) = \langle 0 \rangle$.
2. *If* $A \subset B$ *are subsets of* $\mathbb{A}^n(\Bbbk)$, *then* $\mathrm{I}(A) \supset \mathrm{I}(B)$.
3. *If* A, B *are subsets of* $\mathbb{A}^n(\Bbbk)$, *then*
$$\mathrm{I}(A \cup B) = \mathrm{I}(A) \cap \mathrm{I}(B).$$
4. *If* $(a_1, \dots, a_n) \in \mathbb{A}^n(\Bbbk)$ *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* $A \subset \mathbb{A}^n(\Bbbk)$, *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* $\mathbb{A}^n(\Bbbk)$.

Exercise* 1.5.6. Prove Proposition 1.5.5. □

Not every ideal $I \subset \Bbbk[x_1, \dots, x_n]$ 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:

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

In general, if a power f^m of a polynomial f vanishes on a subset $A \subset \mathbb{A}^n(\Bbbk)$, 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 1.5.7. 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 1.5.8. 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

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

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

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

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

Exercise* 1.5.9. 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, then R/I is Noetherian as well. □

Exercise* 1.5.10. Let I, J be ideals of a ring R. Show:

1. $$\text{rad}\,(IJ) = \text{rad}\,(I \cap J) = \text{rad}\,I \cap \text{rad}\,J.$$
2. $$\text{rad}\,(I + J) = \text{rad}\,(\text{rad}\,I + \text{rad}\,J).$$
3. $$\text{rad}\,I = \langle 1 \rangle \iff I = \langle 1 \rangle.$$
4. If $\text{rad}\,I, \text{rad}\,J$ 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:

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

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 maps I and V are well behaved over an algebraically closed field – provided we restrict our attention to radical ideals:

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 $f_k(a_1, \dots, a_n) \neq 0$ 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 $\Bbbk = \overline{\Bbbk}$ is algebraically closed, then* I *and* V *define a one-to-one correspondence*

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

□

The weak version of the Nullstellensatz adresses the basic **problem of solvability**: Given $f_1, \dots, f_r \in \Bbbk[x_1, \dots, x_n]$, 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 $\overline{\Bbbk}$ iff $1 \in \langle f_1, \dots, f_r \rangle$. The trick of of Rabinowitch allows us to discuss a related problem:

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

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

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 $I \subset \Bbbk[x_1, \ldots, x_n]$ or, in turn, statements on quotient rings $\Bbbk[x_1, \ldots, x_n]/I$. Here is, for instance, a result which extends the weak version of the Nullstellensatz in that it characterizes systems of polynomial equations with at most finally many solutions:

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

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

How to decide algorithmically whether an ideal $I \subset \Bbbk[x_1, \ldots, x_n]$ 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 $I \subset \mathbb{R}[x_1, \ldots, x_n]$ whose locus of zeros in $\mathbb{A}^n(\mathbb{R})$ is finite though $\dim_{\mathbb{R}} \mathbb{R}[x_1, \ldots, x_n]/I = \infty$. □

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. An 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 properly contained in A. Otherwise, A is called **irreducible**, or a **subvariety** of $\mathbb{A}^n(\Bbbk)$, or simply an **affine variety**. The empty set is not considered to be irreducible. □

Proposition 1.7.2. *Let $A \subset \mathbb{A}^n(\Bbbk)$ be an algebraic set. Then the following 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*

$$\{\textit{subvarieties of } \mathbb{A}^n(\Bbbk)\} \underset{\mathrm{V}}{\overset{\mathrm{I}}{\rightleftarrows}} \{\textit{prime ideals of } \Bbbk[x_1,\dots,x_n]\}.$$ □

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, we will show in Exercise 2.9.1 that

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

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

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 $p = (a_1, \dots, a_n) \in \mathbb{A}^n(\Bbbk)$ is a point, every polynomial $f \in \Bbbk[x_1, \dots, x_n]$ 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 $\boldsymbol{f}$ at $\boldsymbol{p}$** 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

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

is the kernel of the evaluation map

$$\Bbbk[x_1, \dots, x_n] \to \Bbbk, \ f \mapsto f(p),$$

so that $\Bbbk[x_1, \dots, x_n]/\mathrm{I}(p) \cong \Bbbk$ by the homomorphy theorem (see Theorem 1.10.5 for a general version of the homomorphy theorem). In particular, $\Bbbk[x_1, \dots, x_n]/\mathrm{I}(p)$ 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 $\Bbbk[x_1, \dots, x_n]$ is of the form $\langle x_1 - a_1, \dots, x_n - a_n \rangle$ for some $a_1, \dots, a_n \in \Bbbk$.*

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

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

□

Exercise* 1.7.10. If $\Bbbk$ is not necessarily algebraically closed, the maximal ideals of $\Bbbk[x_1, \dots, x_n]$ can be described as follows. 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*

$$A = V_1 \cup \cdots \cup V_s$$

*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

$$A = V_1 \cup \cdots \cup V_s = V_1' \cup \cdots \cup V_t'$$

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 $I \subset \Bbbk[x_1, \dots, x_n]$ has a unique **minimal prime decomposition**. That is, rad I can be uniquely written as the intersection of finitely many prime ideals:

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

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 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

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

The decomposition is called **minimal** if the radicals 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 be an irreducible ideal of R. 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 be an arbitrary ideal of R. 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 1.9.3. □

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, $\mathfrak{p}_1 \ldots, \mathfrak{p}_s$ are the mimimal associated primes of I, then

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

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 geometry: if $\Bbbk = \overline{\Bbbk}$ is algebraically closed, the minimal associated primes of an ideal $I \subset \Bbbk[x_1, \ldots, x_n]$ 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

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

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. It is, thus, not an algebraic set (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 $\Bbbk[x_1, \dots, x_n]$. 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 $\mathfrak{q}_i : J^\infty = \Bbbk[x_1, \dots, x_n]$ 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 Lemma 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 Lemma 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,

$$I : J^\infty = \bigcap_{i=1}^{t} (\mathfrak{q}_i : J^\infty) = \left(\bigcap_{\substack{J^m \subset \mathfrak{q}_i \\ \text{for some } m \geq 1}} (\mathfrak{q}_i : J^\infty) \right) \cap \left(\bigcap_{\substack{J^m \not\subset \mathfrak{q}_i \\ \text{for all } m \geq 1}} (\mathfrak{q}_i : J^\infty) \right)$$

$$= \bigcap_{\substack{J^m \not\subset \mathfrak{q}_i \\ \text{for all } m \geq 1}} \mathfrak{q}_i = \bigcap_{\mathrm{V}(J) \not\supset \mathrm{V}(q_i)} \mathfrak{q}_i .$$

Hence, by Step 3,

$$\mathrm{V}(I : J^\infty) = \bigcup_{\mathrm{V}(J) \not\supset \mathrm{V}(\mathfrak{q}_i)} \mathrm{V}(\mathfrak{q}_i) = \bigcup_{\mathrm{V}(J) \not\supset \mathrm{V}(\mathfrak{q}_i)} \overline{\mathrm{V}(\mathfrak{q}_i) \setminus \mathrm{V}(J)}$$

$$= \bigcup_{i=1}^{t} \overline{(\mathrm{V}(\mathfrak{q}_i) \setminus \mathrm{V}(J))} = \overline{\mathrm{V}(I) \setminus \mathrm{V}(J)},$$

as required.

For the second statement, suppose that I is a radical ideal. In this case, we may write I as the intersection of *prime* ideals $\mathfrak{q}_i$. The same arguments as above show, then, that

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

□

Lemma 1.9.2. *If I is an ideal of a Noetherian ring R, the following hold:*

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 $\operatorname{rad} I = \langle f_1, \dots, f_r \rangle$. 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. Prove Theorem 1.8.7.
Hint. As a first step, show that the radicals $\mathfrak{p}_i = \operatorname{rad} \mathfrak{q}_i$ are precisely the prime ideals occuring in the set of ideals $(\operatorname{rad} I) : f$, $f \in R$. □

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. In treating some of the basic operations on modules, we will, in particular, discuss 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 **??** 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:

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

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 following definition extends the notion of a linear map 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

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

ia again an R-module: if $r \in R$, and $\phi, \psi \in \mathrm{Hom}_R(M, N)$, 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 a subgroup I of M such that if $r \in R$ and $m \in I$, then $rm \in I$. In this case, 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 $\boldsymbol{X}$**. 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 $M = \langle m_1, \ldots, m_k \rangle$ for some $m_1, \ldots, m_k \in M$.

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 $M_1 \oplus \cdots \oplus M_k$ of any finite set of R-modules $M_1, \ldots, M_k$. 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, a set of generators which is 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. 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 1.4.4 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

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

□

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:*

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

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

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

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:

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

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** $M_1 \otimes \cdots \otimes M_k$ 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

$$m \otimes m' \mapsto \phi(m) \otimes \phi'(m'). \qquad \square$$

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 1.5.9, 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.

1.11 Coordinate Rings and Morphisms

In this section, we will return to the geometry-algebra dictionary. We will take up a theme which is already familiar to students of introductory courses in mathematics: In order to understand a given class of mathematical objects, it is usually necessary to study the natural maps between these objects. In linear algebra, for instance, we study linear maps between vector spaces, and in topology, we study continuous maps between topological spaces. In algebraic geometry, just as affine algebraic sets are given by polynomials, the natural maps between them are also given by polynomials. Before presenting the general definition, we treat the special case of polynomial functions:

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 $f, g \in \Bbbk[A]$, then

$$(f+g)(p) = f(p) + g(p) \text{ and } (f \cdot g)(p) = f(p) \cdot g(p) \text{ for all } p \in A.$$

We may, then, regard $\Bbbk$ as the subring of all constant functions and, thus, $\Bbbk[A]$ as a $\Bbbk$-algebra.

Two polynomials in $\Bbbk[x_1, \dots, x_n]$ define the same polynomial function on A iff their difference vanishes on A. We, hence, have a natural isomorphism

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

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. Note that the residue classes of the x_i (that is, the coordinate functions on A) generate $\Bbbk[A]$ as a $\Bbbk$-algebra.

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. □

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 $\Bbbk[x_1, \dots, x_n]$ 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 $\boldsymbol{x} = x_1, \dots, x_n$ and $\boldsymbol{y} = y_1, \dots, y_m$.

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
$$\mathrm{I}(X \times Y) = \mathrm{I}(X)\,\Bbbk[\boldsymbol{x}, \boldsymbol{y}] + \mathrm{I}(Y)\,\Bbbk[\boldsymbol{x}, \boldsymbol{y}] \subset \Bbbk[\boldsymbol{x}, \boldsymbol{y}].$$
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 $\Bbbk[x_1, \dots, x_n]$ by an arbitrary algebraic subset $A \subset \mathbb{A}^n(\Bbbk)$ and its coordinate ring $\Bbbk[A]$, respectively. 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 form the closed sets of a topology on A, called 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. Note that the Zariski topology on A is induced by the Zariski topology on $\mathbb{A}^n(\Bbbk)$.

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 $\Bbbk[A] = \Bbbk[x_1, \ldots, x_n]/\mathrm{I}(A)$ to ideals in $\Bbbk[x_1, \ldots, x_n]$ (see Exercise 1.5.9).
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 a variety consisting of more than one point is not Hausdorff:

Proposition 1.11.8. *The following conditions on an algebraic set $A \subset \mathbb{A}^n(\Bbbk)$ 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. □

Here is the definition of the natural maps between affine algebraic sets:

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 $f_1, \ldots, f_m \in \Bbbk[x_1, \ldots, x_n]$ such that $\varphi(p) = (f_1(p), \ldots, f_m(p))$ for all $p \in A$. □

Theorem 1.11.12. *Let $A \subset \mathbb{A}^n(\Bbbk)$ and $B \subset \mathbb{A}^m(\Bbbk)$ be algebraic sets.*

1. *If $\varphi : A \to B$ is a polynomial map, then for each polynomial function g on B, the composition $g \circ \varphi$ is a polynomial function on A. The induced map*
$$\varphi^* : \Bbbk[B] \to \Bbbk[A],\ g \mapsto g \circ \varphi,$$
is a $\Bbbk$-algebra homomorphism.
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,
$$\mathrm{id}_A^* = \mathrm{id}_{\Bbbk[A]}.$$

Proof. Let $x_1, \ldots, x_n$ be the coordinate functions on $\mathbb{A}^n(\Bbbk)$, and let $y_1, \ldots, y_m$ be the coordinate functions on $\mathbb{A}^m(\Bbbk)$.

1. Let $\varphi : A \to B$ be given by polynomials $f_1, \ldots, f_m \in \Bbbk[x_1 \ldots, x_n]$. If $q \mapsto g(q)$ is a polynomial function on B, represented by a polynomial $g(y_1, \ldots, y_m) \in \Bbbk[y_1 \ldots, y_m]$, then $p \mapsto g(\varphi(p))$ is a polynomial function on A, represented by the polynomial $g(f_1, \ldots, f_m) \in \Bbbk[x_1 \ldots, x_n]$. We leave it to the reader to show that the induced map φ^* is a $\Bbbk$-algebra homomorphism.

2. For $j = 1, \ldots, m$, choose a polynomial $f_j \in \Bbbk[x_1, \ldots, x_n]$ representing $\phi(y_j)$. Then $f_1, \ldots, f_m$ define a polynomial map $\varphi : A \to \mathbb{A}^m(\Bbbk)$. We leave it to the reader to show that φ maps A to B, that $\phi = \varphi^*$, and that φ is uniquely determined by ϕ.

3. This is obvious. □

Exercise* 1.11.13. Complete the proof of Theorem 1.11.12. □

Remark 1.11.14. According to Definition 1.11.11, 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 1.11.12 and its proof, 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. If $\phi : \Bbbk[y_1,\dots,y_m] \to \Bbbk[x_1,\dots,x_n]$ is any $\Bbbk$-algebra homomorphism, 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 1.11.19 for an example. □

Remark 1.11.15. Every morphism $A \to B$ of affine algebraic sets is continuous with respect to the Zariski topology. 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 1.11.16. 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 1.11.12 implies:

Corollary 1.11.17. *A morphism $\varphi : A \to B$ of affine algebraic sets is an isomorphism of affine algebraic sets 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.* □

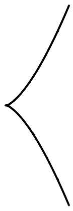
Exercise 1.11.18. Let $\Bbbk$ be infinite. Show:

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 1.11.19. 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* 1.11.20. 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 possibly different affine spaces. On the algebraic side, the vanishing ideal depends on the embedding, but the coordinate ring does not. The coordinate ring is an invariant of an algebraic set up to isomorphism.

Definition 1.11.21. 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 1.11.22. *A morphism $F = (f_1, \ldots, f_n) : \mathbb{A}^n(\Bbbk) \to \mathbb{A}^n(\Bbbk)$ is an automorphism of $\mathbb{A}^n(\Bbbk)$ iff*
$$\Bbbk[x_1, \ldots, x_n] = \Bbbk[f_1, \ldots, f_n].$$

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 of $\mathbb{A}^n(\Bbbk)$ under F can, then, be thought of as the the original set A viewed using the new coordinates.

Definition 1.11.23. An automorphism of $\mathbb{A}^n(\Bbbk)$ is called a **change of coordinates** of $\mathbb{A}^n(\Bbbk)$. □

Example 1.11.24. 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 1.11.25. 1. By results of Jung (1942) and van der Kulk (1953), who treat the cases $\mathrm{char}\,\Bbbk = 0$ and $\mathrm{char}\,\Bbbk > 0$, 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 $\mathrm{char}\,\Bbbk = 0$, 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 1.11.26. 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 $\operatorname{char} \mathbb{k} > 0$. □

Remark 1.11.27. 1. In the language of categories (see, for instance, Mac Lane (1990) for categories), Theorem 1.11.12 can be rephrased as follows. Over an algebraically closed field $\mathbb{k}$, the functor $A \to \mathbb{k}[A]$ induces an arrow-reversing equivalence between the category of affine algebraic sets over $\mathbb{k}$ and the category of reduced affine $\mathbb{k}$-algebras. The subcategory of affine varieties over $\mathbb{k}$ corresponds to that of affine domains over $\mathbb{k}$.

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

Exercise 1.12.2. If $\mathbb{k}$ is not algebraically closed, show that every algebraic subset of $\mathbb{A}^n(\mathbb{k})$ can be defined by a single polynomial equation (see Exercise 1.6.6 for the case $\mathbb{k} = \mathbb{R}$).
Hint. Consider the case of the origin in $\mathbb{A}^2(\mathbb{k})$ 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 $\operatorname{char} \mathbb{k} = p > 0$, show that the map

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

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). For this, 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.

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

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

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

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

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

and call it a syzygy on $f_1, \ldots, f_r$. 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 $f_1, \ldots, f_r$ form a Gröbner basis, the special syzygies obtained in Buchberger's test form a Gröbner basis for the module of all the syzygies on $f_1, \ldots, f_r$. 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 **??** 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 ideals generated by monomials and, thus, to questions which are usually much easier.

2.1 Monomials and Monomial Ideals

According to our conventions, we write monomials using multiindices: a **monomial** in $\mathbb{k}[x_1, \ldots, x_n]$ is a product

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n},$$

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

A number of operations on polynomials are simpler for monomials than for arbitrary polynomials. For instance, if $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$ is another multiindex, the **least common multiple** of x^α and x^β is

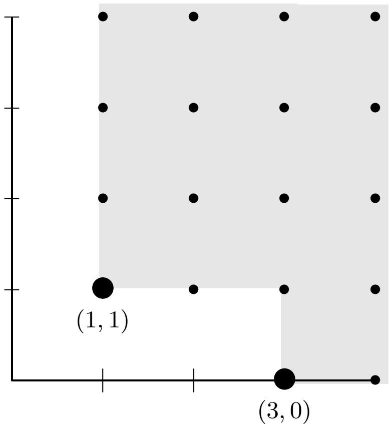
$$\mathrm{LCM}(x^\alpha, x^\beta) = x_1^{\max(\alpha_1,\beta_1)} \cdots x_n^{\max(\alpha_n,\beta_n)},$$

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 manipulate than arbitrary ideals. Consider, for instance, the ideal membership problem: If $I \subset \mathbb{k}[x_1, \ldots, x_n]$ is a monomial ideal, given by monomial generators $m_1, \ldots, m_r$, a term is contained in I iff it is divisible by at least one of the m_i; an arbitrary polynomial $g \in \mathbb{k}[x_1, \ldots, x_n]$ 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 $\mathbb{k}[x, y]$ 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. □

If $I\subset\Bbbk[x_1,\dots,x_n]$ is a monomial ideal, the minimal elements in the partial order defined by divisibility on the set of monomials in I are uniquely determined. They can be obtained from any set of monomial generators by removing those generators which are divisible by others. We will refer to them 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, we have a uniquely determined 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*

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

Indeed, **Euclid's division algorithm** finds the remainder h and the quotient g_1 by successively cancelling leading terms using f. 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. 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 algorithm for computing GCD's in $\Bbbk[x]$. □

Euclidean division with remainder allows us to decide ideal membership in $\Bbbk[x]$ as follows. If nonzero polynomials $g, f_1, \dots, f_r \in \Bbbk[x]$ are given, use Euclid's algorithm to compute $f = \mathrm{GCD}(f_1, \dots, f_r)$. Then $\langle f_1, \dots, f_r\rangle = \langle f\rangle$, so that $g \in \langle f_1, \dots, f_r\rangle$ 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 $\Bbbk[x_1, \dots, x_n]$ is generated by just one element, we ask for an algorithm which divides by several polynomials in $\Bbbk[x_1, \dots, x_n]$ 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 $\Bbbk[x_1, \dots, x_n]$ 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 $\Bbbk[x_1, \dots, x_n]$ is a total order $>$ on the set of monomials in $\Bbbk[x_1, \dots, x_n]$ 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 $\Bbbk[x_1, \dots, x_n]$:

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 $w = (w_1, \dots, w_n)\colon \mathbb{R}^n \to \mathbb{R}$ 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 $\Bbbk[x_1,\dots,x_n]$. 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 $\Bbbk[x_1,\dots,x_n]$, we will abuse notation as follows: If $c,d \in \Bbbk \setminus \{0\}$ are scalars and x^α, x^β are monomials in $\Bbbk[x_1,\dots,x_n]$ 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 $\Bbbk[x_1,\dots,x_n]$, and let $f \in \Bbbk[x_1,\dots,x_n]$ 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 $c \in K$, then c is called the **leading coefficient** of f and x^α is called the **leading monomial** of f. □

Remark 2.2.8. Since a monomial order is defined to be compatible with multiplication,

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

for all $f,g \in \Bbbk[x_1,\dots,x_n]$. Furthermore, if all polynomials involved are nonzero,

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

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 $f = x - x^2, x^2, x^3, \dots$. □

Proposition 2.2.10. *If $>$ is a monomial order on $\Bbbk[x_1,\dots,x_n]$, the following are equivalent:*

1. *$>$ is* **Artinian**, *that is, every nonempty set of monomials has a least element.*
2. *$>$ is* **global**, *that is,*

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

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 $I = \langle X \rangle \subset \Bbbk[x_1, \ldots, x_n]$ 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 $\Bbbk[x_1, \ldots, x_n]$, and let X be a finite set of monomials in $\Bbbk[x_1, \ldots, x_n]$. Prove that there exists a weight order $>_w$ on $\Bbbk[x_1, \ldots, x_n]$ 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 $1, x_1, \ldots, x_n$ 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 $f_1, \ldots, f_r$ 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 syzygy computations (see Corollary 2.3.17).

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

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

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 refer to h as the ***remainder*** *of g on determinate division by $f_1, \ldots, f_r$.*

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.

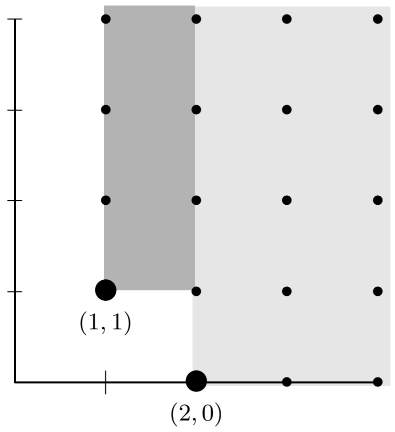
If $f_1, \dots, f_r$ are arbitrary, apply the above to g and $\mathbf{L}(f_1), \dots, \mathbf{L}(f_r)$. If

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

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 $\mathbf{L}(f_2) = xy$. If we group the monomials in $\Bbbk[x, y]$ into different sets as indicated 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 $\Bbbk[x,y,z]$ which yields the leading terms y of $y-x^2$ and z of $z-x^3$, and reconsider part 1 of Exercise 1.5.2. □

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 $g = g_1 f_1 + \ldots + g_r f_r + h$. 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 $>$).

Note that in practical terms, it is often useful to give up uniqueness and to allow choices to be made since some of these 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; 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 $g = g_1 f_1 + \ldots + g_r f_r + h$ 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 $g = g_1 f_1 + \ldots + g_r f_r + h$ 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 $\mathbf{L}(h)$ (which 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 refer to this fact by writing

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

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

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

□

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). Rewrite Theorem 2.2.12, its proof, and Remark 2.2.16 such that they apply to elements of 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 $\Bbbk[x_1, \dots, x_n]$-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 $\Bbbk[x_1, \dots, x_n]$, we refer to $\mathbf{L}(I)$ as the **leading ideal** (or **initial ideal**) of I.
2. A finite subset $\mathcal{G} = \{f_1, \dots, f_r\}$ of I is a **Gröbner basis for *I*** if
$$\mathbf{L}_>(I) = \big\langle \mathbf{L}_>(f_1), \dots, \mathbf{L}_>(f_r)\big\rangle.$$
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 $\Bbbk[x_1, \dots, x_n]$-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 $\{f_1, \dots, f_r\} \subset F \setminus \{0\}$ be a Gröbner basis for the submodule $I := \langle f_1, \dots, f_r\rangle \subset F$. 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 $\mathbf{L}(h) \in \mathbf{L}(I) = \langle \mathbf{L}(f_1), \dots, \mathbf{L}(f_r)\rangle$. 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, $\Bbbk[x_1, \dots, x_n]$ 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 $f_1, \dots, f_r \in I$ such that $\mathbf{L}(I) = \langle \mathbf{L}(f_1), \dots, \mathbf{L}(f_r) \rangle$. That is, $f_1, \dots, f_r$ form a Gröbner basis for I. If $\mathcal{G} \subset F \backslash \{0\}$ 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 by 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 **??** 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 $f_1, \dots, f_r \in F \backslash \{0\}$ 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 a 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

$$\begin{aligned} \mathrm{NF}(f+g,I) &= \mathrm{NF}(f,I)+\mathrm{NF}(g,I), \text{ and} \\ \mathrm{NF}(f\cdot g,I) &= \mathrm{NF}(\mathrm{NF}(f,I)\cdot\mathrm{NF}(g,I),I). \end{aligned}$$ □

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 $g_if_i+g_jf_j$ 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 $\mathbf{L}(g_if_i)$ and $\mathbf{L}(g_jf_j)$ cancel each other in $g_if_i+g_jf_j$:

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**. □

Buchberger's criterion asserts that the simple way of dividing S-vectors by $f_1,\dots,f_r$ allows us 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 of M_i, choose an index $j=j(i,\alpha)<i$ such that $m_{ji}=cx^\alpha$ 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, and show how to use it 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 $\Bbbk[x_1, \dots, z_5]$. The leading terms of the minors and the minimal generators for the corresponding monomial ideals M_i are:

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

Thus, only 15 out of $\binom{10}{2} = 45$ S-vectors are needed in Buchberger's test. In fact, the test shows that the minors form a Gröbner basis (see 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 $f_1, \dots, f_r \in F \setminus \{0\}$, a Gröbner basis for $I := \langle f_1, \dots, f_r \rangle \subset F$ can be computed in finitely many steps.*

Proof. If $f_1, \dots, f_r$ satisfy Buchberger's criterion, we are done. Otherwise, Buchberger's test yields a remainder $0 \neq h \in I$ with $\mathbf{L}(h) \notin \langle \mathbf{L}(f_1), \dots, \mathbf{L}(f_r) \rangle$. That is, $\langle \mathbf{L}(f_1), \dots, \mathbf{L}(f_r) \rangle \subsetneq \langle \mathbf{L}(f_1), \dots, \mathbf{L}(f_r), \mathbf{L}(h) \rangle$. 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 $f_1, \dots, f_r, f_{r+1}, \dots, f_{r'}$ 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

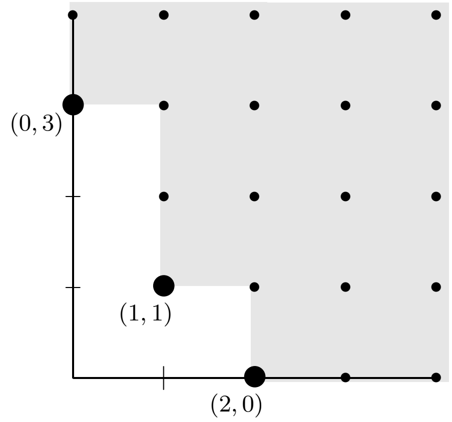
$$\mathrm{S}(f_2, f_1) = x \cdot f_2 - y \cdot f_1 = -xy^2 = 0 \cdot f_1 - y \cdot f_2 - y^3,$$

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 $\mathbb{k}[x_1, \dots, x_n]$ is called a **binomial** if it has at most two terms. An ideal of $\mathbb{k}[x_1, \dots, x_n]$ is called a **binomial ideal** if it is generated by binomials.

Given any global monomial order $>$ and an ideal $I \subset \mathbb{k}[x_1, \dots, x_n]$, show that the following 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 $\mathbb{k}[x_1, \dots, x_n]/I$ 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 $f_1, \dots, f_r$ 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 $(g_1, \dots, g_r)^t \in \mathbb{k}[x_1, \dots, x_n]^r$, and call it a syzygy on $f_1, \dots, f_r$:

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

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

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

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

If $\mathrm{Syz}\ (f_1, \dots, f_r)$ 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 $f_1, \dots, f_r$. □

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

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

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

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,

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

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 $f_1, \ldots, f_r$. 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 $f_1 = x^2, f_2 = xy - y^2, f_3 = -y^3 \in \Bbbk[x, y]$. □

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 $\Bbbk[x_1, \ldots, z_5]$ 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 $\Bbbk[x_1, \ldots, z_5]^{15}$ 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 $\{f_1, \ldots, f_r\} \subset F$ computed with Buchberger's algorithm quite often contains elements whose leading terms are unneeded generators for $\mathbf{L}(\langle f_1, \ldots, f_r \rangle)$. 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 $\mathbf{L}(\langle f_1, \ldots, f_r \rangle)$. In addition, we may "reduce the tail" of each element in the Gröbner basis:

Exercise* 2.3.22. A minimal Gröbner basis $\{f_1, \ldots, f_r\} \subset F$ 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

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

where $m_1, \ldots, m_r$ 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 algorithm:

In the case of a homogeneous linear system of equations, given by polynomials

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

let $>$ be a global monomial order on $\Bbbk[x_1, \dots, x_n]$ 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.

In the case of one variable x, there is precisely one global monomial order: $\cdots > x^2 > x > 1$. 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 2.3.20, we computed the lexicographic Gröbner basis

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

for the ideal $I = \langle f_1, f_2\rangle \subset \Bbbk[x, y]$. 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 (as suggested by the computation in Example 2.3.20), we find that

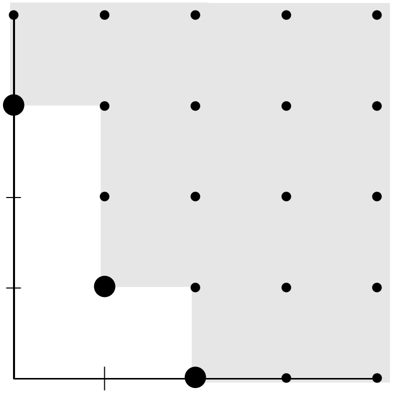
$$g = (x - 1 + y)\cdot f_1 - x \cdot f_2.$$ □

As already remarked earlier, Algorithm 2.4.1 can be used to decide solvability. In fact, by inspecting the Gröbner basis computed in the first step of the algorithm, we get the following information on a set of solutions:

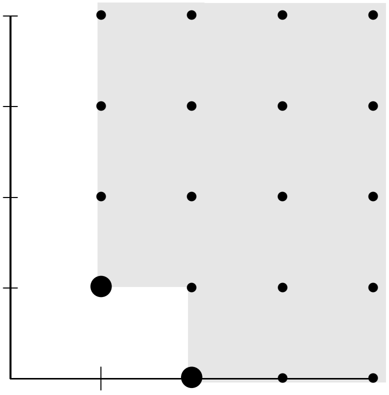
Remark 2.4.3. Let $f_1, \dots, f_r \in \Bbbk[x_1, \dots, x_n] \setminus \{0\}$, let $I := \langle f_1, \dots, f_r \rangle$, and let $\overline{\Bbbk}$ be the algebraic closure of $\Bbbk$. According to Hilbert's Nullstellensatz, Exercise 1.6.5, and Macaulay's Theorem 2.3.5, we can determine whether the system

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

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 every monomial in $\Bbbk[x_1, \dots, x_n]$ 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* 2.4.4. If the system defined by $f_1, \dots, f_r \in \Bbbk[x_1, \dots, x_n]$ has only finitely many solutions in $\mathbb{A}^n(\overline{\Bbbk})$, prove that the number of these solutions is at most $\dim_{\Bbbk} \Bbbk[x_1, \dots, x_n]/\langle f_1, \dots, f_r \rangle$. That is, the number of monomials not in $\mathbf{L}(f_1, \dots, f_r)$ is an upper bound for the number of solutions. Show that these numbers are equal if $\langle f_1, \dots, f_r \rangle$ is a radical ideal. □

Exercise 2.4.5. 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 $\dim_{\mathbb{Q}} \mathbb{Q}[x,y,z]/I$. Indeed, 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 2.4.6. 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 2.4.7 (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 2.4.8. 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 2.4.7, 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). Our first application is an algorithm for computing ideal intersections and, thus, unions of algebraic sets (the correctness of the algorithm is obvious):

Algorithm 2.4.9 (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 $I \cap J$.* □

Our second application of syzygy computations concerns ideal quotients and saturation (geometrically, it concerns the Zariski closure of the difference of two algebraic sets):

Exercise* 2.4.10. 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 2.4.11. 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 need not be a variety.

Remark 2.4.12. 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 decompositions. 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 respect 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. The **degree reverse lexicographic order** on $\mathbb{k}[x_1, \ldots, x_n]$ is defined by

$$x^\alpha >_{\text{drlex}} x^\beta \iff \begin{array}{l} \deg x^\alpha > \deg x^\beta, \text{ or } (\deg x^\alpha = \deg x^\beta \text{ and the} \\ \text{last nonzero entry of } \alpha - \beta \in \mathbb{Z}^n \text{ is negative).} \end{array}$$

It is extended to free $\mathbb{k}[x_1, \ldots, x_n]$-modules as in Remark 2.2.20 (we suggest to give priority to the monomials in $\mathbb{k}[x_1, \ldots, x_n]$). □

Note that as in the case of $>_{\text{lex}}$, we have defined $>_{\text{drlex}}$ such that the variables are ordered according to their appearance when writing $\mathbb{k}[x_1, \ldots, x_n]$. In contrast to $>_{\text{lex}}$, however, $>_{\text{drlex}}$ refines the partial order by total degree:

$$\deg x^\alpha > \deg x^\beta \Longrightarrow x^\alpha >_{\text{drlex}} x^\beta.$$

We will refer to this fact by saying that $>_{\text{drlex}}$ is **degree-compatible**.

Example 2.5.2. With respect to $>_{\text{lex}}$ and $>_{\text{drlex}}$, the monomials of degree 2 in $\mathbb{k}[x, y, z]$ are ordered as follows:

$$x^2 >_{\text{lex}} xy >_{\text{lex}} xz >_{\text{lex}} y^2 >_{\text{lex}} yz >_{\text{lex}} z^2$$

and

$$x^2 >_{\text{drlex}} xy >_{\text{drlex}} y^2 >_{\text{drlex}} xz >_{\text{drlex}} yz >_{\text{drlex}} z^2.$$ □

For monomials of the same degree, the difference between $>_{\text{lex}}$ and $>_{\text{drlex}}$ is subtle but crucial. The use made of these orders depends on their key properties (which, as we will see in Exercise 2.9.4, characterize $>_{\text{lex}}$ and $>_{\text{drlex}}$ among all global monomial orders).

The key property of $>_{\text{drlex}}$ is: $>_{\text{drlex}}$ is degree-compatible, and if $f \in \mathbb{k}[x_1, \ldots, x_n]$ is homogeneous, then

$$\begin{array}{c} >_{\text{drlex}} \text{ chooses the leading term of } f \text{ in a subring } \mathbb{k}[x_1, \ldots, x_k] \\ \text{such that } k \text{ is as small as possible.} \end{array}$$

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 $>_{\text{lex}}$ is that the following holds for all $f \in \mathbb{k}[x_1, \ldots, x_n]$:

$$\mathbf{L}(f) \in \Bbbk[x_{k+1}, \ldots, x_n] \text{ for some } k \implies f \in \Bbbk[x_{k+1}, \ldots, x_n].$$

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 $I \subset \Bbbk[x_1, \ldots, x_n]$ is an ideal, its ***k*****th elimination ideal** is the ideal

$$I_k = I \cap \Bbbk[x_{k+1}, \ldots, x_n]. \qquad \square$$

In particular, $I_0 = I$.

Algorithm 2.5.4 (Elimination Using $>_{\text{lex}}$). *Given $I = \langle f_1, \ldots, f_r \rangle \subset \Bbbk[x_1, \ldots, x_n]$, compute all elimination ideals I_k.*

1. *Compute a Gröbner basis $\mathcal{G}$ for I with respect to $>_{\text{lex}}$ on $\Bbbk[x_1, \ldots, x_n]$.*
2. *For all k, the elements $g \in \mathcal{G}$ with $\mathbf{L}(g) \in \Bbbk[x_{k+1}, \ldots, x_n]$ form a Gröbner basis for I_k with respect to $>_{\text{lex}}$ on $\Bbbk[x_{k+1}, \ldots, x_n]$.*

Proof (of correctness). If $f \in I \cap \Bbbk[x_{k+1}, \ldots, x_n]$, then $\mathbf{L}(f)$ is divisible by $\mathbf{L}(g)$ for some $g \in \mathcal{G}$. Since f does not involve $x_1, \ldots, x_k$, 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

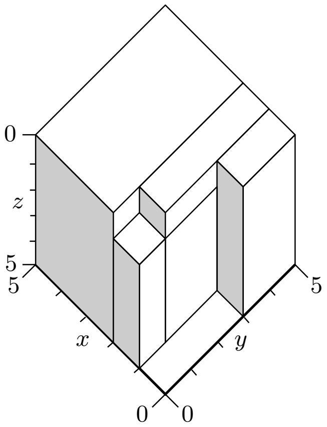
$$\mathrm{S}(f_3, f_2) = y(-xz + y^2) + z(xy - z) = y^3 - z^2 =: f_4.$$

In the next step, $M_4 = \langle x^2, xy, xz \rangle : y^3 = \langle x \rangle$, and

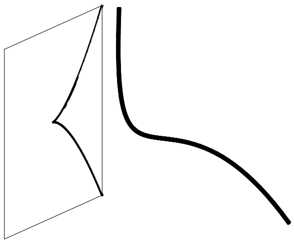
$$\mathrm{S}(f_4, f_2) = x(y^3 - z^2) - y^2(xy - z) = -xz^2 + y^2 z = z(-xz + y^2)$$

is a standard expression with remainder zero. Hence, f_1, f_2, f_3, f_4 form a Gröbner basis for $\langle f_1, f_2 \rangle$.

We visualize the monomials in $\mathbf{L}(f_1, f_2)$ via their exponent vectors:



The computaion shows that the elimination ideal $I_1 \subset \Bbbk[y,z]$ is generated by the polynomial $f_4 = y^3 - z^2$. The geometric interpretation of this is (see Section 2.6 below):



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 sending (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 $\langle f_1, f_2\rangle$. In contrast, show that f_1, f_2, and f_3 form the reduced Gröbner basis with respect to $>_{\mathrm{drlex}}$. □

A single Gröbner basis computation with respect to $>_{\mathrm{lex}}$ yields the whole flag of elimination ideals I_k, $k = 0, \ldots, n-1$. 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

$$\Bbbk[\boldsymbol{x}, \boldsymbol{y}] = \Bbbk[x_1, \ldots, x_n, y_1, \ldots y_m]$$

is an **elimination order** with respect to $x_1, \ldots, x_n$ if the following holds for all $f \in \Bbbk[\boldsymbol{x}, \boldsymbol{y}]$:

$$\mathbf{L}(f) \in \Bbbk[\boldsymbol{y}] \implies f \in \Bbbk[\boldsymbol{y}]. \qquad \square$$

Example 2.5.8. Given monomial orders $>_1$ on $\Bbbk[\boldsymbol{x}]$ and $>_2$ on $\Bbbk[\boldsymbol{y}]$, the **product order** (or **block order**) $> = (>_1, >_2)$ on $\Bbbk[\boldsymbol{x}, \boldsymbol{y}]$, defined by

$$x^\alpha y^\gamma > x^\beta y^\delta \iff x^\alpha >_1 x^\beta, \textit{ or } (x^\alpha = x^\beta \textit{ and } y^\gamma >_2 y^\delta),$$

is an elimination order with respect to $x_1, \dots, x_n$. It is global if $>_1$ and $>_2$ are global. Note that it is often most efficient to pick $>_1$ and $>_2$ to be degree reverse lexicographic. □

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 **??**). □

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, we 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 $\Bbbk[x_1,\dots,x_n]$, and let $\overline{f}_1 = f_1 + I, \dots, \overline{f}_m = f_m + I \in \Bbbk[x_1,\dots,x_n]/I$. Consider the homomorphism*

$$\phi : \Bbbk[y_1,\dots,y_m] \to S = \Bbbk[x_1,\dots,x_n]/I,\ y_i \mapsto \overline{f}_i.$$

If J is the ideal

$$J = I\,\Bbbk[\boldsymbol{x},\boldsymbol{y}] + \langle f_1 - y_1, \dots, f_m - y_m\rangle \subset \Bbbk[\boldsymbol{x},\boldsymbol{y}],$$

then

$$\ker\phi = J \cap \Bbbk[\boldsymbol{y}].$$

Proof. Let $g \in \Bbbk[\boldsymbol{y}] \subset \Bbbk[\boldsymbol{x},\boldsymbol{y}]$. 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 $h \in I\,\Bbbk[\boldsymbol{x},\boldsymbol{y}]$, then $g(f_1,\dots,f_m) = h(x_1,\dots,x_n,f_1,\dots,f_m) \in I\,\Bbbk[\boldsymbol{x},\boldsymbol{y}] \cap \Bbbk[\boldsymbol{x}] = I$.

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

$$g(y) = g(f_1,\dots,f_m) + \sum g_j(f_j - y_j).$$ □

The proposition gives us, in particular, a method for computing in the algebra $\Bbbk[\overline{f}_1,\dots,\overline{f}_m] \cong \Bbbk[y_1,\dots,y_m]/\ker\phi$ since we already know how to compute in affine rings. Once we have the required Gröbner basis for J, we know, in particular, whether $\ker\phi = 0$, that is, whether $\overline{f}_1,\dots,\overline{f}_m$ are algebraically independent over $\Bbbk$.

Besides checking the injectivity of ϕ, we may also check its surjectivity. More generally, we can decide subalgebra membership:

Exercise* 2.5.13 (Subalgebra Membership). Let $\overline{g}, \overline{f}_1,\dots,\overline{f}_m$ be elements of an affine ring $\Bbbk[x_1,\dots,x_n]/I$, let $J \subset \Bbbk[\boldsymbol{x},\boldsymbol{y}]$ be the ideal defined in Proposition 2.5.12 above, and let $>$ be a global elimination order on $\Bbbk[\boldsymbol{x},\boldsymbol{y}]$ with respect to $x_1,\dots,x_n$. Show:

1. We have $\overline{g} \in \Bbbk[\overline{f}_1,\dots,\overline{f}_m]$ iff the normal form $h = \mathrm{NF}(g,J) \in \Bbbk[\boldsymbol{x},\boldsymbol{y}]$ is contained in $\Bbbk[\boldsymbol{y}]$. 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 $\phi : \Bbbk[y_1,\dots,y_m] \to \Bbbk[x_1,\dots,x_n]/I$ is surjective iff $\mathrm{NF}(x_i,J) \in \Bbbk[\boldsymbol{y}]$ for $i = 1,\dots,n$. □

Exercise 2.5.14. 1. Compute the algebra relations on the polynomials

$$f_1 = x^2 + y^2,\ f_2 = x^2y^2,\ f_3 = x^3y - xy^3 \in \Bbbk[x,y].$$

2. Consider the polynomials
$$g = x^4 + y^4,\ g_1 = x + y,\ g_2 = xy \in \Bbbk[x, y].$$
Show that g is contained in the subalgebra $\Bbbk[g_1, g_2] \subset \Bbbk[x, y]$, and express g as a polynomial in g_1, g_2.
3. Consider the endomorphism ϕ of $\Bbbk[x_1, x_2, x_3]$ 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
$$\Bbbk[x_1, x_2, x_3]/\langle x_1x_2x_3 - 1\rangle.$$
This means that the variety $A = \mathrm{V}(x_1x_2x_3 - 1) \subset \mathbb{A}^3(\Bbbk)$ admits a nonlinear automorphism. Determine the fixed points of this automorphism.

□

2.6 The Geometry of Elimination

Geometrically, eliminating variables amounts to projection. To formulate a precise result, fix an integer k, $1 \leq k \leq n - 1$, and consider the projection map
$$\pi_k : \mathbb{A}^n(\Bbbk) \to \mathbb{A}^{n-k}(\Bbbk),\ (a_1, \ldots, a_n) \mapsto (a_{k+1}, \ldots, a_n).$$
Let $I \subset \Bbbk[x_1, \ldots, x_n]$ be an ideal, let
$$I_k = I \cap \Bbbk[x_{k+1}, \ldots, x_n]$$
be its kth elimination ideal, and let $A = \mathrm{V}(I) \subset \mathbb{A}^n(\Bbbk)$. Then
$$\pi_k(A) \subset \mathrm{V}(I_k) \subset \mathbb{A}^{n-k}(\Bbbk) \tag{2.6}$$
since every element $f \in I_k \subset I$ vanishes on A and, thus, on $\pi_k(A)$. Note that the inclusion $\pi_k(A) \subset \mathrm{V}(I_k)$ may well be strict. In fact, even over an algebraically closed field, the image $\pi_k(A)$ needs not be Zariski closed:

Exercise 2.6.1. Let $\Bbbk$ be any field, and let $I = \langle xy - 1, y^2 - z\rangle \subset \Bbbk[x, y, z]$. We project $A = \mathrm{V}(I) \subset \mathbb{A}^3(\Bbbk)$ to the yz-plane: Apply Algorithm 2.5.4 to show that $I_1 = \langle y^2 - z\rangle \subset \Bbbk[y, z]$ is the first elimination ideal of I. Then observe that the origin $o = (0, 0)$ is a point of $\mathrm{V}(I_1) \subset \mathbb{A}^2(\Bbbk)$ which has no preimage in A. □

In Chapter **??**, 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 **??** 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 which vanishes on $\pi_k(A)$ and, thus, on A. Then, by Hilbert's Nullstellensatz, $f^m \in I \cap \Bbbk[x_{k+1}, \dots, x_n] = I_k$ for some $m \geq 1$. It follows that $\mathrm{I}(\pi_k(A)) \subset \mathrm{rad}\, I_k$, so that

$$\overline{\pi_k(A)} = \mathrm{V}(\mathrm{I}(\pi_k(A))) \supset \mathrm{V}(\mathrm{rad}\, I_k) = \mathrm{V}(I_k). \qquad \square$$

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 $\mathbb{k}$ is not algebraically closed, suppose that A is Zariski dense in the locus of zeros of I in $\mathbb{A}^n(\overline{\mathbb{k}})$. Then*

$$\overline{\varphi(A)} = \mathrm{V}(J \cap \mathbb{k}[\boldsymbol{y}]) \subset \mathbb{A}^m(\mathbb{k}).$$

Proof. We write $\overline{\mathrm{V}}$ for taking loci of zeros over $\overline{\mathbb{k}}$ and $\overline{\varphi}$ for the morphism $\overline{\mathrm{V}}(I) \to \mathbb{A}^m(\overline{\mathbb{k}})$ given by $f_1, \dots, f_m$.

From the proof of the preceeding corollary, we already know that

$$\varphi(A) \subset \mathrm{V}(J \cap K[\boldsymbol{y}]) \subset \mathbb{A}^m(\mathbb{k}).$$

To show that $\mathrm{V}(J \cap K[\boldsymbol{y}])$ is the smallest algebraic subset of $\mathbb{A}^m(\mathbb{k})$ containing $\varphi(A)$, let $g \in \mathbb{k}[y_1, \dots, y_m]$ be any polynomial vanishing on $\varphi(A)$. Then the polynomial $g(f_1, \dots, f_m) \in \mathbb{k}[x_1, \dots, x_n]$ vanishes on A and, thus, on $\overline{\mathrm{V}}(I) \subset \mathbb{A}^n(\overline{\mathbb{k}})$ since A is Zariski dense in $\overline{\mathrm{V}}(I)$. So g vanishes on $\overline{\varphi}(\overline{\mathrm{V}}(I)) \subset \mathbb{A}^m(\overline{\mathbb{k}})$ and, thus, on $\overline{\mathrm{V}}(J \cap \mathbb{k}[\boldsymbol{y}])$ by the preceeding corollary. In particular, g vanishes on $\mathrm{V}(J \cap \mathbb{k}[\boldsymbol{y}])$. We conclude that every algebraic subset containing $\varphi(A)$ must contain $\mathrm{V}(J \cap \mathbb{k}[\boldsymbol{y}])$ as well. □

Remark 2.6.5. If $\mathbb{k}$ is infinite, then $\mathbb{A}^n(\mathbb{k})$ is Zariski dense in $\mathbb{A}^n(\overline{\mathbb{k}})$. Indeed, the same argument as in Exercise 1.2.1 shows that if $f \in \overline{\mathbb{k}}[x_1, \dots, x_n]$ is a polynomial vanishing on $\mathbb{A}^n(\mathbb{k})$, then f is zero. □

Exercise 2.6.6 (Steiner Roman Surface). Consider the real 2-sphere

$$S^2 = \mathrm{V}(x_1^2 + x_2^2 + x_3^2 - 1) \subset \mathbb{A}^3(\mathbb{R})$$

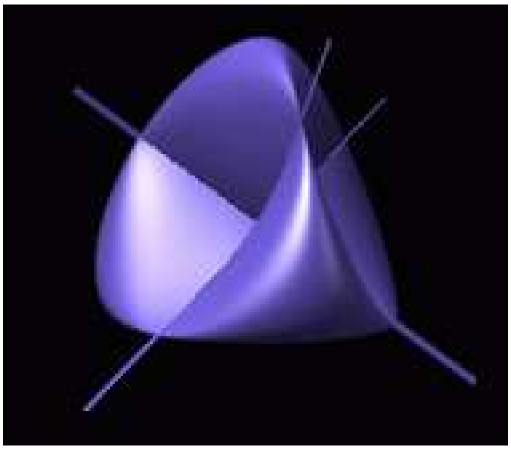
and the morphism

$$\varphi : S^2 \to \mathbb{A}^3(\mathbb{R}),\ (a_1, a_2, a_3) \mapsto (a_1a_2, a_1a_3, a_2a_3).$$

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 an algebraic subset. 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

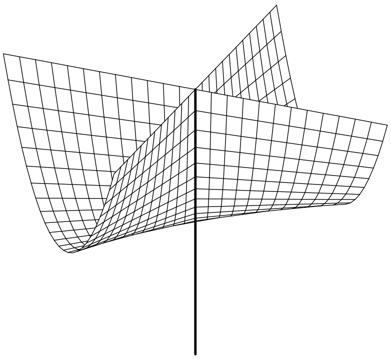
$$J = \langle x - t,\ y - t^2,\ z - t^3 \rangle \subset \Bbbk[t, x, y, z]$$

with respect to the product order $(>_1, >_2)$, where $>_2$ is the degree reverse lexicographic orders 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. □

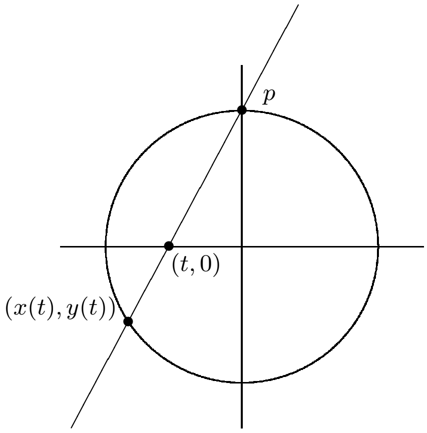
Exercise 2.6.10. If $\operatorname{char} \Bbbk \neq 2$, show that the circle

$$C = \mathrm{V}(x^2 + y^2 - 1) \subset \mathbb{A}^2(\Bbbk)$$

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, using the point $p = (0, 1)$ 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). □

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* $\operatorname{dom}(f)$ *is open and dense in the Zariski topology on V.*
2. *If $\Bbbk = \overline{\Bbbk}$ is algebraically closed, then*

$$\operatorname{dom}(f) = V \iff f \in \Bbbk[V].$$

In other words, a rational function $f \in \Bbbk(V)$ is defined everywhere on V iff f is a polynomial function on V.

Proof. Considering the *ideal I_f of denominators* of f,

$$I_f = \{h \in \Bbbk[V] \mid \text{there is an expression } f = g/h \text{ with } g \in \Bbbk[V]\} \cup \{0\},$$

we find that

$$V \setminus \operatorname{dom}(f) = \mathrm{V}_V(I_f)$$

is an algebraic subset of V. Hence, $\operatorname{dom}(f)$ is Zariski open and, being nonempty, dense in the Zariski topology on V (see Proposition 1.11.8). Furthermore, if $\Bbbk = \overline{\Bbbk}$, then

$$\operatorname{dom}(f) = V \iff \mathrm{V}_V(\mathrm{I}_f) = \emptyset \iff 1 \in \mathrm{I}_f \iff f \in \Bbbk[V]$$

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 $(f_1, \ldots, f_m)$ 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

$$\varphi : \operatorname{dom}(\varphi) \to \mathbb{A}^m(\Bbbk),\ p \mapsto \varphi(p) := (f_1(p), \ldots, f_m(p)).$$

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 $g_i, h_i \in \Bbbk[x_1, \ldots, x_n]$. Suppose that V is Zariski dense in the locus of zeros of I in $\mathbb{A}^n(\overline{\Bbbk})$. Design an algorithm for computing the smallest algebraic subset of $\mathbb{A}^m(\Bbbk)$ containing $\varphi(\operatorname{dom}(\varphi))$.
Hint. Consider the ideal

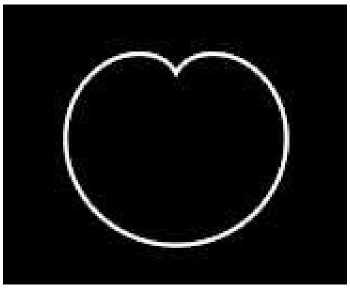
$$J = I\,\Bbbk[\boldsymbol{x}, \boldsymbol{y}] + \langle h_1 y_1 - g_1, \ldots, h_m y_m - g_m, zh - 1\rangle \subset \Bbbk[\boldsymbol{x}, \boldsymbol{y}, z],$$

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 1.11.12, 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 $\varphi^{-1}(\mathrm{dom}(\psi)) \cap \mathrm{dom}(\varphi)$. 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, the difficulty shows as follows. Arguing as in the proof of Theorem 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 $(f_1, \dots, f_m)$ 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 1.11.12 for polynomial maps:

Theorem 2.6.22. *Let $V \subset \mathbb{A}^n(\Bbbk)$ and $W \subset \mathbb{A}^m(\Bbbk)$ be varieties.*

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 $\operatorname{dom}(\psi \circ \varphi)$ contains $\varphi^{-1}(\operatorname{dom}(\psi))$, but may well be larger (see Exercise 2.6.26 below for examples).

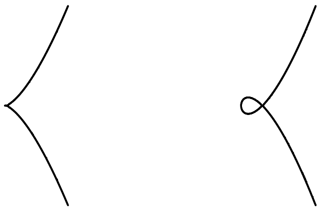
According to our definition in Chapter 1, an isomorphism of algebraic sets is a morphism admitting an inverse morphism. Similarly, we define:

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 there 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 $\varphi^{-1}(\operatorname{dom}(\psi))$. □

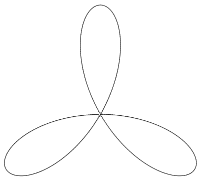
Now, finally, we come to the definition of a rational parametrization:

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 **??**, in Section **??**, 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 Chapter 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 **??** 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 the ground field to be 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 $\Bbbk \subset \mathbb{K}$ 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 2.4.12: 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 $\Bbbk \subset \mathbb{K}$ 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 all the remarks above 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** $\Bbbk$. Furthermore, 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 **quadric of full rank**. 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

$$\ldots \longrightarrow M_{i+1} \xrightarrow{\phi_{i+1}} M_i \xrightarrow{\phi_i} M_{i-1} \longrightarrow \ldots$$

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

$$0 \longrightarrow M' \xrightarrow{\phi} M \xrightarrow{\psi} M'' \longrightarrow 0.$$

[1] not without difficulty

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

$$\ldots \longrightarrow M_{i+1} \xrightarrow{\phi_{i+1}} M_i \xrightarrow{\phi_i} M_{i-1} \longrightarrow \ldots$$

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

$$0 \to M_r \to M_{r-1} \to \ldots \to M_s \to 0$$

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

$$0 \longrightarrow \operatorname{Hom}_R(M, N') \longrightarrow \operatorname{Hom}_R(M, N) \longrightarrow \operatorname{Hom}_R(M, N'')$$

is exact. Give an example which shows that the map $\operatorname{Hom}_R(M, N) \longrightarrow \operatorname{Hom}_R(M, N'')$ 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 and an epimorphism $F_1 \to \ker\pi$. If ϕ is the composite map $F_1 \to \ker\pi \to F_0$, then $M = \operatorname{coker}\phi$.

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**. Furthermore, 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 started 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 $\phi_1 = \phi$).

Remark-Definition 2.8.7. Every exact sequence as above is called a **free resolution** of M. We call $\operatorname{im}\phi_i$ an ***i*th syzygy module** and its elements ***i*th order syzygies** of M (note that these modules depend on the choices made). 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.

It follows from our 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. □

Exercise* 2.8.8. If R is a PID, then every finitely generated R-module M has a presentation

$$0 \longrightarrow F_1 \xrightarrow{\phi} F_0 \longrightarrow M \longrightarrow 0$$

in **Smith normal form**. That is, F_1 and F_0 are free modules with finite ranks satisfying $\operatorname{rank} F_1 \leq \operatorname{rank} F_0$, and ϕ is a syzygy matrix of type

$$\phi = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \ldots & d_m \\ 0 & & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & & \ldots & 0 \end{pmatrix},$$

where d_i divides d_{i+1}, for $i = 1, \ldots, m-1$ (the d_i are usually called the **elementary divisors** of M). In the case where R is Euclidean, prove this result by designing an algorithm which computes the normal form starting from an arbitrary presentation. See Newman (1972) for more information. □

The exercise shows, 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 2.8.9 (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 $1 \leq k \leq n$ 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). Suppose that in Buchberger's test the f_i are arranged such that, for $i > j$, the exponent of x_k in $\mathbf{L}(f_j)$ is strictly smaller 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

$$R^{s_\ell} \xrightarrow{\phi_\ell} R^{s_{\ell-1}} \longrightarrow \dots \longrightarrow R^{s_1} \xrightarrow{\phi_1} R^{s_0} \longrightarrow M \longrightarrow 0$$

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 $x_1, \dots, x_n$. Having chosen a minimal Gröbner basis in the previous step, this implies that $\mathcal{G} = \{0\}$. Thus, $\ker \phi_\ell = 0$ and

$$0 \longrightarrow R^{s_\ell} \xrightarrow{\phi_\ell} R^{s_{\ell-1}} \longrightarrow \dots \longrightarrow R^{s_1} \xrightarrow{\phi_1} R^{s_0} \longrightarrow M \longrightarrow 0$$

is a finite free resolution as desired. □

Example 2.8.10. Our computations in Exercise 2.3.21 show that the affine ring R/I, where $R = \mathbb{k}[x_1, \dots, z_5]$ 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

$$0 \longrightarrow R^6 \longrightarrow R^{15} \longrightarrow R^{10} \longrightarrow R \longrightarrow R/I \longrightarrow 0. \qquad \square$$

Example 2.8.11. Consider the ideal

$$I = \langle f_1, \dots, f_5 \rangle \subset R = \mathbb{k}[w, x, y, z]$$

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 $\mathbb{k}[w, x, y, z]$. We successively obtain three syzygy matrices ϕ_1, ϕ_2, and ϕ_3 which we present in a compact way as follows:

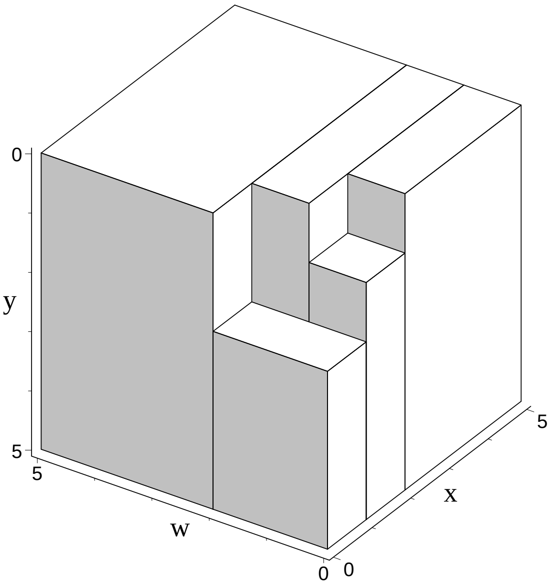
$$\begin{array}{r|cccccc}
\mathbf{w^2} - xz & -x & y & 0 & -z & 0 & -y^2 + wz \\
\mathbf{wx} - yz & \mathbf{w} & -x & -y & 0 & z & z^2 \\
\mathbf{x^2} - wy & -z & \mathbf{w} & 0 & -y & 0 & 0 \\
\mathbf{xy} - z^2 & 0 & 0 & \mathbf{w} & \mathbf{x} & -y & yz \\
\mathbf{y^2} - wz & 0 & 0 & -z & -w & \mathbf{x} & \mathbf{w^2} \\
\hline
 & 0 & y & -x & \mathbf{w} & -z & 1 \\
 & -y^2 + wz & z^2 & -wy & yz & -w^2 & \mathbf{x}
\end{array}$$

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 2.8.12. 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 2.8.13. 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 **??**.
3. Some references for further reading on free resolutions are Serre (1965), Eisenbud (1995), and Avramov (1989). □

Exercise 2.8.14 (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 $\Bbbk[x_1, \ldots, x_n]$. □

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

$$\operatorname{Hom}_R(0 \to N' \to N \to N'' \to 0, M)$$

and

$$M \otimes_R (0 \to N' \to N \to N'' \to 0)$$

respectively.

To compute $\operatorname{Ext}^i_R(M, N)$ 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

$$\operatorname{Ext}^i_R(M, N) \cong \mathrm{H}^i \operatorname{Hom}(M, I^*) \cong \mathrm{H}^i \operatorname{Hom}(F_*, N).$$

Similarly,

$$\operatorname{Tor}^R_i(M, N) \cong \mathrm{H}_i(F_* \otimes N).$$

Since injective modules are rarely finitely presented we work with a projective or even simpler free resolution. We proceed in several steps.

Remark 2.8.15 (Presentation of homomorphism). Any morphism $\varphi \in \operatorname{Hom}_R(M, N)$ 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 2.8.16 (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* $\operatorname{Hom}_R(M, N)$.

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

$$\operatorname{Hom}(E_0, F_1) \oplus \operatorname{Hom}(E_1, F_2) \to \operatorname{Hom}(E_0, F_0) \oplus \operatorname{Hom}(E_1, F_1) \to \operatorname{Hom}(E_1, F_0)$$

defined by

$$(h_0, h_1) \mapsto (b_1 h_0, h_0 a_1 - b_2 h_1) \text{ and } (\varphi_0, \varphi_1) \mapsto \varphi_0 a_1 - \varphi_1 b_1.$$

Note that for free modules F and G the module $\operatorname{Hom}(F, G)$ is free of $\operatorname{rank} \operatorname{Hom}(F, G) = \operatorname{rank} F \operatorname{rank} G$. 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 2.8.17 (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*

$$\mathrm{H} = \mathrm{H}(M \to N \to L) = \frac{\ker(\psi)}{\operatorname{im}(\varphi)}.$$

1. Compute the syzygies matrix $(h_0 \; g_0)^t$ *of* $\begin{pmatrix}\psi_0 & c_1\end{pmatrix}$*:*

$$H_0 \xrightarrow{\begin{pmatrix} h_0 \\ g_0 \end{pmatrix}} F_0 \oplus G_1 \xrightarrow{\begin{pmatrix}\psi_0 & c_1\end{pmatrix}} G_0.$$

2. Compute the syzygy matrix $(h_1 \; f_1 \; e_0)^t$ *in*

$$H_1 \xrightarrow{\begin{pmatrix} h_1 \\ f_1 \\ e_0 \end{pmatrix}} H_0 \oplus F_1 \oplus E_0 \xrightarrow{\begin{pmatrix} h_0 & b_1 & \varphi_0 \end{pmatrix}} F_0.$$

3. $\mathrm{H} = \operatorname{Coker}(H_1 \xrightarrow{h_1} H_0)$.

Exercise 2.8.18 (ker, coker)**.** Give an algorithm to compute $\mathrm{Ker}(\psi : N \to L)$ and $\mathrm{Coker}(\varphi : M \to N)$ by simplifying the computation of homology in these cases.

Exercise 2.8.19. Complete the Algorithm 2.8.16 for the computation of $\mathrm{Hom}(M, N)$ by including a simplified version of homology in this cases.

Exercise 2.8.20. Given a homorphism $N' \xrightarrow{f} N$ and a module M specified by presentations, design an algorithm which computes the presentations of

$$\mathrm{Hom}(N, M) \quad \overset{\mathrm{Hom}(f,M)}{\longrightarrow} \quad \mathrm{Hom}(N', M)$$

and

$$\mathrm{Hom}(M, N') \quad \overset{\mathrm{Hom}(M,f)}{\longrightarrow} \quad \mathrm{Hom}(M, N).$$

Algorithm 2.8.21 (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* $\mathrm{Ext}^i_R(M, N)$*.*

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* $\mathrm{Hom}(F_*, N)$*:*

$$\begin{tikzcd}
\mathrm{Hom}(F_{i-1}, G_1) \arrow[r, "\mathrm{Hom}(F_{i-1},b_1)"] \arrow[d, "\mathrm{Hom}(a_i,G_1)"] & \mathrm{Hom}(F_{i-1}, G_0) \arrow[r] \arrow[d, "\mathrm{Hom}(a_i,G_0)"] & \mathrm{Hom}(F_{i-1}, N) \arrow[r] \arrow[d, "\mathrm{Hom}(a_i,N)"] & 0 \\
\mathrm{Hom}(F_i, G_1) \arrow[r, "\mathrm{Hom}(F_i,b_1)"] \arrow[d, "\mathrm{Hom}(a_{i+1},G_1)"] & \mathrm{Hom}(F_i, G_0) \arrow[r] \arrow[d, "\mathrm{Hom}(a_{i+1},G_0)"] & \mathrm{Hom}(F_i, N) \arrow[r] \arrow[d, "\mathrm{Hom}(a_{i+1},N)"] & 0 \\
\mathrm{Hom}(F_{i+1}, G_1) \arrow[r, "\mathrm{Hom}(F_{i+1},b_1)"] & \mathrm{Hom}(F_{i+1}, G_0) \arrow[r] & \mathrm{Hom}(F_{i+1}, N) \arrow[r] & 0
\end{tikzcd}$$

3. *Compute a presentation of the homology*

$$\mathrm{Ext}^i_R(M, N) = \mathrm{H}^i(\mathrm{Hom}(F_*, N)) = \frac{\ker \mathrm{Hom}(a_{i+1}, N)}{\mathrm{im}\, \mathrm{Hom}(a_i, N)}$$

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 2.8.22 (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* $\mathrm{Tor}_i^R(M, N)$.

1. *Compute $i+1$ steps of a free resolution of M:*

$$F_{i+1} \overset{a_{i+1}}{\longrightarrow} F_i \overset{a_i}{\longrightarrow} F_{i-1} \overset{a_{i-1}}{\longrightarrow} \dots \overset{a_2}{\longrightarrow} F_1 \overset{a_1}{\longrightarrow} F_0$$

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*

$$\mathrm{Tor}_i^R(M, N) = \mathrm{H}_i(F_* \otimes N) = \frac{\ker(a_i \otimes id_N)}{\mathrm{im}(a_{i+1} \otimes id_N)}$$

Exercise 2.8.23. 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

$$\operatorname{Ext}^i_R(M, N'') \to \operatorname{Ext}^{i+1}_R(M, N')$$

$$\operatorname{Ext}^i_R(N', M) \to \operatorname{Ext}^{i+1}_R(N'', M)$$

and

$$\operatorname{Tor}^R_i(N', M) \to \operatorname{Tor}^R_{i-1}(N'', M)$$

of the long exact sequences.

Exercise 2.8.24. 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 $\operatorname{Tor}^R_i(M, N)$.

Besides measuring exactness $\operatorname{Ext}^1_R(M, N)$ is used to describe extensions.

Definition 2.8.25. 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 2.8.26. Prove:

1. An extension
$$0 \to A \to E \to C \to 0$$
is isomorphic to the split extention iff $id_C \in \operatorname{Hom}_R(C, C)$ maps to $0 \in \operatorname{Ext}^1_R(C, A)$ under the connecting homomorphism δ.
2. For every class $e \in \operatorname{Ext}^1(C, A)$ 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

$$f_1 = x_1^q - x_q, \dots, f_n = x_n^q - x_n \in \mathbb{F}_p[x_1 \dots, x_n]$$

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 $\mathbb{F}_p[x_1 \dots, x_n]$).

Exercise 2.9.2. Let $f_1, \dots, f_r \in F = R$ 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, $x_2^2 >_{\mathrm{drlex}} x_1x_3$. Since $>$ is compatible with multiplication, also $x_2^2x_4 >_{\mathrm{drlex}} x_1x_3x_4$. □

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

$$f_1 = x_1 - h_1(x_2, \dots, x_n), \dots, f_{n-1} = x_{n-1} - h_{n-1}(x_n), f_n = h_n(x_n)$$

are called **triangular** (the Gröbner basis in Exercise 2.4.5 gives an example). Such a system has at most $\deg f_n$ solutions in $\mathbb{A}^n(\overline{\mathbb{k}})$. 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.

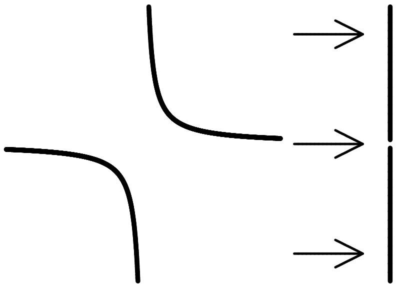
3.1 Proof of the Nullstellensatz

With notation as in Convention 2.7.2, our goal is to show that if $I \subset \Bbbk[x_1, \dots, x_n]$ 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 algebraic. 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 algebraic.



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 an algebraic set.

Proof. The inclusion $\pi(A) \subset \mathrm{V}(I_1)$ holds since $I_1 \subset I$. For the opposite inclusion, let $p' = (a_2, \dots, a_n) \in \mathbb{A}^{n-1} \setminus \pi(A)$ 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 $g \in \Bbbk[x_1, \dots, x_n]$, there is a polynomial $\widetilde{g} \in \Bbbk[x_1, \dots, x_n]$ 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

$$\phi\colon \Bbbk[x_1, \dots, x_n] \to \Bbbk[x_1],\ g \mapsto g(x_1, p').$$

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 $g \in \Bbbk[x_1, \dots, x_n]$ 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 $\Bbbk[x_2, \dots, x_n][x_1]$ 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 $1, x_1, \dots, x_1^{d-1}$, 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 $g_{ij} \in \Bbbk[x_2, \dots, x_n]$ 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, $h := \det(E_d - B) \cdot 1 \in I \cap \Bbbk[x_2, \dots, x_n] = I_1$. 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 $J_1 := I_1\,\mathbb{K}[x_2,\dots,,x_n]$ 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 J_1 which does not vanish at the point p'. Since J_1 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.

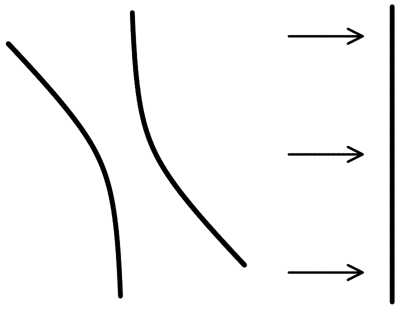
Proof. 1. Let $f = f_d + f_{d-1} + \ldots + f_0$, $f_d \neq 0$, be the decomposition of f into its homogeneous components (the degree of f_k is k if $f_k \neq 0$). After substituting $\widetilde{x}_i + a_i x_1$ for x_i in f, $i = 2,\dots,n$, the coefficient of x_1^d is $f_d(1,a_2,\dots,a_n)$. Since f_d 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

$$\pi = \pi_c \circ \cdots \circ \pi_1 : \mathrm{V}(I) \to \mathbb{A}^{n-c}.$$

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

$$\Phi_k \colon \Bbbk[x_k, x_{k+1}, \dots, x_n] \to \mathbb{K}[x_k],\ x_{k+1} \mapsto a_{k+1}, \dots, x_n \mapsto a_n.$$

The image $\Phi_k(I_{k-1})$ 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 of the desired monic polynomials 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

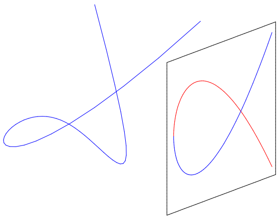
$$\begin{aligned} f_1 &= y^3z - 2y^2z - z^3 + x^2 + z, \\ f_2 &= xy^3z - 2xy^2z - xz^3 + x^3 + y^3 - 2y^2 + xz - z^2 + y. \end{aligned}$$

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.

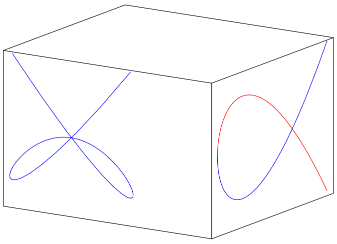


If we reorder the variables as $y > z > x$, the reduced lexicographic Gröbner basis for $\langle f_1, f_2\rangle$ consists of five polynomials:

$$\begin{array}{lll} y^3-2y^2+y-z^2, & y^2x^2-yx^2-z^3, & yz-z-x^2, \\ yx^4-z^4, & z^5-zx^4-x^6. & \end{array}$$

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 **??** 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

$$R = \Bbbk[x_2, \dots, x_n]/I_1 \subset S = \Bbbk[x_1, \dots, x_n]/I$$

which is induced by the inclusion $\Bbbk[x_2, , \dots, x_n] \subset \Bbbk[x_1, \dots, x_n]$. 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

$$\text{finite type } + \text{ integral } = \text{ finite:}$$

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 $s \in \operatorname{rad}(IR[s])$.

$2 \implies 3$: 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 $s \in \operatorname{rad}(IS')$, 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. □

In the situation of the projection theorem,

$$R = \Bbbk[x_2, \dots, x_n]/I_1 \subset S = \Bbbk[x_1, \dots, x_n]/I = R[\overline{x}_1]$$

is a finite (hence integral) ring extension.

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. □

According to our definitions, any extension of affine rings is of finite type. It is, thus, finite iff it is integral.

Exercise* 3.2.6 (Integrality Criterion for Affine Rings). Let I be an ideal of $\Bbbk[x_1, \dots, x_n]$, and let $\overline{f}_1 = f_1 + I, \dots, \overline{f}_m = f_m + I \in \Bbbk[x_1, \dots, x_n]/I$. Consider a polynomial ring $\Bbbk[y_1, \dots, y_m]$, the homomorphism

$$\phi : \Bbbk[y_1, \dots, y_m] \to S = \Bbbk[x_1, \dots, x_n]/I, \ y_i \mapsto \overline{f}_i,$$

and the ideal

$$J = I\,\Bbbk[\boldsymbol{x}, \boldsymbol{y}] + \langle f_1 - y_1, \dots, f_m - y_m \rangle \subset \Bbbk[\boldsymbol{x}, \boldsymbol{y}].$$

Let $>$ be an 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 J with respect to $>$. By Proposition 2.5.12, the elements of $\mathcal{G} \cap \Bbbk[\boldsymbol{y}]$ generate $\ker \phi$. View $R := \Bbbk[y_1, \dots, y_m]/\ker \phi$ as a subring of S by means of ϕ. Show that $R \subset S$ is integral iff for each i, $1 \le i \le n$, there is an element of $\mathcal{G}$ whose leading monomial is of type $x_i^{\alpha_i}$ for some $\alpha_i \ge 1$. □

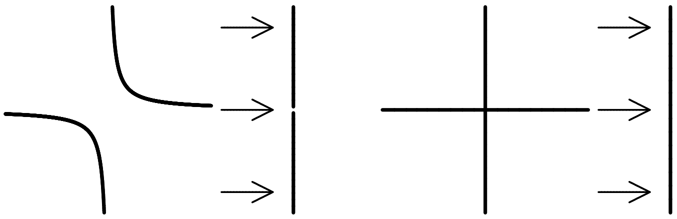
Example 3.2.7. Both ring extensions

$$\Bbbk[y] \subset \Bbbk[x, y]/\langle xy - 1 \rangle, \ y \mapsto \overline{y},$$

and

$$\Bbbk[y] \subset \Bbbk[x, y]/\langle xy \rangle, \ y \mapsto \overline{y},$$

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 $\mathfrak{p}$:*

$$\begin{tikzcd}
{}^{\exists}\,\mathfrak{P} \arrow[d, dash, dotted] \arrow[r, phantom, "\subset"] & S \arrow[d, dash] \\
\mathfrak{p} \arrow[r, phantom, "\subset"] & R
\end{tikzcd}$$

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 $\mathfrak{p}S$. $\square$

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 $\mathfrak{p}$. $\square$

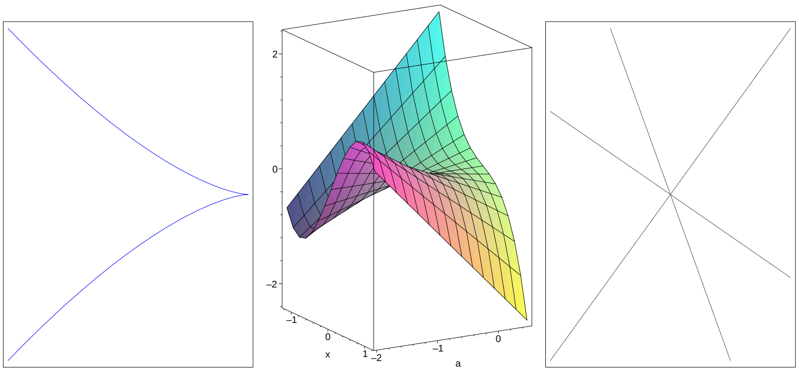
Example 3.2.15. The extension of polynomial rings

$$R = \mathbb{R}[e_2, e_3] \subset T = \mathbb{R}[t_1, t_2]$$

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 $\mathfrak{P}_1 \subset \mathfrak{P}_2$:*

$$\begin{tikzcd}
\mathfrak{P}_1 \arrow[r, phantom, "\subset\,^{\exists}"] \arrow[d, dash] & \mathfrak{P}_2 \arrow[d, dash, dotted] \\
\mathfrak{p}_1 \arrow[r, phantom, "\subset"] & \mathfrak{p}_2
\end{tikzcd}$$

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 now, 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 continuous map between topological spaces is said to be **closed** if it sends closed sets to closed sets.

Corollary 3.2.19. *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 respect to the Zariski topology.*
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 $\varphi(V_2) = W_2$:*

$$\begin{tikzcd}
V_1 \arrow[d, dash] & \supset {}^{\exists} V_2 \arrow[d, dotted, dash] \\
W_1 & \supset W_2
\end{tikzcd}$$

Proof. The assumption on φ is that φ^* defines 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 $\mathfrak{P}$:

$$\begin{tikzcd}
\mathfrak{P} \arrow[d, dash] & \subset {}^{\exists} \mathfrak{M} \arrow[d, dotted, dash] \\
\mathfrak{p} & \subset \mathfrak{m}
\end{tikzcd}$$

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 set $\mathrm{V}(xy^2 - y)$ in 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

$$J = \langle s - x,\ t^2 - 1 - y,\ t(t^2 - 1) - z\rangle,$$

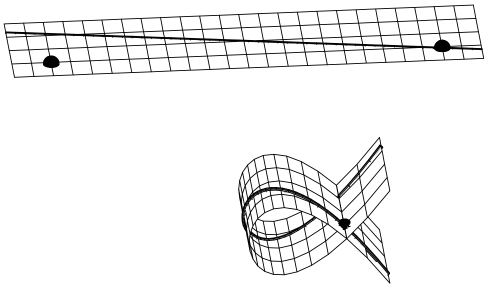
we get the polynomials

$$\begin{array}{lll} y^3 + y^2 - z^2, & tz - y^2 - y, & ty - z, \\ t^2 - y - 1, & s - x. & \end{array}$$

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

$$R = \Bbbk[x, y, z]/\langle z^2 - y^2(y + 1)\rangle \subset S = \Bbbk[s, t]$$

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

$$s^d + r_1 s^{d-1} + \dots r_d = 0$$

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 normalization:

1. The coordinate ring of the plane curve $\mathrm{V}(y^2 - x^{2k+1}) \subset \mathbb{A}^2$, where $k \geq 1$.
2. The coordinate ring of the Whitney umbrella $\mathrm{V}(x^2 - y^2 z) \subset \mathbb{A}^3$. □

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 $K = \mathrm{Q}(R)$ 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 $\operatorname{rad}(\mathfrak{p}\widetilde{R})$, where $\widetilde{R} \subset K$ is the normalization of R. Since $R = \widetilde{R}$ and $\operatorname{rad}\mathfrak{p} = \mathfrak{p}$ 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 $\mathfrak{P}_1 \subset \mathfrak{P}_2$. □

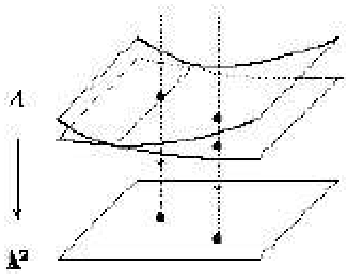
Remark 3.2.29. Even if R is a Noetherian integral domain, its normalization $\widetilde{R}$ need not be Noetherian. In particular, the extension $R \subset \widetilde{R}$ need not be finite (see Nagata (1962), Appendix A1. Examples of bad Noetherian rings). It is finite, however, if R is an affine domain. 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.



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.*

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

$$\Bbbk[x_{c+1}, \dots, x_n] \subset \Bbbk[x_c, \dots, x_n]/I_{c-1} \subset \cdots \subset S$$

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 of Proposition 3.3.3 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 all 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 (apply 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. If $\mathbb{K}[y_1, \dots, y_d] \subset \mathbb{K}[x_1, \dots, x_n]/J$ is a Noether normalization, where $J = I\,\mathbb{K}[x_1, \dots, x_n]$, the composition

$$\phi : \mathbb{K}[y_1, \dots, y_d] \subset \mathbb{K}[x_1, \dots, x_n]/J \to \mathbb{K}[x_1, \dots, x_n]/(\operatorname{rad} J)$$

is injective and, thus, a Noether normalization as well. Indeed, otherwise, we could find a nonzero element $f \in \ker \phi$, and a suitable power of f would define a nontrivial $\mathbb{K}$-algebra relation on $y_1, \dots, y_d \in \mathbb{K}[x_1, \dots, x_n]/J$.

Step 3. If

$$\mathbb{K}[y_1, \dots, y_d] \subset \mathbb{K}[A]$$

is a Noether normalization, and $A = V_1 \cup \dots \cup V_s$ is the decomposition of A into its irreducible components, then each composition

$$\phi_i : \mathbb{K}[y_1, \dots, y_d] \subset \mathbb{K}[A] \to \mathbb{K}[V_i]$$

is either injective (and, thus, a Noether normalization) or the composition of the induced map $\mathbb{K}[y_1, \dots, y_d]/\ker \phi \to \mathbb{K}[V_i]$ 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 since, 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 $\Bbbk$-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. Let $I = \langle xz, yz \rangle$ be the monomial ideal defining the union of the xy-plane and the z-axis. Considering that $I \cap \Bbbk[x, y]$ is zero, we see that $\dim \mathrm{V}(I) = 2$. 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 eimination 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 $\Bbbk = \mathbb{K}$.

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 **??**. We will complete the proof of the theorem in Exercise **??**. □

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 $A \subset \mathbb{A}^n$ 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 equidimensional of dimension $n-1$ iff it is a hypersurface. □

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 $\mathbf{L}(f) = m m_i$, 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 $\Bbbk[x_1, \dots, x_n]$, and let $>$ be a global monomial order on $\Bbbk[x_1, \dots, x_n]$. 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, to the contrary, that 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 defined by the associated primes of I. The result in the general case follows, once more, from Remark 2.7.1 on Buchberger's algorithm and field extensions. Indeed, as we already know, we can use Gröbner bases to check whether $R \to S$ is a Noether normalization. If this is true, and if $m_1, \dots, m_k$ are the monomials considered in the proof of Proposition 3.3.11, the check whether the $\overline{m}_i$ are R-linearly independent amounts to computing that the elimination ideal

$\langle m_1, \ldots, m_k\rangle \cap I \cap \Bbbk[x_{c+1}, \ldots, x_n]$ 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 $\Bbbk$-algebra S **Cohen-Macaulay** if it admits a Noether normalization

$$R = \Bbbk[y_1, \ldots, y_d] \subset S$$

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 $\Bbbk[x_1, \ldots, x_n]$ 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 4.6.22, but we will not discuss this topic any further. □

Example 3.3.14. If I is the monomial ideal

$$I = \langle x_1^2, x_2^2, x_1x_2x_3\rangle \subset \Bbbk[x_1, x_2, x_3],$$

then $R = \Bbbk[x_3] \to S = \Bbbk[x_1, x_2, x_3]/I$ 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

$$\Bbbk[z] \subset \Bbbk[x, y, z]/\mathrm{I}(C) = \Bbbk[C]$$

is a Noether normalization such that $\Bbbk[C]$ is a free $\Bbbk[z]$-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 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 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 $I \subsetneq \Bbbk[x_1, \ldots, x_n]$ coincides with the dimension of its locus of zeros $\mathrm{V}(I) \subset \mathbb{A}^n$. Considering a Noether normalization for $\Bbbk[x_1, \ldots, x_n]/I$, and taking Proposition 3.4.2 above into account, this amounts to showing that the Krull dimension of a polynomial ring over $\Bbbk$ equals the number of its variables.

That the dimension of $\Bbbk[x_1, \ldots, x_n]$ 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 $\Bbbk$-algebra, and let $I \subsetneq S$ be an ideal. There exist integers $\delta \leq d$ and a Noether normalization*

$$\Bbbk[y_1, \ldots, y_d] \subset S$$

such that

$$I \cap \Bbbk[y_1, \ldots, y_d] = \langle y_1, \ldots, y_\delta \rangle.$$

Proof. Let $\Bbbk[x_1, \ldots, x_d] \subset S$ be any Noether normalization. Since the composition of two finite ring extensions is again finite, it is enough to find a Noether normalization $\Bbbk[y_1, \ldots, y_d] \subset \Bbbk[x_1, \ldots, x_d]$ such that $I \cap \Bbbk[y_1, \ldots, y_d] = \langle y_1, \ldots, y_\delta \rangle$ for some $\delta \leq d$. We may, thus, suppose that $S = \Bbbk[x_1, \ldots, x_d]$ 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. Indeed,

$$x_1^k + c_1 x_1^{k-1} + \ldots + c_k - y_1 = 0$$

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

$$\mathbb{k}[y_1, \ldots, y_d] \subset \mathbb{k}[y_1, x_2, \ldots, x_d] \subset \mathbb{k}[x_1, x_2, \ldots, x_d]$$

is a finite ring extension. 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. □

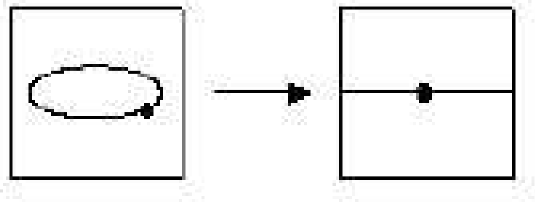
The geometric interpretation of the theorem is as follows: If $A \subset \mathbb{A}^n$ is an algebraic set, and $B \subset A$ is a subvariety, there is a surjective map $\pi : A \to \mathbb{A}^d$ with finite fibers which maps B onto a linear subspace of $\mathbb{A}^d$.

Exercise 3.4.4. If $I \subset S = \mathbb{k}[x_1, \ldots, x_4]$ is the ideal 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

$$\mathfrak{P}_0 \subsetneq \mathfrak{P}_1 \subsetneq \ldots \subsetneq \mathfrak{P}_m \tag{3.2}$$

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 with, $\mathfrak{P}_0 = \langle 0 \rangle$ since $\Bbbk[x_1,\dots,x_n]$ is an integral domain. Furthermore, $\mathfrak{P}_m$ is a maximal ideal (in 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 $\Bbbk[y_1, \ldots, y_d] \subset R$ be a Noether normalization. Then $\dim R = \dim \Bbbk[y_1, \ldots, y_d] = d$ by Proposition 3.4.2 and Theorem 3.4.6. Since also $\operatorname{trdeg}_{\Bbbk} \mathrm{Q}(R) = \operatorname{trdeg}_{\Bbbk} \Bbbk(y_1, \ldots, y_d) = d$, we must have $\dim R = \operatorname{trdeg}_{\Bbbk} \mathrm{Q}(R)$.

Write, now, R as the quotient of a polynomial ring $\Bbbk[x_1, \ldots, x_n]$ 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 $\Bbbk[x_1, \ldots, x_n]$ 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, $\operatorname{codim}_A B = \infty$.

In the example above, the codimension of the point p in $\mathrm{V}(xz, yz)$ is actually the codimension of p in the component $\mathrm{V}(x, y)$ containing p.

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). *If $\Bbbk$ is algebraically closed, and S is an affine domain over $\Bbbk$ with quotient field K, there are $y_1, \dots, y_d \in S$ such that:*

1. *$\Bbbk[y_1, \dots, y_d] \subset S$ is a Noether normalization.*
2. *$\Bbbk(y_1, \dots, y_d) \subset K$ is a separable field extension.*

Proof. In characteristic zero, every field extension is separable. We suppose, therefore, that char $\Bbbk = p > 0$.

Let $S = \Bbbk[x_1, \dots, x_n]/\mathfrak{p}$ for some prime ideal $\mathfrak{p}$ of some polynomial ring $\Bbbk[x_1, \dots, x_n]$. 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 $\Bbbk(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$, 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, $f \in \Bbbk[x_1, \dots, x_i^p, \dots, x_n]$ by Exercise 1.1.3.

Suppose that f is inseparable in each x_i. Then $f \in \Bbbk[x_1^p, \dots, x_n^p]$. Since $\Bbbk$ 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 $\Bbbk[x_1, \dots, x_n]$. That is, there is a polynomial $g \in \Bbbk[x_1, \dots, x_n]$ such that $g^p = f$:

$$f = \sum a_\alpha x_1^{p\alpha_1} \cdots x_n^{p\alpha_n} = (\sum b_\alpha x^\alpha)^p, \text{ where } b_\alpha^p = a_\alpha.$$

This contradicts the irreducibility of f.

So f is separable in at least one of the x_i, say in x_1. Choosing $a_2, \ldots, a_n \in \Bbbk$ sufficiently general as in the proof of Lemma 3.1.3, and expanding $f(x_1, \widetilde{x}_2 + a_2x_1, \ldots, \widetilde{x}_n + a_nx_1)$, we get a polynomial

$$ax_1^e + c_1(\widetilde{x}_2, \ldots, \widetilde{x}_n)x_1^{e-1} + \ldots + c_e(\widetilde{x}_2, \ldots, \widetilde{x}_n)$$

which provides both an integral equation for $\overline{x}_1 \in S$ over $\Bbbk[\widetilde{x}_2, \ldots, \widetilde{x}_n]$ and a separable equation for $\overline{x}_1 \in K$ over $\Bbbk(\widetilde{x}_2, \ldots, \widetilde{x}_n)$. 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). *If V is an affine variety 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. If $\Bbbk$ is a field of definition of V, the arguments of the previous proof, together with the arguments given in what follows, 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. We will, however, ignore these subtleties and simply suppose that $\Bbbk = \mathbb{K}$ is algebraically closed. Let $\overline{x}_1, \ldots, \overline{x}_n$ be generators for $S = \Bbbk[V]$ as a $\Bbbk$-algebra, and choose $y_1, \ldots, y_d \in \Bbbk[V]$ as in the proposition above. Then $\Bbbk[V]$ is a finite $\Bbbk[y_1, \ldots, y_d]$-algebra, and $K = \Bbbk(V)$ is a finite separable field extension of $\Bbbk(y_1, \ldots, y_d)$ which is generated by $\overline{x}_1, \ldots, \overline{x}_n$. By the primitive element theorem from Galois theory (see, for instance, Dummit and Foote (2003), Section 14.4), we can find a $\Bbbk(y_1, \ldots, y_d)$-linear combination y_{d+1} of the $\overline{x}_i$ such that $\Bbbk(V)$ is generated by y_{d+1} over $\Bbbk(y_1, \ldots, y_d)$. Clearing denominators, y_{d+1} can be taken as a $\Bbbk[y_1, \ldots, y_d]$-linear combination of the $\overline{x}_i$ and, thus, as an element of $\Bbbk[V]$.

If $f(y_1, \ldots, y_d, y_{d+1}) = 0$ is an integral equation for y_{d+1} over $\Bbbk[y_1, \ldots, y_d]$ of minimal degree, then f is an irreducible polynomial in $d + 1$ variables which, considered as a univariate polynomial with coefficients in $\Bbbk[y_1, \ldots, y_d]$, is the minimal polynomial of $\Bbbk(V)$ over $\Bbbk(y_1, \ldots, y_d)$. Hence, f defines an irreducible hypersurface $W \subset \mathbb{A}^{d+1}$, and the finite ring inclusion $\phi : \Bbbk[W] = \Bbbk[y_1, \ldots . y_d, y_{d+1}] \to \Bbbk[V]$ extends to an isomorphism $\Bbbk(W) \to \Bbbk(V)$ 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. □

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. A typical application of this lemma is Krull's intersection theorem (which we will treat in a special case).

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 **??**, 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.

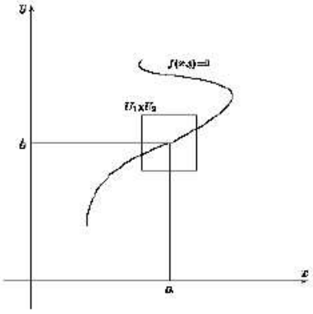
In the final section, we will treat the completion $\widehat{\mathcal{O}_{A,p}}$ of $\mathcal{O}_{A,p}$. This will help us to overcome a drawback of $\mathcal{O}_{A,p}$ which is due to the fact that Zariski open sets are rather large. Since $\mathcal{O}_{A,p}$ consists of (germs of) functions defined 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

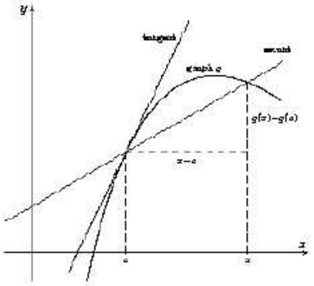
We will define smoothness such that in case $\mathbb{K} = \mathbb{C}$, an algebraic set $A \subset \mathbb{A}^n$ is smooth at a point $p \in A$ iff A is a complex submanifold of $\mathbb{A}^n$ in an Euclidean neighborhood of p. Equivalently, we will require that the hypothesis of the implicit function 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 $\big(\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p)\big)$ 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

$$d_p f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)(x_i - a_i) \in \mathbb{K}[x_1, \dots, x_n].$$

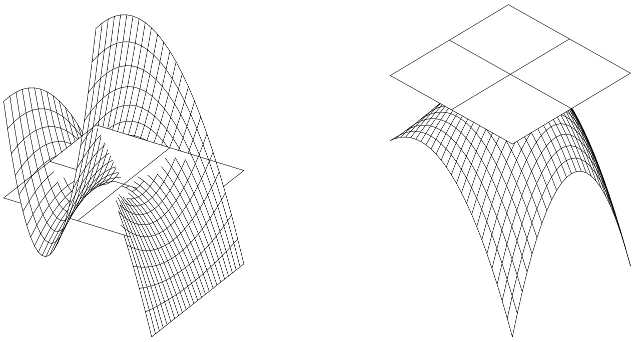
That is, $d_p f$ is the linear part of the Taylor expansion of f at p:

$$f = f(p) + d_p f + \text{terms of degree} \geq 2 \text{ in the } x_i - a_i.$$

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

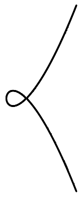
$$T_p A = \mathrm{V}(d_p f) \subset \mathbb{A}^n.$$

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.

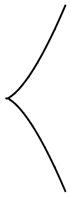


Otherwise, $T_p A = \mathbb{A}^n$, and we call p a **singular point** of A. □

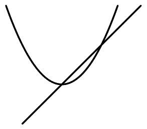
Example 4.1.2. The origin $o = (0,0) \in \mathbb{A}^2(\mathbb{C})$ is a singular point of each cubic curve shown below:



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



$y^2 = x^3$



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

□

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:*

$$A_{\text{sing}} = \mathrm{V}(f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}) \subsetneq A.$$

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 $A_{\text{sing}} = \mathrm{V}(f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})$ 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 $\operatorname{char} \mathbb{K} = 0$, this implies that f is constant, contradicting our assumption that A is a hypersurface. If $\operatorname{char} \mathbb{K} = p > 0$, 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. □

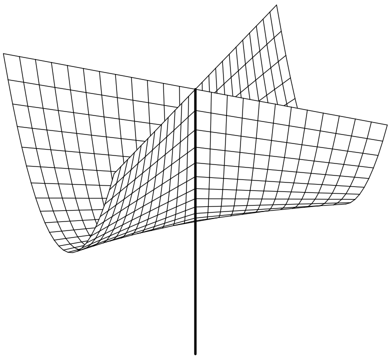
Example 4.1.4. The set of singular points of the Whitney umbrella

$$\mathrm{V}(x^2 - y^2 z) \subset \mathbb{A}^3(\mathbb{C})$$

is the z-axis

$$\mathrm{V}(x^2 - y^2 z, 2x, -2yz, -y^2) = \mathrm{V}(x, y).$$

We show a real picture:



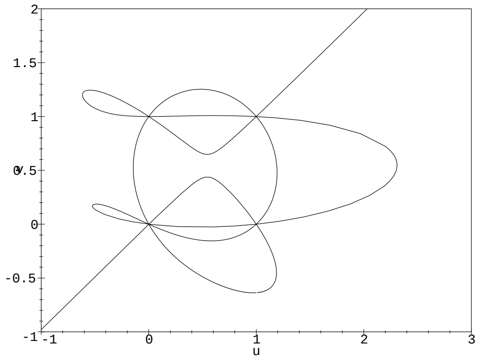
□

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 $\mathrm{I}(A) = \langle f_1, \dots, f_r\rangle$, then

$$T_pA = \mathrm{V}(d_pf_i \mid i = 1, \dots, r) \subset \mathbb{A}^n.$$

In particular,

$$\dim_{\mathbb{K}} T_pA = n - \operatorname{rank}\left(\frac{\partial f_i}{\partial x_j}(p)\right).$$

2. The function

$$A \to \mathbb{N},\ p \mapsto \dim T_pA,$$

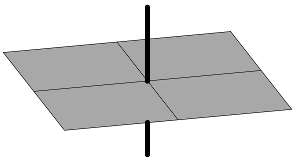
is upper semicontinous in the Zariski topology on A. That is, for any integer k, the subset

$$\{p \in A \mid \dim_{\mathbb{K}} T_pA \geq k\} \subset A$$

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 4.6.19 as a consequence of Krull's principal ideal theorem.

Definition 4.1.10. Let $A \subset \mathbb{A}^n$ be an algebraic set.

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 $A = V_1 \cup \cdots \cup V_s$ is the decomposition of A into its irreducible components, then

$$A_{\text{sing}} = \bigcup_{i \neq j} (V_i \cap V_j) \cup \bigcup_i (V_i)_{\text{sing}}$$

(this will be established in Corollary 4.7.8). 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. We will deduce this in Corollary 4.2.16 from the hypersurface case, making use of the formula in part 2 above and Theorem 3.5.2. □

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 $\dim_{\mathbb{K}} T_p A = n - \operatorname{rank}(\frac{\partial f_i}{\partial x_j}(p))$, 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 set, let $p \in A$ a point, and let $f_1, \dots, f_r \in \mathrm{I}(A)$. Then*

$$n - \operatorname{rank}(\frac{\partial f_i}{\partial x_j}(p)) \geq \dim_p A.$$

If equality holds, then A is smooth at p.

Proof. This follows from the chain of inequalities

$$n - \operatorname{rank}(\frac{\partial f_i}{\partial x_j}(p)) \geq \dim_{\mathbb{K}} T_p A \geq \dim_p A.$$

□

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 $I_{n-d}(\frac{\partial f_i}{\partial x_j})$ denote the ideal generated by the $(n-d) \times (n-d)$ minors of the Jacobian matrix of the f_i. If $I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I = \langle 1 \rangle$, 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 $\mathrm{V}(I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I) = \{p \in A \mid n - \operatorname{rank}(\frac{\partial f_i}{\partial x_j}(p)) > d\} \subset A$ is empty by the assumption on $I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I$ 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 $1 \notin I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I$:

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*

$$\dim \mathrm{V}(I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I) < \dim \mathrm{V}(I) = d,$$

then $I\,\mathbb{K}[x_1, \dots x_n] = \mathrm{I}(A)$ and $\mathrm{V}(I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I) = A_{\mathrm{sing}}$. In particular, I is a radical ideal.

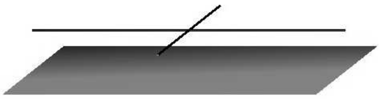
Proof. This will also be established towards the end of Section 4.6. □

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 $f_2 = x_1 x_2 x_3$. 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, $A = \mathrm{V}(x_1) \cup \mathrm{V}(x_1 - 1, x_2 x_3)$ 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

$$D = \begin{pmatrix} x_1 & x_2 & x_3^2 - 1 \\ x_2 & x_3 & x_1 x_2 + x_3 + 1 \\ x_3^2 - 1 & x_1 x_2 + x_3 + 1 & 0 \end{pmatrix}$$

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 $f_2 = x_1 x_3 - x_2^2$ 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. □

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 $d_pc = 0$ 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

$$d_p : \mathrm{I}_A(p) \to T_p^*A,\ \overline{f} \mapsto d_pf|T_pA.$$

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_pf). 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 $d_pf|T_pA = 0$. Then, if $f_1, \dots, f_r$ are generators for $\mathrm{I}(A)$, the differential d_pf is a $\mathbb{K}$-linear combination of the d_pf_i:

$$d_pf = \sum_{i=1}^{r} \lambda_i d_pf_i.$$

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

$$d_p\varphi : T_pA \cong (\mathrm{I}_A(p)/\mathrm{I}_A^2(p))^* \to (\mathrm{I}_B(q)/\mathrm{I}_B^2(q))^* \cong T_qB$$

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, now, 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 same principle was used in Section 2.6 to construct 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

$$\frac{r}{u} + \frac{r'}{u'} = \frac{ur' + u'r}{uu'} \text{ and } \frac{r}{u} \cdot \frac{r'}{u'} = \frac{rr'}{uu'}$$

(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. Show that localization commutes with passing to quotients by ideals: if R and U are as above, $I \subset R$ is an ideal, and $\overline{U}$ is the image of U in R/I, then

$$R[U^{-1}]/IR[U^{-1}] \cong (R/I)[\overline{U}^{-1}].$$ □

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 the 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

$$\mathbb{Z} \subset \mathbb{Z}[1/n] \subset \mathbb{Z}_{\langle p \rangle} \subset \mathbb{Q}.$$ □

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

$$R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}} \cong \mathrm{Q}(R/\mathfrak{p}).$$

□

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. Show that localization commutes with forming radicals: if $I \subset R$ is an ideal, then $\mathrm{rad}\,(IR[U^{-1}]) = (\mathrm{rad}\, I)R[U^{-1}]$. 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** $\mathrm{dom}(f)$ of a rational function $f \in \mathbb{K}(A)$ as in Section 2.6, and view f as a function on $\mathrm{dom}(f)$. Note that $\mathrm{dom}(f)$ 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 d_pf of f at 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*

$$\dim_{\mathbb{K}} \mathfrak{m}_{A,p}/\mathfrak{m}_{A,p}^2 = \dim \mathcal{O}_{A,p}. \qquad \square$$

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. By Remark 4.1.11, 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. 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. □

A few comments on Corollary 4.2.15 are in order. First, we should point out that the general algebraic form of inequality (4.2), which will be proved in Corollary 4.6.19, reads as follows: If $(R, \mathfrak{m})$ is a local Noetherian ring, then

$$\dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2 \geq \dim R. \qquad (4.3)$$

Second, 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 convenient characterization of $\dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2$ in terms of generators of $\mathfrak{m}$. This characterization 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 $h \equiv 1 \mod \mathfrak{m}$. □

Nakayama's lemma allows one to deduce information on modules over local rings from well-known facts on vector spaces. In making this explicit, we use the following notation: If R is any ring, and M is any R-module, a **minimal set of generators** for M is a set of generators for M such that 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 $\overline{m}_i = m_i + \mathfrak{m}M$, $i = 1, \dots, r$, 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 of $M/\mathfrak{m}M$, 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 $\operatorname{span}(\overline{m}_1, \dots, \overline{m}_r) = M/\mathfrak{m}M$. □

The first part of the exercise below shows that the conclusion of the corollary may be wrong over arbitrary rings:

Exercise 4.2.21. 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. □

We can, now, restate Definition 4.2.17 as follows: A local Noetherian ring $(R, \mathfrak{m})$ is **regular** if $\mathfrak{m}$ can be generated by $\dim R$ elements.

For later use, we present another application of Nakayama's lemma in a special case (see Eisenbud (1995), Corollary 5.4 for the general case):

Theorem 4.2.22 (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

$$S = R[\mathfrak{m}t] = R \oplus \mathfrak{m}t \oplus \mathfrak{m}^2 t^2 \oplus \ldots \subset R[t].$$

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 polynomials in $R[t]$ which can be chosen to be homogeneous in t. If r is the maximum degree in t 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 4.2.23. 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 by studying localization in a special case to which we will return in the next section:

Example 4.2.24. Let $I \subset \Bbbk[x_1, \ldots, x_n]$ be an ideal such that $\mathrm{V}(I) \subset \mathbb{A}^n$ consists of a single $\Bbbk$-rational point $p = (a_1, \ldots, a_n)$. Then

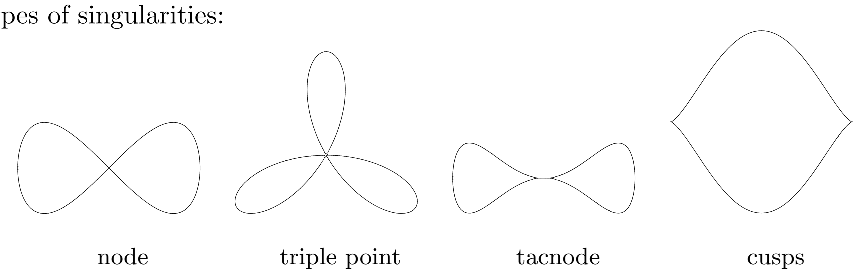
$$\mathcal{O}_p / I\mathcal{O}_p \cong \mathbb{K}[x_1, \ldots, x_n]/I\,\mathbb{K}[x_1, \ldots, x_n] =: R,$$

where $\mathcal{O}_p = \mathcal{O}_{\mathbb{A}^n, p}$ is the local ring of $\mathbb{A}^n$ at p. Indeed, already R has a unique maximal ideal, namely $\overline{\mathfrak{m}} = \langle x_1 - a_1, \ldots, x_n - a_n \rangle / I$. Hence, R is a local ring. By the universal property of localization, $R = R_{\overline{\mathfrak{m}}}$. But $R_{\overline{\mathfrak{m}}} \cong \mathcal{O}_p / I\mathcal{O}_p$ by Exercise 4.2.3. □

4.3 Intersection Multiplicities of Plane Curves

In Section **??**, 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 **??** 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:



□

Plane curves correspond to nonconstant square-free polynomials $f \in \mathbb{k}[x, y]$, 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 $f \in \mathbb{k}[x, y]$ be a nonconstant polynomial, and let $p = (a, b) \in \mathbb{A}^2$ be a point. Let

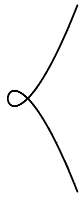
$$f = f_0 + f_1 + f_2 + \ldots + f_d \in \mathbb{K}[x, y]$$

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, $\operatorname{mult}(0, p) = \infty$.

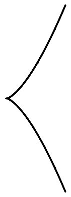
If f is square-free, and $C = \mathrm{V}(f) \subset \mathbb{A}^2$ is the corresponding curve, we write $\operatorname{mult}(C, p) = \operatorname{mult}(f, p)$, and call this number the **multiplicity of *C* at *p***. □

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 $m = 4$.

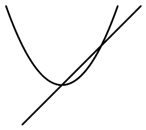
Example 4.3.3. The origin is a double point of each curve shown below:



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



$y^2 = x^3$



$y^2 = xy + x^2y - x^3$ □

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 siutation, 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 $\Longrightarrow$ 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 $\Longrightarrow$ 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 $\mathrm{Q}(R)$. □

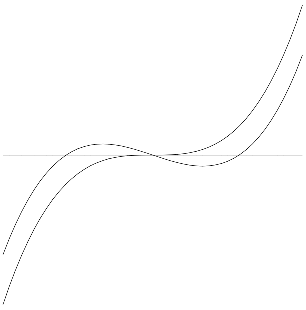
Taking Corollary 4.2.18 into account, we get:

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).

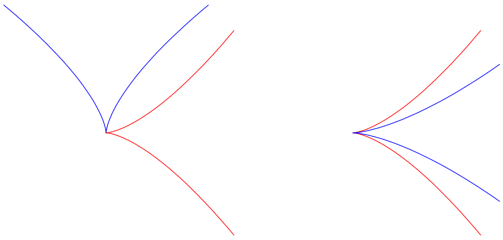
Example 4.3.13. Consider the curves $C = \mathrm{V}(y)$ and $D = \mathrm{V}(y-x^r)$ in $\mathbb{A}^2(\mathbb{C})$. Intuitively, we should count the origin $o = (0,0)$ as an intersection point of multiplicity r. Indeed, if we perturb the equations defining C and D slightly, we get r distinct intersection points near o:



the c ase $r = 3$

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 C and D at p**. □

Example 4.3.16. 1. In accordance with Example 4.3.13, we have

$$i(y, y - x^r; o) = r.$$

Indeed, by Example 4.2.24,

$$\mathcal{O}_{\mathbb{A}^2,o}/\langle y, y - x^r\rangle \mathcal{O}_{\mathbb{A}^2,o} \cong \mathbb{C}[x, y]/\langle y, y - x^r\rangle \cong \mathbb{C}[x]/\langle x^r\rangle.$$

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

$$\langle y^2 - x^3, x^2 - y^3\rangle = \langle y^2 - x^3, x^2 - x^3y\rangle = \langle y^2 - x^3, x^2\rangle = \langle y^2, x^2\rangle \subset \mathcal{O}_{\mathbb{A}^2,o},$$

and the result follows as above from Example 4.2.24. Similarly, for the tangential cusps,

$$i(y^2 - x^3, 2y^2 - x^3; o) = 6$$

since

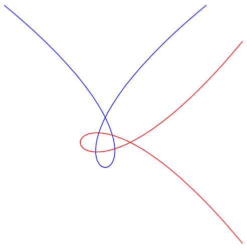
$$\langle y^2 - x^3, 2y^2 - x^3\rangle = \langle y^2, x^3\rangle \subset \mathcal{O}_{\mathbb{A}^2,o}.$$

To see this from the dynamical point of view, consider perturbed equations of type

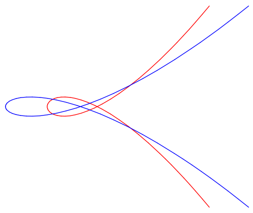
$$y^2 - (x - c)^2(x + c) = x^2 - (y - d)^2(y + d) = 0$$

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 $\operatorname{mult}(f,p)$ and $\operatorname{mult}(g,p)$ are ≥ 1). We say that f and g **intersect transversally at $\boldsymbol{p}$** if $\operatorname{mult}(f,p) = \operatorname{mult}(g,p) = 1$ 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 4.3.17 (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. *$i(f,g;p) \geq \operatorname{mult}(f,p) \cdot \operatorname{mult}(g,p)$, 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 4.4.20 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 Example 4.2.24, 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 lemmma 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,

$$\mathcal{O}_p/\langle f,g\rangle\mathcal{O}_p \cong \mathcal{O}_{C,p}/(\bar{g}) \cong \mathcal{O}_{C,p}/\langle t^k\rangle,$$

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

$$0 \to \mathcal{O}_p/\langle f,h\rangle\mathcal{O}_p \xrightarrow{\phi} \mathcal{O}_p/\langle f,gh\rangle\mathcal{O}_p \xrightarrow{\psi} \mathcal{O}_p/\langle f,g\rangle\mathcal{O}_p \to 0, \tag{4.4}$$

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 $bg \in (f,gh)\mathcal{O}_p$, 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, $b \in (f,h)\mathcal{O}_p$. 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 4.3.16. 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* 4.3.18. 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 $f = ay^2 + bx^3 +$ other terms of degree ≥ 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 are defined (in cases where the original point p is not rational over the originally given ground field, $\Bbbk$ will be an extension of that field). Taking into account that Remark 2.7.1 on field extensions applies to the modified 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 modified 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** $>_{\text{ldrlex}}$ on $\Bbbk[x_1, \dots, x_n]$ is defined by setting

$$x^\alpha >_{\text{ldrlex}} x^\beta \iff \begin{array}{l}\deg x^\alpha < \deg x^\beta, \text{ or } (\deg x^\alpha = \deg x^\beta \text{ and the} \\ \text{last nonzero entry of } \alpha - \beta \in \mathbb{Z}^n \text{ is negative}).\end{array}$$
□

A degree-anticompatible monomial order such as $>_{\rm ldrlex}$ is never global. It is, in fact, local in the following sense:

Definition 4.4.2. A monomial order on $\Bbbk[x_1, \ldots, x_n]$ is **local** if

$$x_i < 1 \quad \text{for} \quad i = 1, \ldots, n.$$

□

Example 4.4.3. A weight order $>_w$ on $\Bbbk[x_1, \ldots, x_n]$ is local iff the coefficients of w are strictly negative. □

Remark 4.4.4. Given a local monomial order $>$ on $\Bbbk[x_1, \ldots, x_n]$, a polynomial $u \in \Bbbk[x_1, \ldots, x_n]$ 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 $g = 1 \cdot f_1 + x^2$, $x^2 = x \cdot f_1 + x^3, \ldots$. 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 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 $\Bbbk[[x_1,\dots,x_n]]$ (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 $g \in \Bbbk[x_1,\dots,x_n]$ 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 $\Bbbk[[x_1,\dots,x_n]]$. 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 $\mathfrak{m}$. □

In what follows, we endow $\Bbbk[[x_1,\dots,x_n]]$ with the $\mathfrak{m}$-adic topology, where $\mathfrak{m}$ is the maximal ideal

$$\mathfrak{m} = \langle x_1,\dots,x_n\rangle \subset \Bbbk[[x_1,\dots,x_n]].$$

Accordingly, we say that a sequence $(f_\nu) \subset \Bbbk[[x_1,\dots,x_n]]$ 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** $(f_\nu) \subset \Bbbk[[x_1,\dots,x_n]]$ 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 $\nu \geq \nu_0$. By the Hausdorff property established in the first part of Proposition 4.4.7 below, the limes f is, then, uniquely determined, and we write $f = \lim_{\nu\to\infty} f_\nu$. A **series** $\sum_{\nu=0}^{\infty} f_\nu$ in $\Bbbk[[x_1,\dots,x_n]]$ is **convergent** and constitutes, thus, an element of $\Bbbk[[x_1,\dots,x_n]]$ if the sequence formed by its partial sums is convergent.

Proposition 4.4.7. *Let $\mathfrak{m} = \langle x_1,\dots,x_n\rangle \subset \Bbbk[[x_1,\dots,x_n]]$. Then:*

1. *The $\mathfrak{m}$-adic topology on $\Bbbk[[x_1,\dots,x_n]]$ is Hausdorff:*
$$\bigcap_{k=0}^{\infty}\mathfrak{m}^k = \langle 0\rangle.$$
2. *With respect to the $\mathfrak{m}$-adic topology, $\Bbbk[[x_1,\dots,x_n]]$ is* **complete**. *That is, every Cauchy sequences converges.*
3. *A series $\sum_{\nu=0}^{\infty} f_\nu$ in $\Bbbk[[x_1,\dots,x_n]]$ converges iff $\lim_{\nu\to\infty} f_\nu = 0$.*
4. *The ring $\Bbbk[[x_1,\dots,x_n]]$ is local with maximal ideal $\mathfrak{m}$.*

5. *There is a natural embedding of local rings $\mathcal{O}_o \subset \Bbbk[[x_1, \ldots, x_n]]$ defined by power series expansion. This embedding sends the maximal ideal of $\mathcal{O}_o$ into the maximal ideal of $\Bbbk[[x_1, \ldots, x_n]]$.*

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 $(f_\nu) = \left(\sum a_\alpha^{(\nu)} x^\alpha\right) \subset \Bbbk[[x_1, \ldots, x_n]]$, define $f = \sum a_\alpha x^\alpha \in \Bbbk[[x_1, \ldots, x_n]]$ 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: with respect to the $\mathfrak{m}$-adic topology, 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 $f \in \Bbbk[[x_1, \ldots, x_n]] \setminus \mathfrak{m}$ is a unit in $\Bbbk[[x_1, \ldots, x_n]]$. 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 $\Bbbk[[x_1, \ldots, x_n]]$. □

Next, we discuss division with remainder and Gröbner bases in $\Bbbk[[x_1, \ldots, x_n]]$. 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 $0 \neq f = \sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha \in \Bbbk[[x_1, \ldots, x_n]]$, we call any $a_\alpha x^\alpha$ with $a_\alpha \neq 0$ a **term** of f. And, given a local monomial order $>$ on $\Bbbk[x_1, \ldots, x_n] \subset \Bbbk[[x_1, \ldots, x_n]]$, 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 $\Bbbk[x_1, \ldots, x_n]$ 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 $\langle X \rangle \subset \Bbbk[x_1, \ldots, x_n]$. 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 4.4.8. *Let $\mathfrak{m} = \langle x_1, \ldots, x_n \rangle$ be the maximal ideal of $\Bbbk[[x_1, \ldots, x_n]]$, and let $>$ be a local monomial order on $\Bbbk[x_1, \ldots, x_n] \subset \Bbbk[[x_1, \ldots, x_n]]$.*

1. *If $(m_\nu)_{\nu \in \mathbb{N}}$ is a sequence of monomials in $\Bbbk[x_1, \ldots, x_n]$ 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 local weight order $>_w$, and $(f_\nu)_{\nu \in \mathbb{N}}$ is a sequence of formal power series in $\Bbbk[[x_1, \ldots, x_n]]$, then, with respect to the $\mathfrak{m}$-adic topology:*

$$\lim_{\nu\to\infty} \mathbf{L}_{>_w}(f_\nu) = 0 \implies \lim_{\nu\to\infty} f_\nu = 0$$ □

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 4.4.9 (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*

$$g = g_1f_1 + \ldots + g_rf_r + h, \text{ with } g_1,\dots,g_r,h \in R,$$

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 the terms of f_j except $\mathbf{L}(f_j)$, $j = 1,\dots,r$. 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 4.4.8. 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 4.4.8. □

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 4.4.10. *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 of g on Grauert division by the elements of 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, we have $h - \sum_{|\alpha|<k} b_\alpha x^\alpha \in \mathfrak{m}^k$, where $\mathfrak{m}$ is the maximal ideal of R. Hence, $g + I \equiv \sum_{|\alpha|<k} b_\alpha x^\alpha + I \mod \mathfrak{m}^k_{R/I}$, 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. □

Finally, based on Grauert division, we have a version of Buchberger's criterion for $\Bbbk[[x_1,\dots,x_n]]$ whose statement and proof read word for word identically to what we did in the polynomial case (we ask the reader to check this, see Exercise 4.4.11 below). 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.

Exercise 4.4.11. Let $R = \Bbbk[[x_1,\dots,x_n]]$.

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. □

We, now, turn from $\Bbbk[[x_1,\dots,x_n]]$ to $\mathcal{O}_o$.

Example 4.4.12. Giving an explicit example, we show that in $\mathcal{O}_o$, we cannot always achieve the strong condition (DD2) of Grauert's division theorem. For this, 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 $>_{\mathrm{ldrlex}}$). 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. □

We are, now, ready to discuss division with remainder and Gröbner bases in $\mathcal{O}_o$. Motivated by what we did in Example 4.4.5, and taking into account that every ideal in $\mathcal{O}_o$ can be generated by polynomials, our statements will be formulated such that they involve polynomial data only.

Theorem 4.4.13 (Mora's Division Theorem). *Let $>$ be a 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 $g \in \Bbbk[x_1, \dots, x_n]$, there exists an expression*

$$u \cdot g = g_1 f_1 + \ldots + g_r f_r + h,$$

where $u, g_1, \dots, g_r, h \in \Bbbk[x_1, \dots, x_n]$, with $\mathbf{L}(u) = 1$, 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 $f_1, \dots, f_r$, 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 4.4.14. Let $>$ be a monomial order on $\Bbbk[x_1, \dots, x_n]$. Given a nonzero polynomial $f \in \Bbbk[x_1, \dots, x_n]$, 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 4.4.15 (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 4.4.16. 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.

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 $I_1 \subset I_2 \dots$ 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

$$x_o^c x^\alpha >_g x_o^d x^\beta \iff \begin{array}{l} \deg x_o^c x^\alpha > \deg x_o^d x^\beta, \text{ or} \\ \deg x_o^c x^\alpha = \deg x_o^d x^\beta \text{ and } x^\alpha > x^\beta. \end{array}$$

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

$$u_\ell \cdot g = g_1^{(\ell)} f_1 + \ldots + g_r^{(\ell)} f_r + h_\ell, \quad \text{with } \mathbf{L}(u_\ell) = 1,$$

$\ell = 0, \ldots, k-1$. 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 $f_1, \ldots, f_r$, or
(b) $f^{(k)}$ is one of $h_0, \ldots, h_{k-1}$.

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 4.4.17. 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 4.4.18. Consider $>_{\mathrm{ldrlex}}$ 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. Let $>$ be a local monomial order on $\Bbbk[x_1, \ldots, x_n]$. Considering the embedding $\mathcal{O}_o \subset \Bbbk[[x_1, \ldots, x_n]]$, 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. **Leading ideals**, **standard monomials**, and **Gröbner bases** for ideals in $\mathcal{O}_o$ are defined as for ideals in $\Bbbk[x_1, \ldots, x_n]$, where we regard the leading ideal as an ideal of $\Bbbk[x_1, \ldots, x_n]$, and ask that the Gröbner basis elements are polynomials (we always may clear denominators).

Proposition 4.4.19. *Let $>$ be a local monomial order on $\Bbbk[x_1, \ldots, x_n]$. Then:*

1. *Let I is an ideal of $\mathcal{O}_o$, and let $f_1, \dots, f_r \in I$ be polynomials. Then the f_i 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 4.4.10 on standard monomials remains true if $\Bbbk[[x_1, \dots, x_n]]$ is replaced by $\mathcal{O}_o$.*

Proof. 1. The implication from right to left is clear since $I \subset I\, \Bbbk[[x_1, \dots, x_n]]$. Conversely, suppose that the f_i form a Gröbner basis for I. In particular, the f_i generate I and, thus, $I\, \Bbbk[[x_1, \dots, x_n]]$. Let $g = \sum_\alpha a_\alpha x^\alpha = \sum g_i f_i$ be any element of $I\, \Bbbk[[x_1, \dots, x_n]]$, and let $g^{(k)} = \sum_{|\alpha| \leq k} a_\alpha x^\alpha$ be obtained from g by truncating above degree k. If k is sufficiently large, then $\mathbf{L}(g^{(k)}) = \mathbf{L}(g)$ and the representation of g in terms of the f_i gives rise to a representation $g^{(k)} = \sum g_i^{(k)} f_i$, with polynomials $g_i^{(k)}$. Then $g^{(k)} \in I$, which implies by assumption that $\mathbf{L}(g) = \mathbf{L}(g^{(k)})$ is divisible by one of the $\mathbf{L}(f_i)$.

2. Given an ideal I of $\mathcal{O}_o$ and an element $g \in \mathcal{O}_o \subset \Bbbk[[x_1, \dots, x_n]]$, we find a power series $h = \sum_\alpha b_\alpha x^\alpha$ as in the proof Proposition 4.4.10 which, modulo $I + \mathfrak{m}_o^k$ can be replaced by the polynomial $\sum_{|\alpha| < k} b_\alpha x^\alpha$, where $\mathfrak{m}_o$ is the maximal ideal of $\mathcal{O}_o$. □

The second part of the proposition implies:

Remark 4.4.20. 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 4.3.17. □

Exercise* 4.4.21 (Multiplicities in Terms of the Local Ring). Let $f \in \Bbbk[x_1, \dots, x_n]$ be a nonconstant polynomial, let $p \in \mathbb{A}^n$ be a point, and let R be the local ring $R = \mathcal{O}_{\mathbb{A},p}/\langle f \rangle\, \mathcal{O}_{\mathbb{A},p}$ 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. □

Buchberger's Criterion for $\mathcal{O}_o$ is next. Taking Remark 2.3.18 into account, we formulate the criterion as follows:

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 $f_1, \dots, f_r \in \Bbbk[x_1, \dots, x_n]$ be nonzero polynomials. Let $X \subset \{\mathrm{S}(f_i, f_j) \mid j < i\}$ be a subset of S-polynomials such that the corresponding relations*

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

generate Syz $(\mathbf{L}(f_1), \dots, \mathbf{L}(f_r))$. *Then* $f_1, \dots, f_r$ *form a Gröbner basis iff any* $\mathrm{S}(f_i, f_j) \in X$ *has a Mora standard expression with remainder zero.*

Proof. The condition on the remainders is clearly necessary. It is also sufficient. Indeed, arguing as in the proof of Buchberger's criterion 2.3.9, this time using the syzygies arising from the Mora standard expressions with remainder zero, 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_i with remainder zero. Hence, $\mathbf{L}(g)$ is divisible by one of the $\mathbf{L}(f_i)$. □

Exercise 4.4.23. Consider $\Bbbk[x, y]$ with $>_{\mathrm{ldrlex}}$ and the ideals

$$I = \langle x^3 - y^3, x^2y^2 + xy^3 \rangle, \; J = \langle x^3 - y^3, x^2y^2 \rangle, \; \text{ and } \; K = \langle x^3 - y^4, x^2y^2 \rangle.$$

Compute that

$$\{x^3 - y^3, x^2y^2 + xy^3, xy^4 - y^5, y^6\}, \; \{x^3 - y^3, x^2y^2, y^5\}, \; \text{ and } \; \{x^3 - y^4, x^2y^2, y^6\}$$

are Gröbner bases for I, J, and K, respectively. In the proof below, we will make use of the ideals I, J, and K to illustrate the main arguments by examples. □

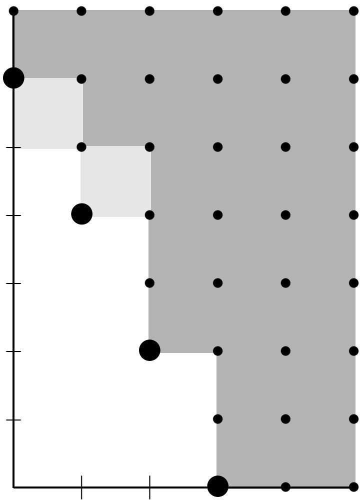
Proof of Theorem 4.3.17, Part 3. Let $f, g \in R = \Bbbk[x, y]$ be nonconstant polynomials, let $m = \mathrm{mult}(f, o)$ and $n = \mathrm{mult}(f, o)$ be their multiplicities at the origin o, and let f_m and g_n be the homogeneous summands of f and g of degrees m and n, respectively. We have to show that $i(f, g; p) \geq m \cdot n$, with equality occuring iff f and g have no tangent line in common at p. This is clear if $i(f, g; p) = \infty$. We assume, therefore, that $i(f, g; p) < \infty$. By part 2 of Theorem 4.3.17, the geometric meaning of this is that f and g do not have a common component passing through p.

Algebraically, writing $I = \langle f, g \rangle \subset \mathcal{O}_p$, our assumption is that $i(f, g; p) = \dim_{\Bbbk} \mathcal{O}/I\mathcal{O}_p < \infty$. Given any local monomial order on $\Bbbk[x, y]$, it follows from Remark **??** that $i(f, g; p)$ is precisely the number of standard monomials for I. To compute this number, we fix the local degree reverse lexicographic order. Then, since $>_{\mathrm{ldrlex}}$ is degree-anticompatible, $\mathbf{L}(f)$ 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 three cases.

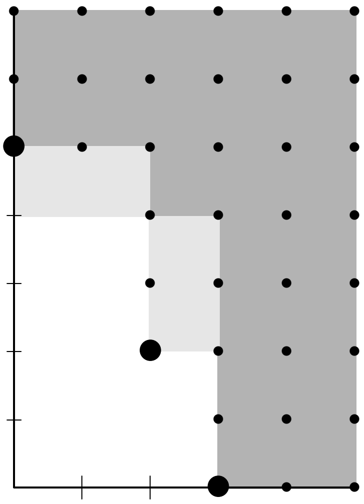
Case 1: f and g are homogeneous. That is, $f = f_m$ and $g = g_n$. Then f and g have no common tangent line at p since every such line would be a common component of f and g at p. This means: we have to show that the number of standard monomials for I is $m \cdot n$.

If $\beta_1 = 0$, we are done right away: f, g form a Gröbner basis for I, and the monomials $x^{\alpha_1}y^{\alpha_2}$ with $0 \leq \alpha_1 \leq m - 1$ and $0 \leq \alpha_2 \leq n - 1$ are precisely the standard monomials. Indeed, if $\beta_1 = 0$, then $y^n = \mathbf{L}(g) = g$, and the S-polynomial $S(g, f)$ is a multiple of g.

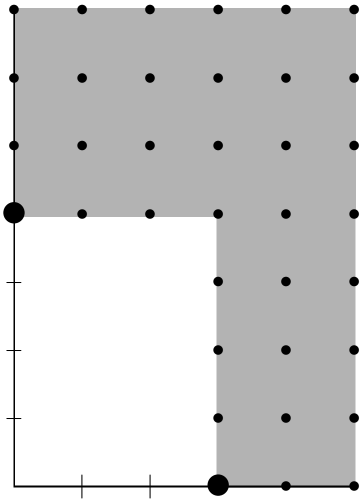
If $\beta_1 > 0$, however, then f, g do not form a Gröbner basis since this would imply that there are infinitely many standard monomials. Hence, the remainder of $S(g,f) = x^{(m-\beta_1)}g - y^{\beta_2}f$ on Mora division by f, g is nonzero and gives a new homogeneous Gröbner basis element f_3 for I 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, f_3, the only S-polynomial to check is $S(f_3, g)$ (indeed, $x^{m-\gamma_1}$ is divisible by $x^{\beta_1 - \gamma_1}$). Continuing in this way if necessary, we get Gröbner basis elements $f_1 = f, f_2 = g, f_3, \ldots$, where at each stage, the degree of the new generator f_{k+1} coincides with that of the S-polynomial $S(f_k, f_{k-1})$ leading to f_{k+1}. The process stops with an element f_r such that $\mathbf{L}(f_r)$ is a scalar times a power of y. If we visualize the monomials in $\Bbbk[x,y]$ by printing their exponent vectors as in Chapter 2, the leading monomials of the f_i determine a staircase which connects the x-axis with the y-axis. An elementary inductive argument shows that the area under the stair has size $m \cdot n$, as in case $\beta_1 = 0$.



4 Groebner basis elements



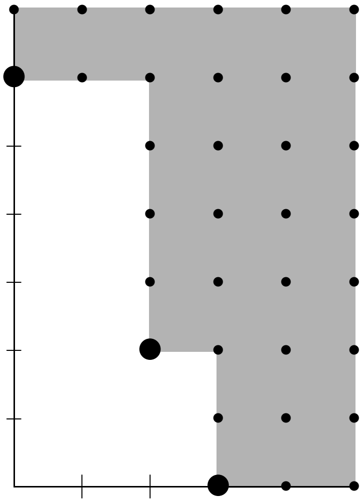
3 Groebner basis elements



2 Groebner basis elements

Case 2: f and g are nonhomogeneous, but do not have a common tangent line at p. Then, since $>_{\mathrm{ldrlex}}$ chooses the leading term of a polynomial from among the terms of lowest degree, the staircase obtained from f and g coincides with that obtained from f_m and g_n. (The only difference is that the Gröbner bases elements for $I = \langle f, g \rangle$ contain also terms of degree larger than the leading term).

Case 3: f and g are nonhomogeneous and do have a common tangent line at p. Then $\dim_{\Bbbk} \mathcal{O}_p/\langle f_m, g_n\rangle \mathcal{O}_p = \infty$. Hence, at least at one stage, the Gröbner basis computation yields a new element whose degree is larger than that of the monomial syzygy considered, and the area below the stair is strictly larger than $m \cdot n$.



This completes the proof of Theorem 4.3.17. □

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 radia $r_1, \ldots, r_n \in \mathbb{R}_{>0}$ such that the series $\sum_\alpha |f_\alpha| r_1^{\alpha_1} \cdots r_n^{\alpha_n}$ converges. In this case, f converges absolutely on the polydisc $\{|x_1| \leq r_1, \ldots, |x_n| \leq r_n\{$. 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 4.4.24. *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 $\mathbb{C}[[x_1, \ldots, x_n]]$ generated by convergent power series consists of convergent power series, too.* □

Exercise 4.4.25. Proof the proposition.

Let $>_w$ be local weight order on $\mathbb{C}[x_1, \ldots, x_n]$ given by $\mathbb{Q}$-linear negative weights such that $\mathbf{L}(g) = \mathbf{L}(g)$ and $\mathbf{L}(f_i) = \mathbf{L}(f_i)$. Let $r_1, \ldots, r_n \in \mathbb{R}_{>0}$ be radia such that $g, f_1, \ldots, f_r$ are convergent on the polydisc $\{|x_1| \leq r_1, \ldots, |x_n| \leq r_n\{$. Set $\widetilde{r_i} = \min(\log(-w_i), r_i$. Show, that the g_i (and, thus, h) converge in the polydisc defined by the $\widetilde{r_i}$. □

The rings $\Bbbk[[x_1, \ldots, x_n]]$ and $\mathbb{C}\{x_1, \ldots, x_n\}$. 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 4.4.26 (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

$$p = x_n^d + a_1(x_1, \dots, x_{n-1})x_n^{n-1} + \dots + a_d(x_1, \dots, x_{n-1}) \in \Bbbk[[x_1, \dots, x_{n-1}]][x_n] \text{ with } p(0, x_n) = x_n^d,$$

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 4.4.27. Complete the proof of the fact that $\Bbbk[[x_1, \dots, x_n]]$ is factorial. □

Exercise 4.4.28. 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 $\boldsymbol{M}$ at $\boldsymbol{U}$**.

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: 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*

$$I[U^{-1}] = \bigcap_{\mathfrak{q}_i \cap U = \emptyset} \mathfrak{q}_i[U^{-1}] \quad \textit{and} \quad \iota^{-1}(I[U^{-1}]) = \bigcap_{\mathfrak{q}_i \cap U = \emptyset} \mathfrak{q}_i$$

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 $\Longrightarrow$ 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.

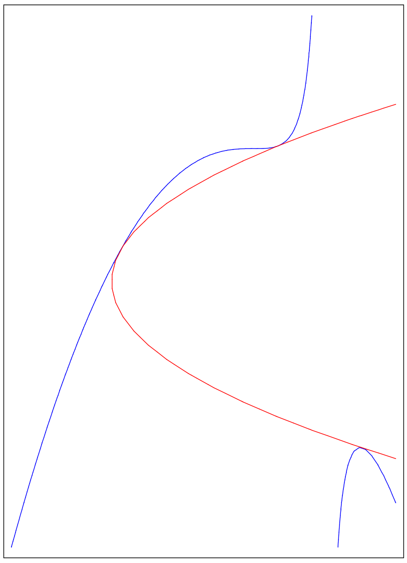
The surjectivity part follows in the same way. □

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 $\frac{x^3-6x^2+9x+1}{6-2x}$.



three 2-fold intersection points

Both curves are defined over $\mathbb{Q}$. Plugging in y for x^2 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 the picture above, 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

$$R/\langle \overline{y}^3 - 3\overline{y} + 1\rangle \cong \mathbb{Q}[y]/\langle y^3 - 3y + 1\rangle.$$

□

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+1}/M_i 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 $\ell > 2$, 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 $k - 1 \leq \ell - 1$. □

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 4.6.5. 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 1.4.4 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 4.6.6. *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. $1 \implies 2$: Iff $\ell(M) < \infty$, 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 $M_1 \subsetneq M$, which is Noetherian as well. Applying the same argument to M_1 and so forth, we get a descending chain $M = M_0 \supsetneq M_1 \supsetneq \ldots$ which, since M is Artinian, is eventually stationary. It is, hence, a composition series of M. □

Exercise* 4.6.7. 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* 4.6.8. 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. An affine $\mathbb{k}$-algebra $\mathbb{k}[x_1, \dots, x_n]/I$ is of finite length as a ring iff it has finite dimension as a $\mathbb{k}$-vector space. Geometrically, this is the case where the vanishing locus $\mathrm{V}(I) \subset \mathbb{A}^n$ consists of finitely many points.
3. The $\mathbb{k}[x]$-module $M = \mathbb{k}[x, x^{-1}]/\mathbb{k}[x]$ is Artinian, but not Noetherian.

Hint. For part 2, use Theorem 4.6.14 below. □

Definition 4.6.9. Let $f, g \in \mathbb{k}[x, y]$ be nonconstant polynomials, and let $\mathfrak{m}$ be a maximal ideal of $\mathbb{k}[x, y]$. The **intersection multiplicity of *f* and *g* at $\mathfrak{m}$**, written $i(f, g; \mathfrak{m})$, is defined to be

$$i(f, g; \mathfrak{m}) = \operatorname{length} \mathbb{k}[x, y]_{\mathfrak{m}}/\langle f, g\rangle \mathbb{k}[x, y]_{\mathfrak{m}}$$

□

Remark 4.6.10. 1. If $\mathrm{V}(f, g) < \infty$, then $i(f, g; \mathfrak{m})$ is the number of times in which the residue field $\mathbb{k}[x, y]/\mathfrak{m}$ occurs in a composition series of the $\mathbb{k}[x, y]$-module $\mathbb{k}[x, y]/\langle f, g\rangle$.

2. $i(f, g; \mathfrak{m}) = \infty \iff f$ and g have a common factor contained in $\mathfrak{m}$, $i(f, g; \mathfrak{m}) = 0 \iff$ one of the curves does not contain $\mathrm{V}(\mathfrak{m})$.
3. If the extension $\mathbb{k}[x, y]/\mathfrak{m}$ over $\mathbb{k}$ is seperabale, then $\mathrm{V}(\mathfrak{m})$ consist of an Galois orbit of $d = [\mathbb{k}[x, y]/\mathfrak{m} : \mathbb{k}]$ many points $p_1, \dots, p_d$ and
$$i(f, g; p_j) = i(f, g; \mathfrak{m})$$
in each of them.
4. If $\mathbb{k}[x, y]/\mathfrak{m}$ over $\mathbb{k}$ is not seperable, then $\mathrm{V}(\mathfrak{m})$ consist of $d = [\mathbb{k}[x, y]/\mathfrak{m} : \mathbb{k}]_{sep}$ many points, and the intersection multiplicity is
$$i(f, g; p_j) = i(f, g; \mathfrak{m})([\mathbb{k}[x, y]/\mathfrak{m} : \mathbb{k}]_{insep})$$
in each of them. □

Example 4.6.11. The ring

$$R = \mathbb{Q}[x, y]/\langle y^2 - x, x^3 - 6x^2 + 2xy + 9x - 6y + 1\rangle \cong \mathbb{Q}[y]/\langle y^3 - 3y + 1\rangle^2$$

considered in Example 4.6.1 has finite length since it has finite dimension as a $\mathbb{Q}$-vector space. In fact, $R \supsetneq \langle \overline{y}^3 - 3\overline{y} + 1) \supsetneq \langle 0\rangle$ is a composition series. Note that both factors are isomorphic to $L = \mathbb{Q}[y]/\langle y^3 - 3y + 1\rangle$. Taking part 3 of Remark 4.6.10 into account, this bbbbbbbbbbbbb

This implies that the curves intersect in three points 2-fold. □

Exercise 4.6.12. Let $I \subset \mathbb{k}[x_1, \dots, x_n]$ be an ideal with $\mathrm{V}(I) < \infty$, and let $\mathfrak{m} \subset \mathbb{k}[x_1, \dots, x_n]$ a maximal ideal. Express

$$\operatorname{length} \mathbb{k}[x_1, \dots, x_n]_{\mathfrak{m}}/I\mathbb{k}[x_1, \dots, x_n]_{\mathfrak{m}}$$

in terms of the sequence $\dim_{\mathbb{k}} \mathbb{k}[x_1, \dots, x_n]/I_k$, where $I_0 = I$ and $I_k = I_{k-1} : \mathfrak{m}$. □

Exercise 4.6.13. 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 4.6.14. *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

$$\Gamma := \{I \subset R \ \text{ ideal } \mid R/I \ \text{ is not of finite length}\}$$

is nonempty since $\langle 0\rangle \in \Gamma$. Hence, since R is Noetherian, Γ 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 $\mathfrak{p} + \langle f\rangle \supsetneq \mathfrak{p}$, the module $R/(\mathfrak{p} + \langle f\rangle)$ must have finite length by the maximality of $\mathfrak{p}$ as an element of Γ. 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 4.6.7, 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, wie 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 $f \in \mathfrak{p}_2/\mathfrak{p}_1 \subsetneq R/\mathfrak{p}_1$ 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

$$\operatorname{rad}\langle 0\rangle = \mathfrak{m}_1 \cap \ldots \cap \mathfrak{m}_s. \tag{4.10}$$

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 (4.10). This contradiction proves that $I = \langle 0\rangle$, as claimed.

Step 4. Each of the successive quotients in the descending chain of ideals

$$R \supset \mathfrak{m}_1 \supset \ldots \supset \mathfrak{m}_1^N \supset \mathfrak{m}_1^N\mathfrak{m}_2 \supset \ldots \supset \prod_{i=1}^{s} \mathfrak{m}_i^N = \langle 0\rangle \tag{4.11}$$

is a vector space over some field $R/\mathfrak{m}_i$. Hence, taking part 1 of Exercise 4.6.7 and part 1 of Exercise 4.6.8 into account, we get the following chain of eqivalences: R is Artinian $\iff$ each quotient in (4.11) is Artinian $\iff$ each quotient in (4.11) is Noetherian $\iff$ R is Noetherian. This concludes the proof. □

Next, we establish a structure result for Artinian rings. Then, following Krull, we will apply Theorem 4.6.14 above to prove the principal ideal theorem which is fundamental to the dimension theory of Noetherian rings.

Theorem 4.6.15 (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*

$$R \cong \bigoplus_{i=1}^{s} R_{\mathfrak{m}_i}$$

is a finite direct sum 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 1.5.10). Hence, the natural map

$$R \to \bigoplus_{i=1}^{s} R/\mathfrak{m}_i^N \tag{4.12}$$

is an isomorphism by the Chinese remainder theorem (see Exercise 1.3.9). To conclude the proof, we localize both sides of (4.12) 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 Example 4.2.24:

Corollary 4.6.16. *Let $I \subset \Bbbk[x_1, \dots, x_n]$ be an ideal such that $\mathrm{V}(I) \subset \mathbb{A}^n$ is finite, say $\mathrm{V}(I) = \{p_1, \dots, p_s\}$. Then*

$$\mathbb{K}[x_1, \dots, x_n]/I\,\mathbb{K}[x_1, \dots, x_n] \cong \bigoplus_{i=1}^{s} \mathcal{O}_{p_i}/I\mathcal{O}_{p_i}.$$

□

Theorem 4.6.17 (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 4.6.14. 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 codim $\mathfrak{p} \geq 1$. □

Theorem 4.6.18 (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*

$$\text{codim } \mathfrak{p} \leq c.$$

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,

$$f_i^N = a_i f_c + g_i \text{ with } g_i \in \mathfrak{q} \text{ snd } a_i \in R,\ i = 1, \dots, c-1.$$

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 (4.2) in its general form (4.3):

Corollary 4.6.19. *Let $(R, \mathfrak{m})$ be a local Noetherian ring. Then*

$$\dim R = \min\{d \mid \textit{ there exists an } \mathfrak{m}\textit{-primary ideal } \langle y_1, \ldots, y_d \rangle\}. \tag{4.13}$$

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 (4.13). 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 4.6.19 an important result of commutative algebra, where, in the situation of the corollary, a sequence of $d = \dim R$ elements $y_1, \ldots, y_d \in \mathfrak{m}$ 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 4.6.20. *The formal power series ring $\Bbbk[[x_1, \ldots, x_n]]$ is regular of dimension n. In fact, $x_1, \ldots, x_n$ form a regular system of parameters.*

Proof. Since $\Bbbk[[x_1, \ldots, x_n]]$ is an integral domain, $\dim \Bbbk[[x_1, \ldots, x_n]]/\langle x_n \rangle = \dim \Bbbk[[x_1, \ldots, x_n]] - 1$ by Krull's principal ideal theorem. On the other hand, $\Bbbk[[x_1, \ldots, x_n]]/\langle x_n \rangle \cong \Bbbk[[x_1, \ldots, x_{n-1}]]$. Hence,we conclude by induction on n that $\dim \Bbbk[[x_1, \ldots, x_n]] = n$. The result follows. □

The same argument shows that $\Bbbk[x_1, \ldots, x_n]_{\langle x_1, \ldots, x_n \rangle}$ is a regular local ring of dimension n. This is, of course, already clear from our discussion on smoothness.

Remark 4.6.21. Using induction on $\dim R$ and Nakayama's lemma, one can prove that every regular local ring $(R, \mathfrak{m})$ is an integral domain. This implies that if $y_1, \ldots, y_d$ is a regular system of parameters for R, then $y_1, \ldots, y_d$ is a **regular sequence** on R. That is, each y_i represents a nonzerodivisor of $R/\langle y_1, \ldots, y_{i-1} \rangle$, $i = 1, \ldots, d$. See Eisenbud (1995), Corollaries 10.14, 10.15 for details. □

For the sake of completeness, returning to Remark 3.3.13 on Cohen-Macaulay rings, we will, now, give the general definition of a Cohen-Macaulay ring. According to this definition and the remark above, every regular local ring is Cohen-Macaulay.

Definition 4.6.22. A local 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 ring is called **Cohen-Macaulay** iff its lcoalization $R_{\mathfrak{p}}$ is Cohen-Macaulay for every prime ideal $\mathfrak{p}$ of R. □

shows, in particular, that the local ring of an algebraic set at a smooth point is an integral domain. We will use this fact only in the special case treated directly in the following two corollaries:

Proposition 4.6.23. *Let $1 \leq r \leq n$, let $f_1, \dots, f_r \in \Bbbk[x_1, \dots, x_n]$ be polynomials vanishing at the origin $o \in \mathbb{A}^n$, and let $R = \mathcal{O}_{\mathbb{A}^n,o}/\langle f_1, \dots, f_r\rangle \mathcal{O}_{\mathbb{A}^n,o}$. Suppose that the matrix $M = \left(\frac{\partial f_i}{\partial x_j}(o)\right)_{1 \leq i,j \leq r}$ has maximal rank r. Then R is isomorphic to a subring of $\mathbb{K}[[x_{r+1}, \dots, x_n]]$. In particular, R is an integral domain.*

Proof. We write $M^{-1} = (a_{ij})$ and set $g_i = \sum_{j=1}^{r} a_{ij} f_j$, $i = 1, \dots r$. Then each g_i is of type $x_i +$terms of degree ≥ 2. In particular, by Buchberger's criterion, the g_i form a Gröbner basis for the ideal generated by the f_i in $\mathbb{K}[[x_1, \dots, x_n]]$ (fix any 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_i 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 4.6.24. *Let $A \subset \mathbb{A}^n$ be an algebraic set, let p be a point of A, and let $r = n - \dim_p A$. Suppose that we can find polynomials $f_1, \dots, f_r \in$ $\mathrm{id}(A)$ such that the matrix $M = \left(\frac{\partial f_i}{\partial x_j}(o)\right)_{1 \leq i,j \leq r}$ has maximal rank r. Then $\mathcal{O}_{A,p} \cong \mathcal{O}_{\mathbb{A},^n,p}/\langle f_1, \dots, f_{n-d}\rangle$ and this ring is a regular local ring.*

Proof. We have an epimorphism $\phi : R = \mathcal{O}_{\mathbb{A},^n,p}/\langle f_1, \dots, f_{n-d} \to \mathcal{O}_{A,p}$, so that $d = \dim \mathcal{O}_{A,p} R \leq R \leq \dim \mathfrak{m}_R/\mathfrak{m}_R^2 = d$ by assumption. Since R is a domain by the preceeding proposition, the kernel of ϕ must be zero since, otherwise, dim

by Krull's principal ideal theorem and the inequalitie.

Beweis vorziehen. □

Our final goal in this section is to show the corollaries to the Jacobian criterion:

Proof of Corollary 4.1.13, conclusion. Given an ideal $I = \langle f_1, \dots, f_r\rangle \subset \Bbbk[x_1, \dots, x_n]$ such that $A = \mathrm{V}(I) \subset \mathbb{A}^n$ is equidimensional of dimension d, we have to show that if

$$I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I = \langle 1\rangle, \tag{4.14}$$

then $I = \mathrm{I}(A)$. Here, $I_{n-d}(\frac{\partial f_i}{\partial x_j})$ is the ideal generated by the $(n-d) \times (n-d)$ minors of the Jacobian matrix of the f_i.

Let $\mathfrak{m} \subset \Bbbk[x_1, \dots, x_n]$ be any maximal ideal, and let $p \in \mathbb{A}^n$ be the corresponding point. Since $I \subset \mathrm{I}(A) \subset \Bbbk[x_1, \dots, x_n]$, also $I_\mathfrak{m} \subset \mathrm{I}(A)_\mathfrak{m} \subset \mathcal{O}_{\mathbb{A}^n,p}$ 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 $I_\mathfrak{m} = \mathrm{I}(A)_\mathfrak{m}$. For this, we distinguish two cases.

If $p \in \mathbb{A}^n \setminus A$, there is a polynomial $f \in I \subset \mathrm{I}(A)$ which is not contained in $\mathfrak{m}$. Then f is a unit in $\mathcal{O}_{\mathbb{A}^n,p}$, which implies that $I_\mathfrak{m} = \mathrm{I}(A)_\mathfrak{m} = \mathcal{O}_{\mathbb{A}^n,p}$.

If $p \in A$, then $I \subset \mathrm{I}(A) \subset \mathfrak{m}$. By assumption, there is at least one $(n-d) \times (n-d)$ minor of the Jacobian matrix $\left(\frac{\partial f_i}{\partial x_j}(p)\right)$ not vanishing at p, say $\det\left(\frac{\partial f_i}{\partial x_j}(p)\right)_{1 \le i,j \le n-d} \neq 0$. Then $d_p f_1, \dots, d_p f_{n-d}$ are $\Bbbk$-linearly independent. For the algebraic set $B = \mathrm{V}(f_1, \dots, f_{n-d}) \subset \mathbb{A}^n$, this implies that $\dim T_p B \le d$. On the other hand, $\dim_p B \ge d$ by the generalized principal ideal theorem. Corollary 4.6.19 applied to the local ring $R = \mathcal{O}_{\mathbb{A}^n,p}/\langle f_1, \dots, f_{n-d}\rangle$ shows that equality holds. In particular, R is a regular local ring and, hence, an integral domain by Exercise **??**.

Since $\mathcal{O}_{\mathbb{A}^n,p}/\mathrm{I}(A)_\mathfrak{m}$ is a quotient of R of the same dimension (by assumption), we conclude that $\langle f_1, \dots, f_{n-d}\rangle_\mathfrak{m} = I_\mathfrak{m} = \mathrm{I}(A)_\mathfrak{m}$. □

Proof of Corollary 4.1.14. Given an ideal $I = \langle f_1, \dots, f_r\rangle \subset \Bbbk[x_1, \dots, x_n]$ of dimension d, we, now, suppose that $\Bbbk[x_1, \dots, x_n]/I$ is Cohen-Macaulay (in particular, by the Unmixedness Theorem 3.3.12, A is equidimensional). We have to show: if

$$\dim \mathrm{V}(I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I) < \dim \mathrm{V}(I) = d,$$

then $I = \mathrm{I}(A)$ and $\mathrm{V}(I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I) = A_{\mathrm{sing}}$.

Arguing as in the previous proof, we see that the equality $\langle f_1, \dots, f_r\rangle_\mathfrak{m} = \mathrm{I}(A)_\mathfrak{m}$ holds for the maximal ideal $\mathfrak{m}$ of any point $p \notin B := \mathrm{V}(I_{n-d}(\frac{\partial f_i}{\partial x_j}) + I)$. Since $I = \langle f_1, \dots, f_r\rangle$ and $\mathrm{I}(A)$ are equidimensional and $\dim B < \dim A$ by assumption, $I = \mathrm{I}(A)$ and, hence, $B = A_{\mathrm{sing}}$. □

4.7 Analytic Type and Tangent Cone

So far, given an algebraic set A and a point $p = (a_1, \dots, a_n) \in A$, we have defined two invariants of A at p: the local ring $\mathcal{O}_{A,p}$ with its maximal ideal $\mathfrak{m}_{A,p}$, and the Zariski tangent space $T_pA \cong (\mathfrak{m}_{A,p}/\mathfrak{m}_{A,p}^2)^*$. Both invariants have their drawbacks. Whereas T_pA approximates A near a smooth point, it fails to do this if A is singular at p. In fact, in the latter 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. The local ring $\mathcal{O}_{A,p}$, on the other hand, is too fine at a singular point in that it distinguishes between algebraic sets which should be considered locally isomomorphic:

Example 4.7.1. Intuitively, the plane curves

$$C = \mathrm{V}(y^2 - x^2 - x^3) \subset \mathbb{A}^2 \quad \text{and} \quad D = \mathrm{V}(v^2 - u^2) \subset \mathbb{A}^2$$

should be regarded isomorphic near the origin o. Nevertheless, the local rings $\mathcal{O}_{C.o}$ and $\mathcal{O}_{D.o}$ are not isomorphic. Indeed, 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 components of D, the ring $\mathcal{O}_{D.o}$ contains zerodivisors: $(v-u)(v+u) = 0 \mod \langle v^2 - u^2 \rangle$. □

In this section, motivated by the problems just discussed, we will introduce two further invariants of A at p. These carry the same information as $\mathcal{O}_{A,p}$ respectively T_pA at a smooth point, but are better behaved than these at singular points.

Our first new invariant is the completion $\widehat{\mathcal{O}_{A,p}}$ of $\mathcal{O}_{A,p}$ with respect to the $\mathfrak{m}_{A,p}$-adic topology. Reconsidering Example 4.7.1, our intuitive understanding is that both curves C and D consist of two branches near the origin. The curve C, however, does not decompose in a *Zariski* neighborhood of the origin. In terms of functions, $y^2 - x^2 - x^3$ 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 arbitrary small Euclidean neighborhoods and allow convergent power series as functions on these:

$$y^2 - x^2 - x^3 = (y + x\sqrt{1+x}) \cdot (y - x\sqrt{1+x}),$$

and the expansion

$$\sqrt{1+x} = \sum_{k=0}^{\infty} \binom{k}{1/2} x^k$$

converges 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 analog to the Euclidean topology, and it is not meaningful to speak of convergent power series. We, may, however, consider the ring

$$\mathbb{K}[[x_1 - a_1, \dots, x_n - a_n]]/\mathrm{I}(A)\,\mathbb{K}[[x_1 - a_1, \dots, x_n - a_n]].$$

As we will show, this ring is nothing but the completion $\widehat{\mathcal{O}_{A,p}}$.

To begin with, we discuss the completion in a more general algebraic context. For our purposes, it is enough to consider a local Noetherian ring $(R, \mathfrak{m})$ together with the $\mathfrak{m}$-adic topology on R. In this situation, we call two Cauchy sequences (f_ν) and (g_ν) in R equivalent if the sequence of differences $(f_\nu - g_\nu)$ converges to zero. Then, the set $\widehat{R}$ of all equivalence classes carries a natural ring structure: if (f_ν) and (g_ν) 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). The ring $\widehat{R}$ is called the **completion** of R (with respect to the $\mathfrak{m}$-adic topology). 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 $\bigcap_{k=0}^{\infty} \mathfrak{m}^k$ which, in our situation, is zero by Krull's intersection theorem. Thus, ι is injective, and we may consider R as a subring of $\widehat{R}$. It is easy to see that $\widehat{R}$ is, again, a local ring, with maximal ideal $\mathfrak{m}\widehat{R}$. For the completion of $\mathcal{O}_{A,p}$, this is also clear from the result already announced above:

Proposition 4.7.2. *Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p = (a_1, \dots, a_n) \in A$ be a point. Then*

$$\widehat{\mathcal{O}_{A,p}} \cong \Bbbk[[x_1 - a_1, \dots, x_n - a_n]]/\mathrm{I}(A)\,\Bbbk[[x_1 - a_1, \dots, x_n - a_n]].$$

Proof. To ease our notation, we suppose that $p = o$ is the origin. To a given power series $\sum a_\alpha x^\alpha \in \Bbbk[[x_1, \dots, x_n]]$, we associate the Cauchy sequence $(f_\nu) = (\sum_{|\alpha| \le \nu} a_\alpha x^\alpha)$. This defines a map

$$\phi : \Bbbk[[x_1, \dots, x_n]] \to \widehat{\mathcal{O}_o}.$$

By division with remainder in $\mathcal{O}_{\mathbb{A}^n,p} \subset \Bbbk[[x_1, \dots, x_n]]$ every element of $\mathcal{O}_{A,p}/\mathfrak{m}_{A,p}^\nu$ is represented by a polynomial of degree $< \nu$ with terms not contained in $\mathbf{L}(\mathrm{I}(A))$ (with respect to a degree-anticompatible local order). Thus, the map above is surjective. Since every element mod $\mathrm{I}(A)\Bbbk[[x_1, \dots, x_n]]$ is uniquely represented by a power series with terms not in $\mathbf{L}(\mathrm{I}(A))$, the kernel of the map is $\mathrm{I}(A)\Bbbk[[x_1, \dots, x_n]]$. □

Definition 4.7.3. *Given algebraic sets A and B together with points $p \in A$ and $q \in B$, the pairs (A, p) and (B, q) are called* **analytically isomorphic** *if $\widehat{\mathcal{O}_{A,p}} \cong \widehat{\mathcal{O}_{B,q}}$.* □

Example 4.7.4. In Example 4.7.1, (C, o) and (D, o) are analytically isomorphic. Indeed, the substitution homomorphism

$$\Bbbk[[u, v]] \to \Bbbk[[x, y]]$$

defined by

$$u \mapsto x\sqrt{1+x} = x\sum_{k=0}^{\infty}\binom{k}{1/2}x^k,\ v \mapsto y$$

induces an isomorphism $\Bbbk[[u, v]] \to \Bbbk[[x, y]]$
$\widehat{\mathcal{O}_{A,p}} \cong \widehat{\mathcal{O}_{B,p}}$. □

In particular the analytic type is a coarser invariant than the local ring. It is finer than the tangent space, because, if $R = \widehat{\mathcal{O}_{A,p}}$ and $\mathfrak{m}$ the maximal ideal of R, then $\mathfrak{m}/\mathfrak{m}^2 \cong \mathfrak{m}_{A,p}/\mathfrak{m}_{A,p}^2$. Indead 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. The composition is surjective and has kernel is $\mathfrak{m}_{A,p}^2$.

Concerning dimension we have

Proposition 4.7.5. *Let $I \subset \Bbbk[x_1, \dots, x_n]$ be an ideal, and let $p = (a_1, \dots, a_n) \in \mathbb{A}^n$ be a $\Bbbk$-rational point. Then*

$$\dim \mathcal{O}_p/I\mathcal{O}_p = \dim \Bbbk[[x_1 - a_1, \dots, x_n - a_n]]/I\Bbbk[[x_1 - a_1, \dots, x_n - a_n]].$$

where $\mathcal{O}_p$ denotes the local ring of the origin $p \in \mathbb{A}^n$ and $I \subset \Bbbk[x_1, \dots, x_n]$. In particular, if $A \subset \mathbb{A}^n$ is an algebraic set, then

$$\dim \mathcal{O}_{A,p} = \dim \widehat{\mathcal{O}_{A,p}}.$$

Proof. $\dim \Bbbk[[x_1, \dots, x_n]] \geq n$, because

$$\langle 0\rangle \subsetneq \langle x_1\rangle \subsetneq \langle x_1, x_2\rangle \subsetneq \dots \subsetneq \langle x_1, x_2 \dots, x_n\rangle$$

is a chain of prime ideals of length n. By Corollary 4.6.19 $\dim \Bbbk[[x_1, \dots, x_n]] \leq n$, since the maximal ideal is generated by n elements. For the second statement we note that if $y_1, \dots, y_d$ is a system of parameters in $\widehat{\mathcal{O}_{A,p}}$ or $\mathcal{O}_{A,p}$ then we can truncate the powerseries expansion of the y_j's sufficiently high to obtain a system of represented by polynomials, which represent a system of parameters in both rings.

Exercise 4.7.6. Formulate and proof the analogous statements of Proposition 3.3.3, Theorem **??** and Corollary 3.3.12 in the local and complete case.

Proposition 4.7.7. *Let $A \subset \mathbb{A}^n$ be an algebraic set, and let $p \in A$ be a point.*

1. *If A is smooth at p, then*

$$\widehat{\mathcal{O}_{A,p}} \cong \Bbbk[[t_1, \dots, t_d]], \quad \textit{where}\ \ d = \dim_p A.$$

In particular, in this case, $\mathcal{O}_{A,p}$ is an integral domain.

2. More generally, if p *is arbitrary, we have*

$$\widehat{\mathcal{O}_{A,p}} \cong \Bbbk[[t_1, \dots, t_e]]/J, \quad \textit{where} \ \ e = \dim T_pA,$$

and where J *is an ideal of* $\Bbbk[[t_1, \dots, t_e]]$ *such that* $J \subset \langle t_1, \dots, t_e\rangle^2$.

Proof. 2. If $\mathrm{I}(A) = \langle f_1, \dots, f_r\rangle$, then $T_pA = \mathrm{V}(d_pf) \subset \mathbb{A}^n$. In conveniently chosen coordinates $x_1, \dots, x_n$ of $\mathbb{A}^n$, we may, hence, suppose that $d_pf_i = x_i$, for $i = 1 \dots, n-e$, and $f_i \in \mathrm{I}(p)^2$, for $i > n-e$. Sending t_i to x_i and composing with the canonical projection, we get a ring homomorphism

$$\varphi : \Bbbk[[t_1, \dots, t_e]] \to \Bbbk[[x_1, \dots, x_n]] \to \widehat{\mathcal{O}_{A,p}}.$$

is surjective by the Division Theorem 4.4.9 for formal power series, because $\mathbf{L}(f_j) = x_j$ for $j = 1, \dots, n-e$ with respect to $<_{lrx}$. The kernel $J = \ker\varphi \subset (t_1, \dots, t_e)^2$, because $f_{n-e+1}, \dots, f_r \in (x_1, \dots, x_n)^2$. This completes the second part.

1. For the first it remains to prove that any quotient $R = \Bbbk[t_1, \dots, t_d]]/J$ by a nonzero ideal J has dimension $\dim R < d$. This follows from Proposition 4.7.5, because a chain of prime ideals in R correspond to a chain of prime ideals $\mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \dots \subsetneq \mathfrak{p}_k$ in $\Bbbk[[t_1, \dots, t_d]]$ containing $\langle 0\rangle \subsetneq J$. So $k < d = \dim \Bbbk[[t_1, \dots, t_d]]$. □

We are, now, ready to prove part 2 of Remark 4.1.11:

Corollary 4.7.8. *Let* $A \subset \mathbb{A}^n$ *be an algebraic set. If* $A = V_1 \cup \dots \cup V_s$ *is the decomposition of* A *into its irreducible components, then*

$$A_{\text{sing}} = \bigcup_{i \neq j} (V_i \cap V_j) \cup \bigcup_i (V_i)_{\text{sing}}.$$

Proof. Let $p \in A$ be a smooth point of A. Then, since $\mathcal{O}_{A,p}$ is an integral domain, p lies in a unique component V_i of A and is a smooth point of that component. Thus, $A - A_{\text{sing}} \subset (\bigcup_i V_i \setminus (V_i)_{\text{sing}}) \setminus \bigcup_{i \neq j}$. The converse inclusion is clear. □

Remark 4.7.9. Because of this proposition $e = \dim_{\Bbbk} T_pA$ is frequently called the **embedding dimension** of (A, p). We say that $(A, p) \subset (\mathbb{A}^n, p)$ is **minimally embedded**, if $n = e$.

Remark 4.7.10. In case of we have $\Bbbk = \mathbb{C}$ as a ground field, and a pair of analytically isomorphic minimally embedded singularities $(A, p) \subset (\mathbb{A}^e, p)$ and $(B, q) \subset (\mathbb{A}^e, q)$ the isomorphism $\widehat{\mathcal{O}_{B,q}} \to \widehat{\mathcal{O}_{A,p}}$ might be induced by an e-tuple of convergent power series $(z_1, \dots, z_e)$. In this case there are neighbarhoods 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 $z(A \cap U) = B \cap V$.

We can now make the notion of a cusp and tacnode precise:

Definition 4.7.11. For *char*$\Bbbk \neq 2, 3$, a plane curve singularity analytically isomorphic to

$$\mathrm{V}(y^2 - x^2),\ \mathrm{V}(y^2 - x^3) \text{ or } \mathrm{V}(y^2 - x^4)$$

is called an **(ordinary) node**, **(ordinary) cusp** or an **(ordinary) tacnode** respectively.

Exercise 4.7.12. Prove, that every ordinary triple point over an algebraically closed field is analytically isomorphic to $\mathrm{V}(xy(x-y))$.

Our second new invariant of A at p is the tangent cone. 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 secants in calculus. Mimicing this construction of tangent lines if A is not necessarily smooth at p gives the tangent cone.

For simplicity, we suppose that $p = o = (0, \ldots, 0)$ is the origin. Consider

$$B = \{(a, t) \in \mathbb{A}^n \times \mathbb{A}^1 \mid at \in A\}.$$

B is reducible. One component is $B_1 = \mathbb{A}^n \times \{o\}$. Let B_2 be the union of the remaining components of B. Then the **tangent cone** of A at o is

$$TC_oA = B_1 \cap B_2 \subset \mathbb{A}^n \times \{o\} \cong \mathbb{A}^n$$

regarded as a subspace of $\mathbb{A}^n$. To obtain equations for the tangent cone we expand any $f \in I_A \subset \Bbbk[x_1, \ldots, x_n]$ in homogenous parts

$$f = f_m(x) + f_{m+1}(x) + \ldots + f_d(x)$$

with $f_j \in \Bbbk[x_1, \ldots, x_n]$ homogeneous of degree j and $f_m \neq 0$.

Proposition 4.7.13. *Let $o \in A \in \mathbb{A}^n$ be an algebraic set containing the origin, and let $I \subset \Bbbk[x_1, \ldots, x_n]$ be a defining ideal of A Then*

$$J = (\{f_m \mid f_m \text{ is the smallest degree part of an equation } f = f_m + \ldots + f_d \in I\})$$

defines the tangent cone TC_pA.

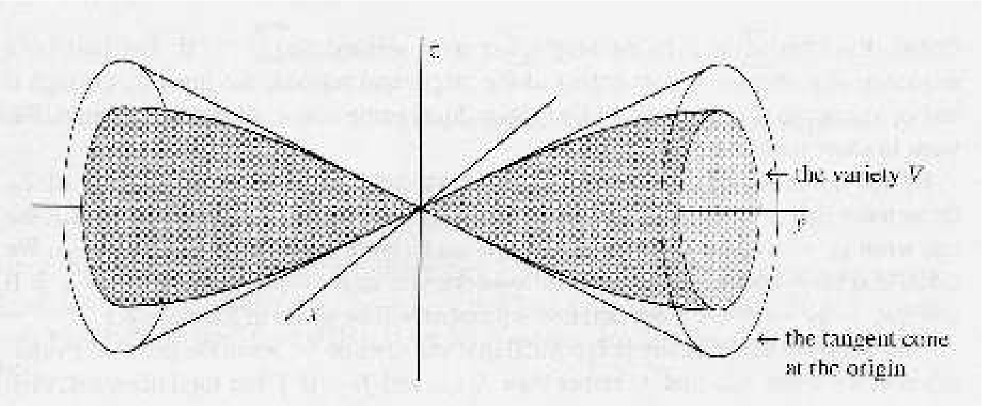
Proof. Defining equations of $B \subset \mathbb{A}^n \times \mathbb{A}^1$ are

$$f(tx) = t^m f_m(x) + t^{m+1} f_{m+1}(x) + \ldots + t^d f_d(x) = 0,$$

for $f \in I$. Since $o \in A$ we have $m \geq 1$ for all $f \in I$. The factor t^m vanishes on B_1 the remaining factor $f_m(x) + tf_{m+1}(x) + \ldots + t^{d-m} f_d(x)$ gives the defining equations for the union of the remaing components. Restricted to B_1 we obtain equations $f_m(x) = 0$ for $f \in I$ for the tangent cone.

In particular, if $A = \mathrm{V}(f) \subset \mathbb{A}^n(\Bbbk)$ is a hypersurface with $o \in A$, then $\mathrm{TC}_o\, A$ is defined by the vanishing of the lowest degree part of f.

Example 4.7.14. If $A = \mathrm{V}(x^2+y^2-z^2+z^4)$, then $\mathrm{TC}_o\, A = \mathrm{V}(x^2+y^2-z^2)$.



Remark 4.7.15. Note that $\mathrm{TC}_p\, A$ is a cone, since it is defined by homogenous equations. The linear equations in J define the tangent space T_pA.

A method to compute the tangent cone gives Mora algorithm: Given $I \subset \mathcal{O}_p$ generated by polynomials $f_1, \ldots, f_r$ compute a Groebner basis $f_1, \ldots, f_s$ of I with respect to a local monomial order, which refines multiplicity. Then the tangent cone is defined by

$$g_j = \sum_{|\alpha| = \mathrm{mult} f_j} f_{j,\alpha} x^\alpha \text{ for } j = 1, \ldots, s.$$

Remark 4.7.16. In more abstract terms the ring of the tangent cone can be defined as the graded ring

$$grR = R/\mathfrak{m} \oplus \mathfrak{m}/\mathfrak{m}^2 \oplus \mathfrak{m}^2/\mathfrak{m}^3 \oplus \ldots,$$

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 4.7.17. Let $X \subset \mathbb{A}^4$ be the algebraic set defined by the polynomials

$$\begin{gathered} x_2^3 - x_1^2x_3 + x_1x_2x_4 - x_1x_3x_4 - x_2x_4^2 - x_1x_2, \\ x_1x_2^2 - x_1x_3^2 + 2x_2x_3x_4 - x_3^2x_4 - x_2x_3, \\ x_1^3 - x_1x_2x_3 + x_2^2x_4 + x_1x_4^2 - x_4^3 - x_1x_4, \\ x_1^2x_3 - x_2x_3^2 + x_1x_2x_4 + 2x_3x_4^2 - x_3x_4. \end{gathered}$$

One can check that these equations are a local Gröbner basis. Thus the tangent cone at the origin is defined by

$$\langle x_1x_2, x_2x_3, x_1x_4, x_3x_4 \rangle = \langle x_1, x_3 \rangle \cap \langle x_2, x_4 \rangle,$$

two planes in $\mathbb{A}^4$ which intersect in a point.

Exercise 4.7.18. Check the assertion about the Gröbner basis in Example 4.7.17. Prove, that (X, o) and $(\mathrm{TC}_p\, X, o)$ are analytically isomorphic. A singularity of this type is called an **improper node** .

Remark 4.7.19. In general the tangent cone is not analytically isomorphic to the original singularity. For example the tangent cone of $\mathrm{V}(y^2 - x^3)$ is the double line $\mathrm{V}(y^2)$

Exercise 4.7.20. Prove, that over an algebraically closed field an ordinary quadrupel point is analytically isomorphic to one of the curves

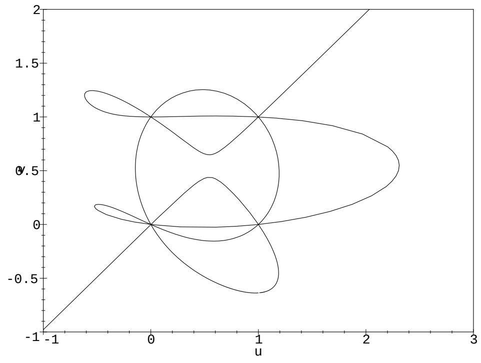
$$\mathrm{V}(xy(y-x)(y-\lambda x)), \text{ with } \lambda \in \mathbb{k} \setminus \{0, 1\}$$

and that two such curves for λ, λ' 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?



□

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